THE VIRTUAL MANUFACTURING STATION
A Framework for Collaborative Assessment of Manual Assembly Tasks

MICHAEL OTTO
aus Memmingen

DISSERTATION

zur Erlangung des Doktorgrades
- *Dr. rer. nat.* -
der Fakultät für Ingenieurwissenschaften, Informatik und Psychologie
der Universität Ulm

Institut für Medieninformatik
Fakultät für Ingenieurwissenschaften, Informatik und Psychologie
Universität Ulm

2020
Version 1.0

ABSTRACT

In the automotive industry, markets are demanding more product models, derivatives and extra equipment with shorter life-cycles. Due to these effects, planning of manual assembly is becoming more complex and diverse. With the current mostly physical mock-up production validation methods, these changes cause considerable increases in production planning costs, product preparation time and and put required quality levels at risk. The use of virtual assessment methods during the production validation phase is a promising countermeasure for these effects.

As of yet, there is no holistic view on virtual production validation in the literature since related publications either offer self-contained, practical approaches or theoretical constructs without direct applicability. In order to bridge this gap, this doctoral thesis focuses on the analysis, development, integration and evaluation of collaborative, virtual methods for assessments of manual assembly processes in the manufacturing industry.

This research focuses on the question whether collaborative virtual environments can support production validation workshops, so that verification criteria can be assessed in the same quality, less time and with lower costs compared to hardware-based workshops.

A new system is being developed and proposed, called the "Virtual Manufacturing Station" (VMS). It is a framework for holistic virtual production validation. The VMS consists of a multi-display environment, sensors and software components so that it can be used in interactive, collaborative, virtual production validation workshops. In order to provide production validation engineers with such a virtual framework, six theoretical key properties are derived for the VMS: "collaborative virtual environments", "multi-user support", "true to scale visualization", "natural user interfaces", "integration of physical and digital mock-ups" and "asymmetric/symmetric output." This theoretical framework is based on four research areas with each contributing to at least one of the theoretical key properties. These areas are "VR simulation software", "markerless, full-body motion capture", "large high-resolution displays" and "spatial augmented reality."

This doctoral thesis presents advances in basic human computer interaction research, technology, production validation methodology substantiated by the following studies: Two contextual inquiry studies on virtual production validation, two technological evaluations using a markerless full-body motion capture system presented, a systematic design space analysis for spatial augmented reality, a standardized benchmark for VR assessments of manual assembly tasks, a

size perception study, and five studies on basic research related to virtual production validation. The latter research studies cover a broad investigation scope, such as measurement of task completion times, error rates and qualitative feedback.

Overall, these studies have demonstrated that the VMS framework is reliable and applicable for collaborative virtual production validation workshops. Although this research has been conducted for the automotive sector, the presented VMS framework is also applicable to the manufacturing industry in general. The VMS methods and tools discussed contribute to higher workshop collaboration performance, lower task completion times, reduced preparation work and a reduced dependency on physical mock-ups. The VMS reduces the overall costs in production validation while simultaneously maintaining the validation quality.

PUBLICATIONS

Some ideas, texts and figures have appeared previously in the following list of directly related core [C] publications :

[C1] Michael Otto, Michael Prieur, and Enrico Rukzio. "Using Scalable, Interactive Floor Projection for Production Planning Scenario." In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*. ACM ITS 2014. Dresden, Germany: Association for Computing Machinery, Nov. 16, 2014, pp. 363–368. DOI: 10.1145/2669485.2669547.

[C2] Michael Otto, Philipp Agethen, Florian Geiselhart, and Enrico Rukzio. "Towards Ubiquitous Tracking: Presenting a Scalable, Markerless Tracking Approach Using Multiple Depth Cameras." In: *Proceedings of the EuroVR 2015*. EuroVR 2015. **Best Industrial Paper Award**. Lecco, Italy: European Association for Virtual Reality and Augmented Reality, 2015. URL: https://www.eurovr-association.org/wp-content/uploads/2019/11/06112017_Proceedings-1.pdf.

[C3] Florian Geiselhart, Michael Otto, and Enrico Rukzio. "On the Use of Multi-Depth-Camera Based Motion Tracking Systems in Production Planning Environments." In: *Proceedings of the 48th CIRP Conference on Manufacturing Systems*. CIRP CMS 2015. Vol. 41. CIRP CMS. Ischia, Italy: Elsevier, 2015, pp. 759–764. DOI: 10.1016/j.procir.2015.12.088.

[C4] Michael Otto, Michael Prieur, Philipp Agethen, and Enrico Rukzio. "Dual Reality for Production Verification Workshops: A Comprehensive Set of Virtual Methods." In: *Proceedings of the 6th CIRP Conference on Assembly Technologies and Systems*. CIRP CATS 2016. Vol. 44. CIRP CATS 2016. Gothenburg, Sweden, 2016, pp. 38–43. DOI: 10.1016/j.procir.2016.02.140.

[C5] Michael M. Otto, Philipp Agethen, Florian Geiselhart, Michael Rietzler, Felix Gaisbauer, and Enrico Rukzio. "Presenting a Holistic Framework for Scalable, Marker-Less Motion Capturing: Skeletal Tracking Performance Analysis, Sensor Fusion Algorithms and Usage in Automotive Industry." In: *Journal of Virtual Reality and Broadcasting, JVRB* 3.13 (2016), p. 10. DOI: 10.20385/1860-2037/13.2016.3.

[C6] Philipp Agethen, Felix Gaisbauer, Michael Otto, and Enrico Rukzio. "Interactive Simulation for Walk Path Planning within the Automotive Industry." In: *Proceedings of the 51st CIRP Conference on Manufacturing Systems*. CIRP CMS 2018. Vol. 72. Stock-

holm, Sweden, Jan. 1, 2018, pp. 285–290. DOI: 10.1016/j.procir.2018.03.223.

[C7]   Michael Otto, Eva Lampen, Philipp Agethen, Mareike Langohr, Gerald Masan, and Enrico Rukzio. "Evaluation on Perceived Sizes Using Large-Scale Augmented Floor Visualization Devices." In: *Proceedings of the 8th ACM International Symposium on Pervasive Displays*. ACM PerDis 2019. Palermo, Italy: Association for Computing Machinery, June 12, 2019, pp. 1–7. DOI: 10.1145/3321335.3324951.

[C8]   Michael Otto, Eva Lampen, Felix Auris, Felix Gaisbauer, and Enrico Rukzio. "Applicability Evaluation of Kinect for EAWS Ergonomic Assessments." In: *Proceedings of the 52nd CIRP Conference on Manufacturing Systems*. CIRP CMS 2019. Vol. 81. Ljubljana, Slovenia, 2019, pp. 781–784. DOI: 10.1016/j.procir.2019.03.194.

[C9]   Michael Otto, Eva Lampen, Philipp Agethen, Mareike Langohr, Gabriel Zachmann, and Enrico Rukzio. "A Virtual Reality Assembly Assessment Benchmark for Measuring VR Performance & Limitations." In: *Proceedings of the 52nd CIRP Conference on Manufacturing Systems*. CIRP CMS 2019. Vol. 81. Ljubljana, Slovenia, 2019, pp. 785–790. DOI: 10.1016/j.procir.2019.03.195.

[C10] Michael Otto, Eva Lampen, Philipp Agethen, Gabriel Zachmann, and Enrico Rukzio. "Using Large-Scale Augmented Floor Surfaces for Industrial Applications and Evaluation on Perceived Sizes." In: *Personal and ubiquitous computing, PAUC*. Theme Issue on Pervasive Displays tbd.2020 (2020), p. 28. DOI: 10.1007/s00779-020-01380-9.

Further [F] co-authored publications that are not directly related to the thesis' topic are:

[F1]   Philipp Agethen, Michael Otto, Felix Gaisbauer, and Enrico Rukzio. "Presenting a Novel Motion Capture-Based Approach for Walk Path Segmentation and Drift Analysis in Manual Assembly." In: *Proceedings of the Sixth International Conference on Changeable, Agile, Reconfigurable and Virtual Production*. CIRP CARV 2016. Vol. 52. CIRP CARV 2016. Jan. 1, 2016, pp. 286–291. DOI: 10.1016/j.procir.2016.07.048.

[F2]   Philipp Agethen, Michael Otto, Stefan Mengel, and Enrico Rukzio. "Using Marker-Less Motion Capture Systems for Walk Path Analysis in Paced Assembly Flow Lines." In: *Proceedings of the 6th CIRP Conference on Learning Factories*. CIRP CLF 2016. Vol. 54. 2016, pp. 152–157. DOI: 10.1016/j.procir.2016.04.125.

[F3]   Martin Manns, Michael Otto, and Markus Mauer. "Measuring Motion Capture Data Quality for Data Driven Human Motion Synthesis." In: *Proceedings of the 48th CIRP Conference on Manufacturing Systems*. CIRP CMS 2016. Vol. 41. Jan. 1, 2016, pp. 945–950. DOI: `10.1016/j.procir.2015.12.068`.

[F4]   Ulrich Bockholt, Thomas Bochtler, Volker Enderlein, Manuel Olbrich, Michael Otto, Michael Prieur, Richard Sauerbier, Roland Stechow, Andreas Wehe, and Hans-Joachim Wirsching. "Motion Capturing." In: *Web-basierte Anwendungen Virtueller Techniken*. Ed. by Werner Schreiber, Konrad Zürl, and Peter Zimmermann. Berlin, Heidelberg: Springer Vieweg, Berlin, Heidelberg, 2017, pp. 219–262. DOI: `10.1007/978-3-662-52956-0_5`.

[F5]   André Antakli, Pablo Alvarado Moya, Beat Brüderlin, Ulrich Canzler, Holger Dammertz, Volker Enderlein, Jürgen Grüninger, Andreas Harth, Hilko Hoffmann, Eduard Jundt, Peter Keitler, Felix Leif Keppmann, Roland Krzikalla, Sebastian Lampe, Alexander Löffler, Julian Meder, Michael Otto, Frieder Pankratz, Sebastian Pfützner, Matthias Roth, Richard Sauerbier, Werner Schreiber, Roland Stechow, Johannes Tümler, Christian Vogelgesang, Oliver Wasenmüller, Andreas Weinmann, Jochen Willneff, Hans-Joachim Wirsching, Ingo Zinnikus, and Konrad Zürl. "Virtuelle Techniken und Semantic-Web." In: *Web-basierte Anwendungen Virtueller Techniken*. Ed. by Werner Schreiber, Konrad Zürl, and Peter Zimmermann. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, pp. 17–116. DOI: `10.1007/978-3-662-52956-0_2`.

[F6]   Philipp Agethen, Viswa Subramanian Sekar, Felix Gaisbauer, Thies Pfeiffer, Michael Otto, and Enrico Rukzio. "Behavior Analysis of Human Locomotion in the Real World and Virtual Reality for the Manufacturing Industry." In: *ACM Transactions on Applied Perception* 15.3 (July 18, 2018), pp. 1–19. DOI: `10.1145/3230648`.

[F7]   Felix Gaisbauer, Jannes Lehwald, Philipp Agethen, Michael Otto, and Enrico Rukzio. "A Motion Reuse Framework for Accelerated Simulation of Manual Assembly Processes." In: *Proceedings of the 51st CIRP Conference on Manufacturing Systems*. CIRP CMS 2018. Vol. 72. Stockholm, Sweden, 2018, pp. 398–403. DOI: `10.1016/j.procir.2018.03.282`.

[F8]   Felix Gaisbauer, Philipp Agethen, Michael Otto, Thomas Bär, Julia Sues, and Enrico Rukzio. "Presenting a Modular Framework for a Holistic Simulation of Manual Assembly Tasks." In: *Proceedings of the 51st CIRP Conference on Manufacturing Systems*. CIRP CMS 2018. Vol. 72. Stockholm, Sweden, 2018, pp. 768–773. DOI: `10.1016/j.procir.2018.03.281`.

ACKNOWLEDGEMENTS

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AR   Augmented Reality

CAD   Computer Aided Design

CSCW   Computer-Supported Cooperative Work

CVE   Collaborative Virtual Environment

DHM   Digital Human Model

DMU   Digital mock-up

DoF   Degrees of Freedom

EAWS   Ergonomic Assessment Worksheet

HMD   Head-Mounted Display

HCI   Human Computer Interaction

ICP   Iterative Closest Point

JT   Jupiter Tesselation

KVM   Keyboard, Video and Mouse

LHRD   Large High-Resolution Displays

MDE   Multi-Display Environment

MoCap   Motion Capture

OEM   Original Equipment Manufacturer

OST   Optical See-Through

PDM   Product Data Management

PDP   Product Development Process

PLM   Product Lifecycle Management

PMU   Physical Mock-up

PMTS   Predetermined Motion Time System

PPR   Product, Process and Resource

PV   Production Validation

REST   Representational State Transfer

SAR   Spatial Augmented Reality

SDK   Software Development Kit

SOP   Start of Production

ToF   Time of Flight

UI   User Interface

VE   Virtual Environment

VMS   Virtual Manufacturing Station

VR   Virtual Reality

VRPN   Virtual-Reality Peripheral Network

VT   Virtual Technology

INTRODUCTION

The success of automotive Original Equipment Manufacturers (OEMs) depends on their ability to create customer-oriented products and services that can be delivered faster to customers than by their competitors. In the context of saturated markets, customer demands are continuously changing and increasing. Demand for electrified vehicles, mobility services and highly customizable products is increasing compared to the purchase of pre-defined products. Therefore, automotive manufacturers have to face these market demands and have to react faster to changes than the competition. Faced with disruptive changes in customer demands and digital transformation, production systems must have the flexibility to produce a wide range of products [1] such as diversified models, derivatives, extra equipment and features.

1.1 ECONOMIC IMPACT OF THE GLOBAL AUTOMOTIVE INDUSTRY

The automotive industry is a key driver of GDP growth and employment [2] in developed countries. Following OICA, the international organization of motor vehicle manufacturers, in 2019 67.14 million passenger cars and 24.63 million commercial vehicles were produced worldwide [3] compared to 39.76 million cars and 16.50 million commercial vehicles in 1999 [4]. In 2019, the top three passenger car producing countries were China (21.36 million), followed by Japan (8.32 million) and Germany (4.66 million) [3]. The average annual revenue of the "world automobile industry is more than 2.75 trillion Euros, which corresponds to 3.65% of the world GDP" [2]. Therefore, in these countries, the automotive industry has a huge share of the local GDP. For example, in Germany, the automotive industry has a share of 14% of the GDP and therefore holds a share of 6% of world production. In Germany, this industry has 807,000 direct employees and 1,800,000 indirect employees [2]. The automotive industry is a globalized market with significant value for the producing countries. For developed countries, they generate a substantial taxable base and revenues for state budgets. In a globalized world, all OEMs see themselves in an increasingly competitive market environment.

1.2 TECHNOLOGICAL CHANGES IN THE AUTOMOTIVE INDUSTRY

The digital transformation in the automotive industry leads to disruptive changes for both the products and their manufacturing processes. Wedeniwski describes "The Mobility Revolution in Automotive Industry" [5] and "how not to miss the digital turnpike." One of the key drivers is the ongoing "revolution in digitalization" and information technology. Not only is the product itself changing with these enabling technologies, but novel mobility concepts are also emerging. Faced with radical changes, large shares of the OEM's revenue are spent on the research and development of future products as described in the whitepaper entitled "Five trends transforming the Automotive Industry" by PWC. They summarize five main changes in the automotive industry: "electrified, autonomous, shared, connected and yearly updated" [6]. Almost all OEMs have started initiatives for these major disruptive changes, including Daimler AG's "CASE (connected, autonomous, shared, electric) strategy" [7]. These five disruptive changes are explained below:

Automakers are preparing to shift from building cars solely powered by internal combustion engines to **electric vehicles**, such as hybrid electric vehicles and battery electric vehicles. By 2025 this share is estimated to be 30% of all vehicle sales, compared to 1% in 2016 [8].

Autonomous driving is clustered in six levels ranging from Level 0 "driver support" to Level 5 "vehicle on demand." This taxonomy for driving automation systems is standardized by "SAE International Mobilus" in the document J3016B [9]. New application scenarios are enabled, such as completely driverless cars. Overall, autonomous driving requires multiple new components within the products, such as sensors, computing power and novel user interfaces [6, p. 20].

For **shared vehicles**, McKinsey has proposed to produce solution specific vehicles for each purpose, rather than offering a one-fits-all purpose vehicle. "The shift to shared mobility, enabling consumers to use the optimal solution for each purpose, will lead to new segments of specialized vehicles designed for very specific needs" [10], such as vacation, commuting, shopping, leisure and business vehicles.

Connected and yearly updated products also have implications on their production. Currently, in the automotive industry, the average expected life cycle of a product is seven years. Shorter time-to-market periods and yearly product updates are major change factors. This holds true for both the hardware and software of the products. **Connectedness** makes new business models feasible, such as over-the-air enabled features: Hardware parts are pre-installed in the products, even though the software feature is not enabled at the time of sale. The feature can be purchased after sales via an over-the-air update. For example, DAB+ radio is pre-installed in all cars, even though the

feature is not enabled by default and can be activated via a remote software update.

1.3 CHALLENGES IN AUTOMOTIVE PRODUCTION

All of these product change factors also directly influence automotive manufacturing. In order to stay competitive in market environments with saturation effects for automotive products, manufacturers have to "align their products and production with market demands" [11]. Diversification such as additional assembly parts, novel powertrains and extra equipment have to be integrated into the manufacturing system to produce these novel products with the aforementioned properties. As customers demand these novel features, more functions and regional adaptations, product variety increases [12]. Additionally, customers demand shorter product life cycles [11].

1.3.1 *Increasing product variety*

An ongoing trend to shorter life cycles and more highly individualized products can be observed [13]. Therefore, OEMs are continuously offering more car models, derivatives and variants. In the automotive industry, a model family consists of several models, such as sedan, wagon or convertible. Göpfert shows that the number of model variants has increased continuously over the past few years [14, p. 248]. For instance, in 1993, Mercedes-Benz offered nine main product variants, whereas in 2012 there were already 22 [15]. Overall, the number of car models in Germany has risen from 101 in 1990 to 453 in 2014 [16]. Along with the rising number of models, optional extra equipment for any given model has increased in a similar manner. A typical C-Class sedan offered 66 options in 1992, whereas in 2015 there were 211 options [17].

This growing product variety has a direct impact on all business units of an OEM such as research and development, production, logistics, brand, marketing, sales and after-sales [16]. Maropoulos & Ceglarek describe the increased efforts regarding the verification and validation of products in complex manufacturing systems [18]. The impact on the production system is one of the main reasons for carrying out the research in this doctoral thesis, since this impact results in a higher complexity in the production systems as well as increases in time and costs, which, in turn, has similar adverse effects on production planning departments.

1.3.2 *Mass customization*

This variety is the consequence of production for a diversified customer base, which demands low cost and high quality goods with

highly customized features. In contrast to mass production, mass customization is one promising approach to achieve this objective. In production systems, there is a continuous trade-off between productivity, quality, efficiency and costs [14, p. 249]. In enhancing the consumer's value through variety, the manufacturing industry in general must deal with increased product variety. It aims to achieve the overall efficiency of mass production while producing small batches of highly customized products.

In accordance with Koren [19], the development of predominant manufacturing principles in history is depicted in Figure 1.1. Before 1930, "craft production" was the prevalent production principle, which is represented by low volume per product variant. When reaching "mass production", significantly fewer variants are offered, but with high production volumes per variant. Henry Ford summarized this production principle in his famous quote: "Any customer can have a car painted any colour that he wants so long as it is black" [20]. So far, "mass customization" has increased the number of variants while only marginally reducing product volumes per variant. Regionalization, personalized production and other manufacturing paradigms are diversifying future manufacturing approaches.
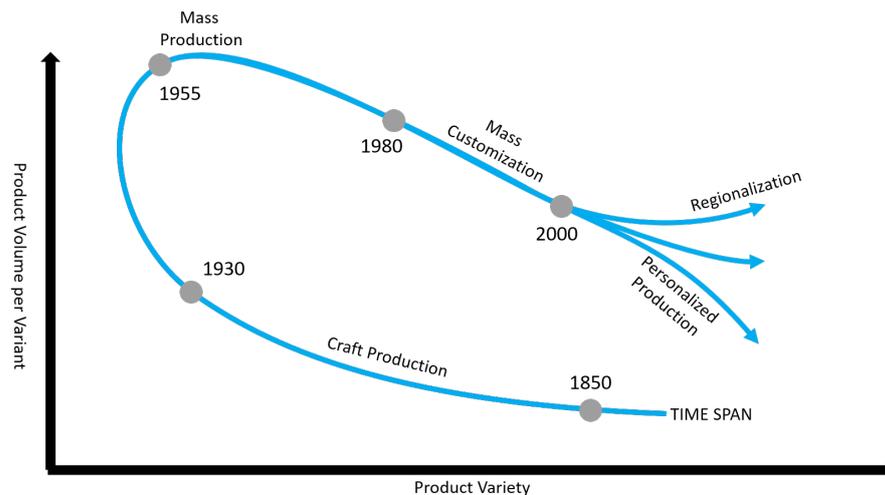
Figure 1.1: From craft production to mass customization (in accordance with "The global manufacturing revolution" by Koren [19])

ElMaraghy et al. show ways to manage product variety throughout the product life cycle. They discuss approaches for producing variety "efficiently including modularity, commonality and differentiation" [12]. This implies that large portions and multiple parts of the product are not varied and provide a common ground for assembling huge batch sizes [12]. Customer specific wishes are realized in the final assembly stage by the addition or removal of extra equipment for customized cars, models and derivatives.

1.3.3 *Need for flexibility*

As product features change, production systems must also be adapted so that they can produce these complex goods. Therefore, in addition to mass customization initiatives, assembly systems must become more flexible. Chryssolouris [21] names multiple aspects of flexibility: Machine, process, product, routing, volume, expansion, operation and production flexibility. The most important clusters for automakers are described below:

1. **Product flexibility:** "The ability to change over to produce new products economically and quickly."

2. **Operational flexibility:** "The ability to interchange ordering of several operations for each part type."

3. **Volume flexibility:** "The ability to operate profitably at different production volumes."

4. **Expansion flexibility:** " The ability to expand the system easily and in a modular fashion."

In general, a potential drop in demand is costly, time intensive and difficult. This is why flexibility strategies in production must be pursued: "Examples from automotive industry are proving that companies leveraging flexibility effects in their plants and an optimal capacity utilization are having a decisive competitive advantage" [14, p. 249].

1.3.4 *Global production networks*

Production facilities are spread all over the world so that they can produce different models in all of their plants anywhere in the world at the same time (compare Shimokawa et al. [22]). In a globalized production network, physically dispersed and geographically spread product ramp-ups must be dealt with. For example, automakers integrate a sedan variant of a new model family in a production line while still producing convertibles from the previous model family generation on the same production line. Such model-mix production systems allow flexible production but also require complex production planning methods. UGS describes the large "potential of assembling any product in their portfolio at any plant anywhere in the world, and to be able to change the production mix quickly while still maintaining high quality" [23].

All these partly disruptive, partly incremental changes put pressure on production planning departments to achieve these requirements, such as reduced time-to-market periods, more frequent production

ramp-ups, highly flexible production systems for mass-customized products and mixed-model production lines.

In their "Global Auto Executive Summary 2009" [24] KPMG found that the biggest (68% accordance) cost saving opportunity for OEMs lies in the domain of "manufacturing process and technology innovations." Overall, to ensure efficient and high quality production of products while having the same resources for planning, novel technologies and methods are required. "In order to have the most flexible and efficient production assembly capabilities, manufacturers must have the ability to continuously validate and optimize their manufacturing processes" [23].

1.4 MOTIVATION FOR VIRTUAL PRODUCTION VALIDATION

Physical assembly assessments are cost intensive and as such are a main cost driver. Therefore, while planning more products, models and options, costs rise accordingly. Nevertheless, products and processes still must be validated. Virtual assessments using digital mock-ups must fill this gap.

Virtual technologies and simulation approaches already partly support production validation processes for the manual final assembly stage. Since there are continuously fewer or even no physical prototypes available throughout the Product Development Process (PDP) (see Weber [25]), virtual assembly aims to offer similar capabilities for the assessment of verification tasks - just like in the physical domain. Even though virtual assembly has a long history in the literature on "digital factory" (see Gomes de Sa and Zachmann [26]), there are drawbacks. The shortcomings below are in accordance with the whitepaper presented by UGS [23] and Walla's doctoral thesis [27]:

- Virtual validation still lacks a **systematic process for the validation of all options**. Due to the high number of permutations in product variance, not all variants can currently be assessed by production planning.

- Virtual prototyping and assessments are also **cost intensive**, since authoring of the virtual environments requires a lot of manual effort. Therefore, only critical work tasks are validated in the virtual domain. Much previous knowledge is required in order to know which tasks could be critical. Batch assessment methods for assessing entire production lines are not available.

- Virtual assessments **lack interactivity** as there is no holistic framework for virtual validation of manual assembly tasks in either the literature or real-life applications. Advances in virtual technology are not immediately adopted by production planning.

- Production engineers cannot carry out holistic virtual assessments on their own due to the **complexity of authoring and simulation software handling**. Virtual assessments can only be carried out by digital factory experts with special knowledge.

- For efficient **data provisioning**, standardized data formats are still missing. Therefore, the interoperability between assessment tools is limited. Similarly, heterogeneous simulation environments are required for certain assessment aspects as they are highly focused on singular assessment scopes, i.e. ergonomic assessments. This requires additional training for virtual environment specialists.

- Interactive assessments often require **cumbersome preparation efforts**.

- Virtual assessment environments are **not optimized for collaborative assessments**.

- Required information is oftentimes either entirely **unavailable** in the virtual domain or is already **out-dated**. Some models lack realism and are too static.

- **Lack of access** to information due to restrictive data access policy or high costs for Product Data Management (PDM) systems.

- **Lack of simulation capabilities**, such as the simulation of flexible parts, holistic workstation visualization, rendering speeds for mass data visualization, interactivity, etc.

UGS summarizes the optimal scenario for production validation, having overcome all aforementioned limitations: "Optimize the design configurations of the building, tooling, carriers, material handling devices, operator walk path and more. Manufacturers can actually run a plant before they ever put a shovel in the ground to build it." [23]. All of the aforementioned deficiencies directly and indirectly have negative impacts on costs, time and product quality.

1.5 THESIS OUTLINE

This doctoral thesis is structured as depicted in Figure 1.2. The first four chapters include the motivation, research objectives, domain analysis and contextual inquiry study on the state of the art:

1. As presented above, Chapter 1 introduces the motivation for change in the automotive industry and the general shortcomings of virtual assembly assessment methods.

2. In Chapter 2, the research hypothesis of this doctoral thesis is formulated along with multiple research questions. Research
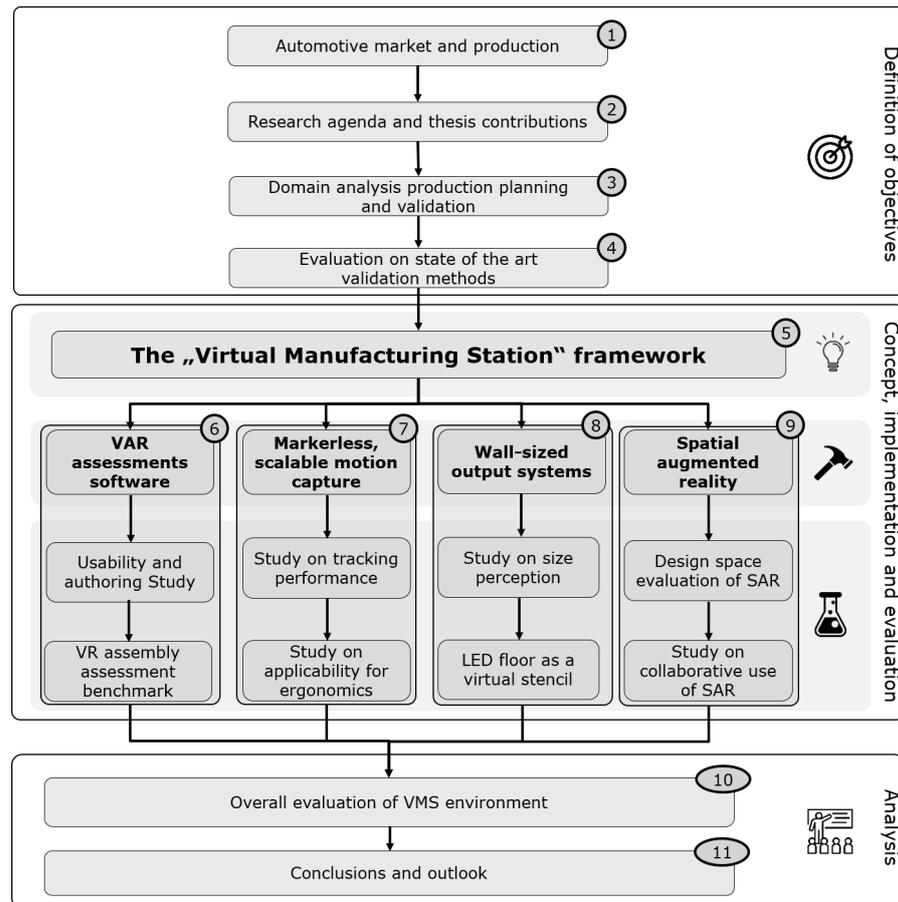
Figure 1.2: Outline and structure of this doctoral thesis. The numbers indicate the corresponding chapters in this document.

methods are summarized briefly. In addition, this research is compared with and delimited to other research topics.

3. Chapter 3 describes an in-depth domain analysis of production planning and validation. In order to elucidate virtual assembly assessments, the following topics are presented: Manufacturing principles of automotive production, final assembly characteristics, digital factory for manual final assembly and production validation workshops.

4. Chapter 4 presents a contextual inquiry study with subsequent expert interviews. The generated qualitative results underline the deficiencies of state-of-the-art production validation workshops.

Chapter 5 to Chapter 9 describe the implementations and research studies of the "Virtual Manufacturing Station":

5. The theoretical concepts of the "Virtual Manufacturing Station" framework are presented in Chapter 5. Objectives and key prop-

erties of the framework are described in the context of a literature review.

6. Chapter 6 shows the necessities of virtual batch production validation in simulation systems. An implementation of such a system is introduced. A subsequent research study focuses on the applicability of Virtual Reality (VR) assembly assessments. A novel research benchmark is proposed to determine the overall VR system's performance and limitations for assembly assessments.

7. Chapter 7 presents an implementation and evaluation of a markerless, scalable full-body motion capture system. An upstream evaluation provides insights into full-body tracking performance using Microsoft Kinect. The description of the implementation is followed by two studies on tracking performance and the applicability of the presented system for standardized ergonomic assessments.

8. Chapter 8 analyzes research on large-scale high-resolution displays and multi-display environments. Application scenarios, a prototype implementation, a large-scale LED implementation and two studies are presented subsequently. Furthermore, a basic research study presents generalized insights into size perception using augmented floor displays. A second evaluation describes an application-driven evaluation using large floor visualizations as a virtual stencil in cardboard workshops.

9. Chapter 9 presents research on projection-based Spatial Augmented Reality (SAR) in production validation using physical mock-ups. A literature review reveals gaps in industrial application scenarios for projection-based augmented reality, and a design space evaluation shows the practical limitations of this interface and compares it with optical see-through head-mounted Augmented Reality (AR) devices. A concluding research study quantifies the benefits using different types of computer-mediated communication in abstract collaboration tasks.

Having presented all implementations and research studies in the context of the Virtual Manufacturing Station (VMS), the final chapters summarize these works:

10. In accordance with Chapter 4, Chapter 10 also presents a contextual inquiry study. Production validation workshops using the final VMS framework and implementation are attended, evaluated and expert interviews carried out. This study qualitatively evaluates the overall performance of production validation workshops with respect to the planning results' quality, task completion time and overall costs.

11. Chapter 11 summarizes the outcomes of the VMS and picks up the research questions presented in Chapter 2. Finally, an outlook is provided on future developments in the interactive validation and automatic simulation of manual assembly tasks.

RESEARCH AGENDA AND THESIS CONTRIBUTIONS

This chapter presents the research agenda for this doctoral thesis, including the fields of contribution, research hypothesis, research questions, research methodology and a contrast and comparison with other research. This doctoral thesis is carried out in research cooperation between the Ulm University, Institute of Media Informatics, and Daimler AG.

2.1 FIELDS OF CONTRIBUTION

This thesis contributes to both fundamental research domains and applied sciences. Its main research area is "collaborative virtual environments for validation of manual assembly processes" affecting several research domains:

- Production Engineering

- Human Computer Interaction (HCI)

- Computer Supported Cooperative Work (CSCW)

As depicted in Figure 2.1, both fundamental research domains and fields of applied sciences interact. Contemporary scientific issues and latest research are applied to manufacturing industry use cases, and real automotive use cases are providing authentic work-related context to fundamental research.
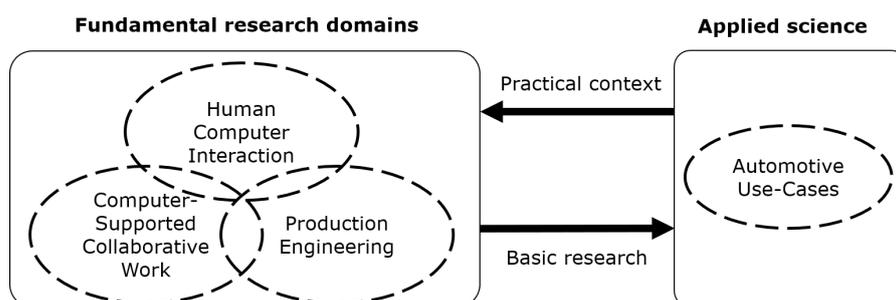
Figure 2.1: Fields of contribution and research context in the automotive industry

Through the analysis of deficiencies in **Human Computer Interaction (HCI) fundamental research**, this doctoral thesis closes multiple gaps in theoretical concepts for co-located, collaborative virtual environments. Key properties for collaborative virtual environments

are derived. Carrying out multiple empirical studies on size perception, collaboration performance and VR research, this doctoral thesis contributes to basic HCI research questions, utilizing state-of-the-art interaction technologies.

In the domain of **production engineering** and **Computer-Supported Cooperative Work (CSCW)**, this doctoral thesis proposes a framework of methods for the collaborative production validation of manual assembly tasks. Processes, technical optimizations, application scenarios and requirements are derived for the virtual validation of manual assembly tasks.

2.2 RESEARCH HYPOTHESIS AND RESEARCH QUESTIONS

This doctoral thesis aims to clarify fundamental questions in applied virtual production validation and to increase overall productivity in Production Validation (PV) workshops by presenting a framework of virtual, collaborative methods. In order to achieve the above-mentioned contributions and to overcome the existing deficiencies, a research hypothesis is formulated:

> **Research hypothesis**
>
> Utilizing collaborative virtual environments in production validation workshops for manual assembly tasks, verification criteria can be assessed in the same quality, less time and lower costs compared to hardware-based workshops.

This research hypothesis implies that providing production engineers with a specific set of virtual methods will have an impact on the overall verification task. Breaking down this hypothesis, three performance measures are analyzed, namely quality, time and costs.

The first performance measure **quality** can be measured directly by using objective error metrics, such as "achievement rates", "problem recognition rates" and "error amounts." For the second performance measure **time**, "task completion times" of individual or collective validations can be measured directly. **Costs** are analyzed as the change in efficiency on the basis of qualitative reports using "task completion times" and "event chains."

Besides the quantifiable performance measures, qualitative optimizations are sought: The goal is for each stakeholder to obtain a better understanding of the complex products, processes and resources by using the proposed framework. Therefore, production engineers are expected to increase their usage frequency, user experience and satisfaction with such virtual environments.

The research hypothesis is sub-divided into multiple concrete research questions. They are clustered with respect to corresponding research domains to which they contribute and the appearance se-

quence presented in this thesis. In the following sections, each research question is assigned to a dedicated chapter in this doctoral thesis that contributes to the respective research question:

Question 1 - Production Engineering

How is assembly validation presented in the literature and carried out in industrial practice? Which assessment criteria must be evaluated in the automotive production validation process?
Chapter 3 Domain Analysis

Question 2 - Production Engineering

Where are the deficiencies in current physical and virtual automotive verification processes, methods and tools? Which criteria can already be assessed in the virtual domain?
Chapter 4 Contextual inquiry study I

Question 3 - Human Computer Interaction & CSCW

Which requirements can be derived for a collaborative virtual assessment framework for the production validation of manual assembly tasks? What is the design space for a framework for virtual, mixed and car assemblies for batch assessments?
Chapter 5 VMS Framework

Question 4 - Human Computer Interaction

Which components are required in a VR batch assembly assessment simulation software and how can the performance and limitations of such a VR assembly assessment system be quantified?
Chapter 6 VR Assembly Assessment

Question 5 - Human Computer Interaction

How can a markerless, scalable tracking system be realized and what advantages of motion capture can be achieved? What are the limitations of markerless tracking systems and what tracking performance can be determined?
Chapter 7 Markerless Motion Capture

Question 6 - Human Computer Interaction

How do wall-sized displays and floor visualization displays influence spatial perception? Does the variation of interaction techniques have any influence on spatial perception and task performance?
Chapter 8 Large High-Resolution Displays (LHRD)s

> **Question 7 - Human Computer Interaction & CSCW**
>
> What is the design space for using projective spatial augmented reality in co-located Physical Mock-up (PMU) and Digital mock-up (DMU)-based environments and how does computer-mediated communication influence workshop performance in terms of task completion time and errors?
>
> **Chapter 9 Projection SAR**

> **Question 8 - Production Engineering**
>
> What implications are there regarding time, costs and quality compared to hardware based workshops? Can integrated virtual environments help production planning to achieve goals more quickly and reliably during the planning and validation of automotive manual end assembly?
>
> **Chapter 10 Contextual inquiry study II**

Therefore, the general shortcomings presented in Section 1.4 are addressed by these research questions.

2.3 RESEARCH METHODOLOGY

Since heterogeneous research areas are involved in this doctoral thesis, multiple research methodologies are applied as follows:

- An **exhaustive literature** review in the research areas of manufacturing and production engineering is carried out. This is followed by a literature review on digital factory and state-of-the-art virtual production validation for the automotive industry.

- In order to obtain insights into the context of use throughout production validation workshops, multiple **on-site attendances** permit insights in organization, process and methods. Silent attendance with a systematic observation of these processes helps to understand optimization potentials and to derive complex requirements.

- Following a user-centered design approach, the aforementioned observation methodology is combined with **semi-structured expert interviews** to form a **"contextual inquiry study"**. Semi-structured expert interviews are conducted to obtain qualitative insights and to understand the stakeholders' needs and their personal opinions. All interviews are recorded using audio recorders, subsequently transcribed, coded and thematically clustered to present the findings. Participants come from representative populations within the application domain, normally

customers for the expected system. Depending on the scope, production engineers, VR specialists and digital factory simulation specialists are chosen since not all production planners inherit leading-edge knowledge on virtual environments.

- **Quantitative evaluations** are carried out in multiple chapters. Objective measures are chosen for the respective research question. The presented HCI research studies are carried out under laboratory conditions as well as real-life environments. To generate quantifiable results, experiment leaders invite representative groups of people. In the studies discussed, "task completion times" and "error metrics" are measured to generate quantifiable results, such as spatial deviations, accuracy, precision and feasibility rates. When required, special tooling is applied such as industrial robots in Chapter 7 for highly reproducible trajectories to measure spatial accuracy and precision.

- For additional quantitative results, use is made of **standardized questionnaires**. Commonly used questionnaires for usability studies are the "System Usability Scale" (SUS) [28], "The Post-Study System Usability Questionnaire" (PSSUQ) (compare [29]) or for VR applications the "Presence Questionnaire" [30] are applied. If required and applicable, non-standardized questionnaires are used in addition to standardized questionnaires including tools such as paper-based questionnaires, Microsoft Info-Path, Microsoft Sharepoint or the LimeSurvey online survey tool.

2.4 SIMILARITIES AND DIFFERENCES

This work focuses on a framework for the collaborative virtual validation of manual assembly tasks. As multiple research domains influence this work, the focus of this doctoral thesis can be contrasted with other works as follows:

- The value creation chain in automotive production comprises several steps, such as body shop, paint shop and final assembly. The presented framework focuses on the application domain of **passenger car final assembly stage**. Spatial dimensions, processes and validation tasks are described for passenger car production only. Nevertheless, these concepts can be generalized and re-used for multiple manufacturing industries having a final assembly stage, such as commercial vehicle production, shipyards, aerospace and other final assembly stages of original equipment manufacturers.

- This work focuses on **manual assembly processes** only. Arteaga et al. state that "manual assembly processes comprise all assem-

bly related operations carried out by a human worker without the use of automatic machines to bring assembly parts onto a base part in order to create a final product. The area where assembly takes place includes the space required for equipment and workers, as well as the space required for the storage of components and finished products." [15]. Even though the final assembly stage has several dozens of automated processes without human labor, final assembly value creation primarily consists of manual human labor which is the only scope of the proposed framework. Therefore, virtual engineering and virtual commissioning (see [31, 32]) are not considered in this doctoral thesis. Human-robot collaboration is not considered either, even though this is a widely researched field.

- The **stakeholders** of the presented framework are limited to automotive production planning departments. These production engineers aim to optimize products, processes and resources for the final assembly stage [33]. By presenting novel virtual assessment methods, other departments besides production planning can use these methods in a similar manner, such as research and development, prototype building, training and maintenance departments. None of these departments are considered as native stakeholders in this doctoral thesis, even though the proposed methods could be transferred with slight changes. For instance, when validating virtual assembly in automotive production planning, after sales departments validate disassembly. As these application scenarios have large overlaps, these methods could be transferred. A generalization of the presented methods is discussed in the outlook in Section 11.3.

- The creation of a fully functional final assembly stage requires **auxiliary stakeholders**: Logistics and factory planning. Logistics departments are necessary for continuous material provisioning as they have to plan and deliver material to the final assembly workstations. These so-called "material zones" and "pre-assemblies" at the continuous flow line workstations are within the optimization scope of this doctoral thesis. In contrast, warehouse planning, material flow, receiving and factory layout are excluded in the proposed framework. The same holds for factory planning: Only local geometric collisions within a workstation are regarded.

- Nowadays, automotive products consist of both **software and hardware components**. Neither software validation processes nor electrical validation is described in the context of this doctoral thesis. Additionally, no "end of line" contents are regarded in the final assembly stage, such as rain simulation, flashing all electronic devices and filling up fluids.

- **Virtual assembly simulation techniques** continue to be an on-going active research area. No deformable, flexible and fluid simulations are included in the research focus of this doctoral thesis: Deformable and realistic material behavior is still an actively researched field, so called "deviation in form and dimensions" during manufacturing processes, such as "plastic deformation, thermal expansion, tool wear, inadequate tooling" [34].

These contrasts and cooperations provide a focus on relevant research aspects in the context of the collaborative virtual validation of manual assembly tasks without failing to address related research.

DOMAIN ANALYSIS PRODUCTION PLANNING AND VALIDATION

Manufacturing industries, especially carmakers worldwide, are facing new challenges with regard to product complexity, globalization of markets, digital transformation, connectivity, autonomous driving, alternative powertrains and stricter global environmental regulations as well as the need for even more environmentally friendly products. Product-related requirements also have an impact on future production systems: "Today manufacturing and production engineering is undergoing an enormous innovation process worldwide" [35].

In order to contextualize the following concepts in this doctoral thesis, this chapter provides an overview of the fundamentals of the automotive development processes with a strict focus on final assembly. Thus the fundamentals of the structure of automotive production, relevance of manual labor in final assembly stages, production paradigms and factory arrangements are presented in this chapter. In addition, this chapter discusses the need for DMUs and PMUs. An in-depth description of the industrial practice of PV workshops is provided along with their corresponding verification goals.

3.1 DOMAIN ANALYSIS OF AUTOMOTIVE PRODUCTION

The automotive PDP is described in the following section, followed by explanations on product life cycles, global production networks and automotive final assembly.

3.1.1 Automotive development process

Large scale companies, such as automotive original equipment manufacturers, require systematic processes with clearly defined responsibilities to bring a product successfully to market. Following Stark et al., the literature uses multiple product life cycle definitions, depending on the stakeholders, e.g. users, marketing, environmental viewpoints or manufacturers [36, p. 6]. In this doctoral thesis, a product lifecycle is defined as the "period starting with a product idea and ending with disposal at end of useful life" as defined in the CIRP "Dictionary of Production Engineering" [35, p.447]. Another definition for product lifecycle is described in the "Dictionary of Production Engineering" as the "period of time for which a product is in the market." The product lifecycle consists of the following five phases: Development phase, introduction phase, growth phase, maturity phase and

declining phase." [35]. The assumed sales volumes for such a lifecycle are depicted in Figure 3.1.
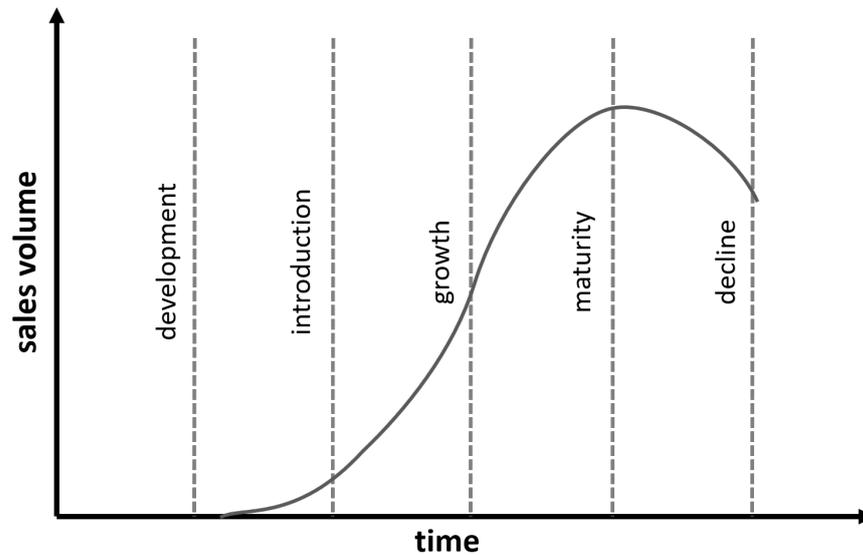
Figure 3.1: Idealized sales volume over time throughout the product lifecycle (as presented in C.I.R.P. Dictionary of Production Engineering [35])

In contrast to the market's supposed sales volumes over time, a typical car model lifecycle from introduction to decline is expected to last approximately seven years for premium OEMs, even though the market demands more frequent product updates, such as yearly updated models [6]. Since OEMs aim for overall continuous sales volumes and continuous development efforts, they try to become independent of a single product's lifecycle. For each OEM, automotive development projects are arranged in a staggered manner so as to continuously bring new products to market. For the manufacturing industry, the PDP mainly takes place during the first two stages of the product lifecycle, from the first idea to the Start of Production (SOP). Maintenance and aftersales support customers during the usage of the product.

Automotive industrial development is usually organized in projects. Weber [25] presents five types of design levels for development projects, which vary significantly in the required overall financial effort, length of time and technical content:

- **Complete redesign:** Redefining or creating a completely new product including all visible parts.

- **Derivative design:** Reusing parts from the same platform and system architecture

- **Variant design:** Building model families by changing as little of the product as possible in order to offer a wider range of products, such as sedan and hatchback.

- **Model updates:** These face lifts are carried out to increase the perceived value with as few changes as possible while offering customers novel features and an updated product

- **Model year:** Reduction of cost and increase in product quality throughout the product lifecycle

An independent procedural development model is described in VDI-2221 [37]. It presents a systematic procedure for developing and designing technical systems. This general method consists of seven detailed steps from "problem definition" to "realization" of a product, and each step provides input and output documents throughout the four general phases "planning, concepts, design and development." In detail, they propose seven in-depth steps namely "clarification and definition of the problem" (1), "determination of functions and their structures" (2), "search for solution principles and their structures" (3), "dividing into realizable modules" (4), "form design of the most important modules" (5), "form design of the entire product" (6) and finally "compilation of design and utility data" (7). Overall, the VDI-2221 proposes this as an iterative process with progressions and regressions.
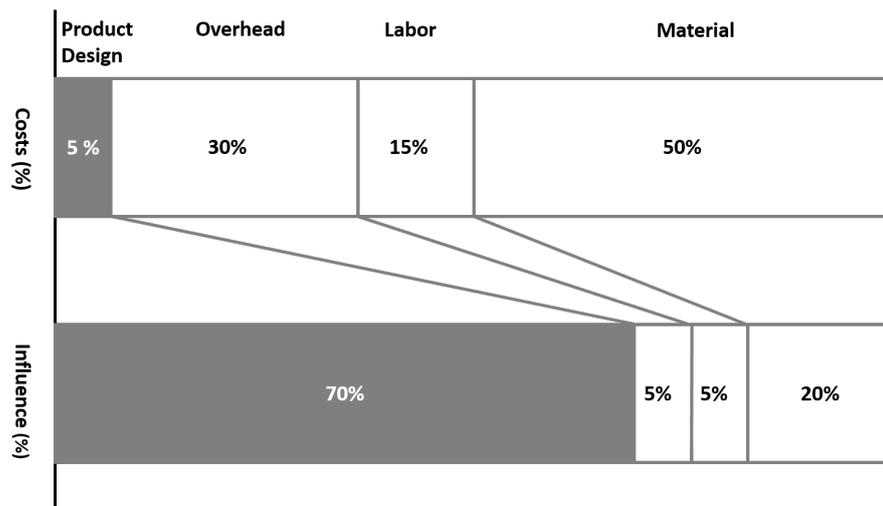
Figure 3.2: Example of the resulting cost influence on the product (Munro and Associates, Inc. [38] and Lotter and Wiedendahl [39])

In such design and development processes of automotive projects, the vast majority of costs are already determined during the early stages. "It is now widely accepted that over 70% of the final product costs are determined during design." [38, 40]. Figure 3.2 illustrates this example. Thus manufacturing and assembly should be taken into account as early as possible, even during the design cycle. Figure 3.3 underlines the importance of the so-called "simultaneous engineering" and "concurrent engineering" concepts. Parallelized production planning and development phases enable a bilateral interaction of

these departments. This method reduces mistakes in advance of SOP, such as wrong design considerations or wrong logistics concepts. The quality and efficiency of the production system improve when complying with "concurrent engineering" principles [41]. Analogously, "Design for Manufacturing and Assembly" (DFMA) is a process and a qualitative method to assess product design with the goal of reducing overall costs and increasing profitability. In 1994, Boothroyd found that using the DFMA methodology "shortens the time taken to bring the product to market" [42]. He describes the former attitude of designers towards the manufacturing engineers as "we design it, you build it" [42] or also called this an "over the wall approach." There, designers passed their designs to the manufacturing engineers without any feedback loops. For instance, manufacturing engineers had to deal with manufacturing and assembly problems even though they originated in the design department. DFMA aims to design products with less complexity, higher standardization, fewer parts and ease of assembly in mind. "Hidden waste" can be found in product designs with regard to complexity, time, energy, labor, defective production and many more [42]. Concurrent engineering, simultaneous engineering and DFMA are countermeasures to improve quality and to reduce waste.

Due to the complexity within the automotive development processes, structuring and organizing these projects is inevitable. Every automotive OEM develops production processes that follow certain phases, quality gates and milestones throughout the PDP [25]. Figure 3.3 shows a generalized automotive development process (e.g.VDI-2221 [37]) with respective milestones (compare Walla [27]). When a SOP (Milestone A) date is set, all other milestones and quality gates are backdated, forming a concrete time frame for all development phases. After each development phase, a specific quality gate must be reached. An OEM-specific example of a PDP is presented by Geissel [43], namely for Mercedes-Benz Cars.
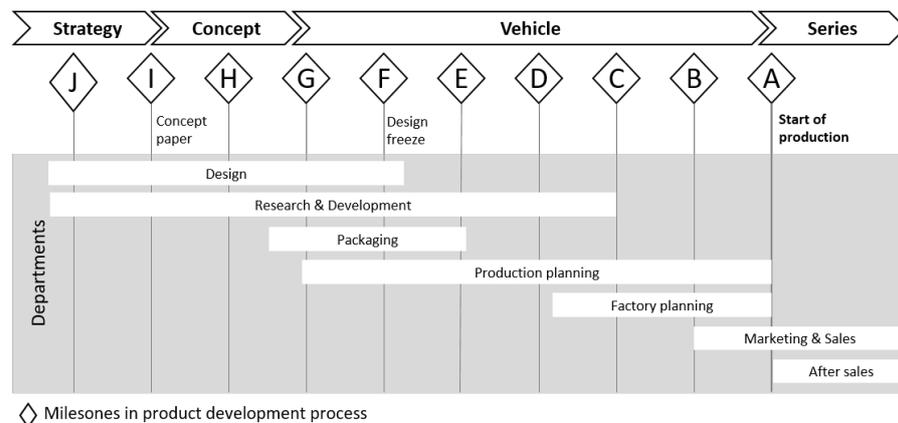
Figure 3.3: Generalized time plan for an automotive PDP (adapted from Walla [27] in accordance with Weber [25])

In Figure 3.3, interdependencies between multiple departments are depicted throughout a product development project. The time overlap of responsibilities between "research & development" and "production planning" allows concurrent engineering and design of manufacturing and assembly. Time frames do not represent firm starting or end points, but represent the time frame with main workloads for the respective departments during the development projects. For example, even though Figure 3.3 indicates that production planning begins at quality gate G, in reality the responsible department already starts out with relatively little effort at quality gate J by writing down manufacturing requirements for the concept papers. On the other hand, research and development does not stop working on a product at quality gate C as long as product improvements can be implemented.

3.1.2 *Automotive production*

Automotive factories are typically structured in four different assembly stages as shown in Figure 3.4: Press shop, body shop, paint shop and final assembly. Sometimes the literature combines the press shop and the body shop in one joint production stage.
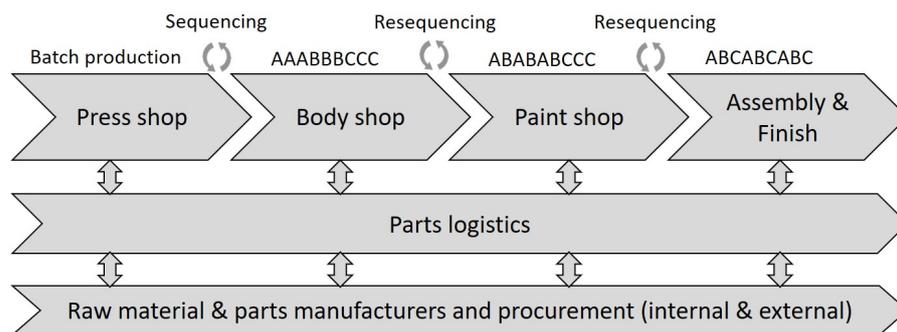
Figure 3.4: Essential stages during vehicle production process (according to Weber [25, p. 288])

In the **press shop**, raw material such as steel is delivered and formed into basic car body parts. Large manufacturing tools such as stamps and presses are utilized. Raw steel is stamped and bent to smaller units of the car body. In general, the press shop works on parallel sequences and creates similar parts in batch lots since machining has changeover times. Body parts are stored in local logistics areas. Typically, this stage is highly automated with low manual effort - at least in high-wage countries [44].

In the **body shop** these parts from the press shop and from suppliers are welded together in order to form the raw car body. Of all vehicle production steps, the body shop has the highest degree of automation. Depending on the definition of an automation degree, it inherits approximately 90% or 95% of automated processes [25].

In the **paint shop**, essentially multiple layers of coating, wax and paint are applied to the car body. After finishing the paint shop stage, the production sequence is changed again with the help of re-sequencing buffers. The output from the paint shop stage is the so called body-in-white (BIW). This production stage also has a high degree of automation.

In **assembly and finish**, the painted car body is assembled and finished for both interior and exterior parts. To date, assembly has a low automation degree [45]. This is why, of all four stages of the vehicle production process, manual labor is most prevalent in final assembly and finish. The final assembly stage is the central step for enabling product variants in the vehicle production process [46].

Between each production step, there are buffers used to decouple all stages and to permit reliable production, shift decoupling and re-sequencing between the production stages [25, 44]. The stream of car bodies is re-sequenced in logistics buffers. Flexibility within the production sequence enables efficient processes in final assembly stage. All four stages of the vehicle production process depend on the complex handling of parts within a plant and on-time delivery of raw materials by logistics departments. Throughout all production stages, logistics handles parts from internal and external suppliers, such as tier 1 and tier 2 suppliers. Great product variety often does not allow a direct material supply of all variants to the place of assembly. In this case, parts must be delivered just-in-time or even "just-in-sequence" to reduce the need for storage capacity and thus save costs.

Overall, Weber states that during the production process, the production plant ideally creates a "constant stream of parts, components and eventually complete vehicles of perfect quality" [25]. This stream must be fitted into appropriate production steps. In general, automakers must set up production systems including assembly systems to produce products, such as cars, trucks or vans.

These production steps are embedded in a global architecture of production facilities. Together they form a global production network with interdisciplinary tasks and divided responsibilities: "Products and related services are provided by production networks where autonomous enterprises are linked by relatively stable material, information, and financial flows. A production network typically includes nodes of suppliers and manufacturers involved in direct value-adding activities, distribution centers and logistics service providers, as well as facilities and channels for reverse logistics." [47]

Figure 3.5 shows the organizational hierarchy and connection from a globally distributed production network to an assembly workstation. In the case of automotive factories, parts and services are bought by tier 1 and tier 2 suppliers from all over the world. They deliver these parts to the assembly buildings and assembly lines at the OEM's factory.
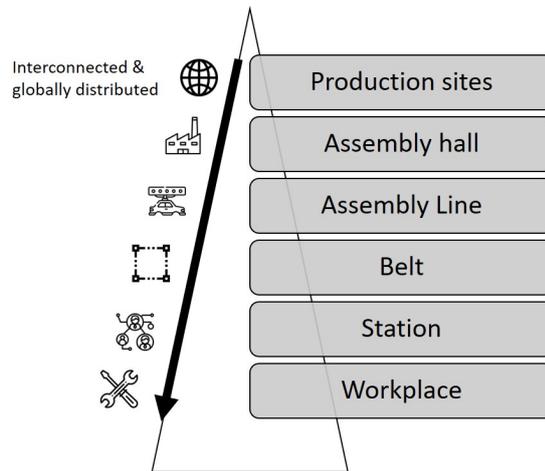
Figure 3.5: Hierarchy of a global production network break-down for assembly systems

3.1.3 *Automotive final assembly*

This doctoral thesis focuses on the validation of manual processes in the final assembly stage. This stage is considered to be the most expensive production process, as the automation degrees in all other stages are higher [46].

That is why the general principles of an **assembly** are described in greater detail: The "CIRP Encyclopedia of Production Engineering" defines an assembly system as "one of the subsystems in a manufacturing system – factory - where the individual components of a product are joined together and thus integrated into a semi-finished or into the final product" [48]. Various production concepts are arranged in multiple geometric ways, such as assembly stations, assembly cells, assembly lines, etc. Assembly is defined as a central part in production engineering namely the organizations in the manufacturing industry in which the product is finished or "the gateway to the customer" [35]. Assembly is part of the entire production system.

Manual assembly is defined as "the assembly of products and subassemblies manually or without the use of automatic assembly machines" [35]. Final assembly is needed in order to make products with higher complexity out of single parts which are produced at different times with a lower degree of complexity [39]. In the automotive industry, each passenger car inherits thousands of assembly parts, packages or pre-assembled parts from automotive suppliers including internal supply.

For carmakers, this **final assembly stage** usually consists of one central **continuous flow line production** with several related preassembly systems. Work tasks are distributed throughout the cells and workstations along the production line, as depicted in a generalized process chart in Figure 3.6.
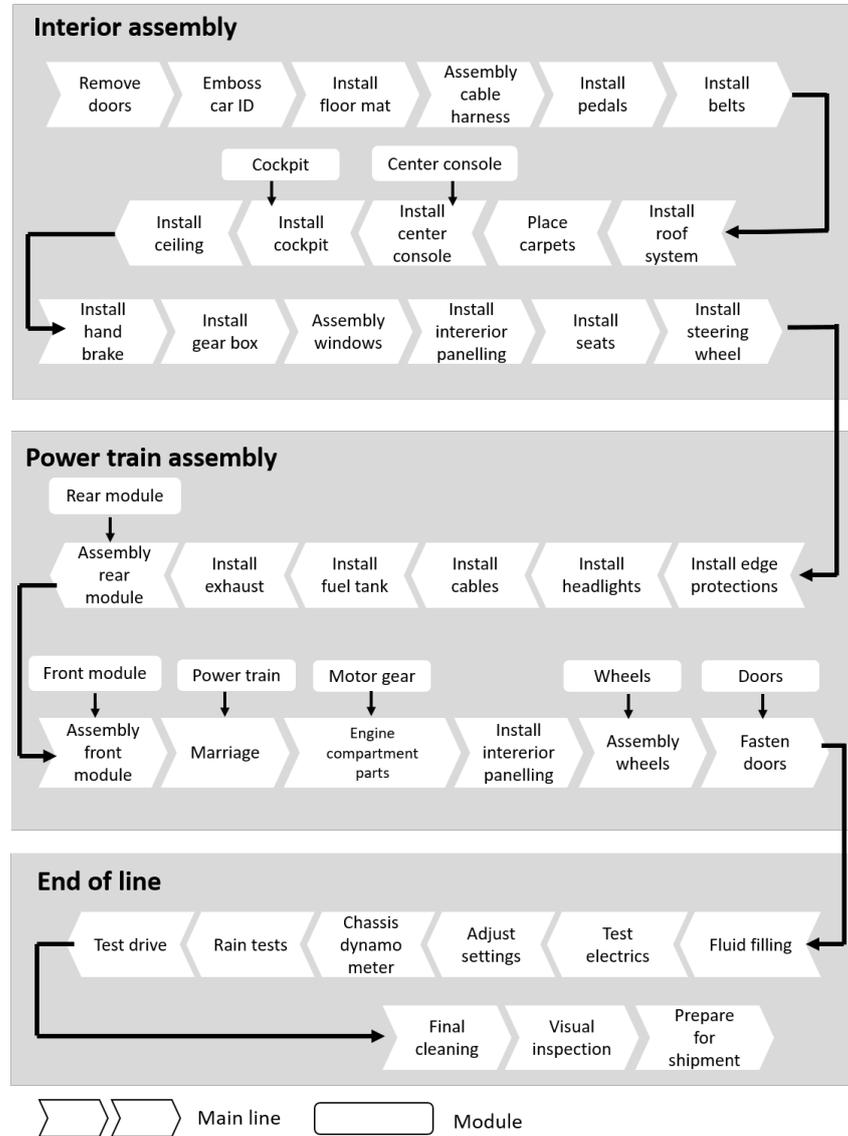
Figure 3.6: Generalized structure of automotive final assembly work contents (according to Lars Weyand [49])

Such a **main assembly line** consists of conveyors or hanging brackets (so called C-hangers), which continuously transport the partly assembled products through several hundreds of so-called cells at a constant speed. Given a certain conveyor speed and a fixed cell length, the predefined cycle time can be deduced, which typically ranges from several seconds to several minutes. This production system's pace influences the overall throughput also measured in "jobs per hour." In 1913, Ford implemented the first paced automotive assembly line, reducing production time from more than 12 hours to approximately 90 minutes [50]. Today, an Audi A3 model production in Györ still follows continuous flow line principles, where final assembly consists of 146 cells with "exactly two minutes" cycle time [51]. As shown in Figure 3.6, buffers add additional cycle times to this calculation. Due to limited space and practicability, one cell typically inherits up to six workstations where people work together on their assigned tasks.

The products continuously arrive at the worker's place at a constant speed. The worker is assigned to perform the tasks on each car with a limited amount of responsibility during the cycle time. This idea of an optimized division of labor in continuous flow line production systems is the key success factor for making production processes efficient. Consecutive work tasks also cause interdependencies at all consecutive workstations. Changing one cell's work content has an impact on the overall efficiency of a production system.

Since not all components can be produced in the main consecutive production line, several pre-assembly cells deliver sub-assemblies to the main production line (see modules in Figure 3.6). Some of these pre-assembly cells also follow the flow line production principles, such as the main assembly line for doors, the cockpit and the combination of drivetrain and suspension" [51]. Bringing together multiple sub-assemblies in such a manner is referred to as the "herringbone principle" [51].

According to the VDI 2860 and DIN 8593-0 standards, **typical manual work contents** within automotive final assembly are handling and joining, adjusting and inspection tasks, and complementary tasks [52, 53]. According to Lotter et al. [39], handling tasks comprise storing, modifying quantities, shift arrangements and securing. Complementary tasks comprise marking, changing temperature, cleaning, deburring, printing, covering, peeling off, unpacking and sealing. Joining tasks are clustered in piecing together, filling, welding soldering, joining by adhesives, mechanical means, forming processes and others. One of the reasons for carrying out assembly tasks as **manual labour** is to achieve greater flexibility and changeability in assembly lines.

According to Küber et al. [54], the main causes of changes can be clustered in volume changes, model mix changes and vehicle deriva-

tive changes. They discuss the interdependency of economic efficiency vs. flexibility and assembly vs. logistics processes. Therefore they introduce a novel decision making process for strategic planning of logistics and assembly processes. The so-called **strategic decision square** takes into account that for multiple time dynamic scenarios and for different optimization goals there is no overall optimum in economic efficiency. The decision space described there mainly consists of two trade-offs. The first trade-off must be made between assembly and logistics processes whereas the second is between flexibility and economic efficiency. Different scenarios must be generated to find strategic decisions for each factory. They exemplify that a basket within logistics brings a lot more flexibility into the assembly line whereas it is much more complicated for logistics because costs increase as a result of additional material handling steps.

This need for flexibility in automotive final assembly is enabled by a large portion of human labour, since this continues to be the most flexible way in final assembly production systems. The German Industrial Standard DIN IEC 60050-31 describes the **degree of automation** as the "proportion of automatic functions to the entire set of functions of a system or plant" [55]. Additionally, Fujimoto et al. proposes a set of definitions for "automation ratio." Since there is no single definition of automation ratios applicable for heterogeneous tasks, he proposes to categorize the operational definitions of automation: machine-based definition, worker-based definition, material-based definition and process-step-based definition. For final assembly he proposes to measure the automation ratio by "the number of parts assembled automatically in the main line (excluding bolts and fasteners) in comparison to the total number of parts assembled" and "the ratio between workers or person-hour saved by automation and those necessary for a totally non-automated process." [56] While they found a high automation ratio (average around 90%) in stamping, welding, engine machining and engine forging, in final assembly areas an average automation ratio of 10% has been found. Typically in European automotive factories, press shops as well as body shops and coating production stages have a large percentage of the value creation automatized, while in the final assembly stage, the automation degree is still low [57], even when producing large volume models. Lay and Schirrmeister discuss whether nowadays the automation degree in the final assembly stage is even too high [45]. Reduced lot sizes, capacity flexibility, lower invests and higher product flexibility are the most relevant reasons for reduced automation degrees [45]. This underlines that human labor is still a major variety enabler in final assembly stage.

Lotter summarizes the optimization goals in automotive final assembly as the minimization of assembly and training time, efficient quality assurance and simplification of the assembly tools and tasks.

He concludes that this is only achievable through the use of simulation tools [46].

3.2 DOMAIN ANALYSIS OF AUTOMOTIVE PRODUCTION PLANNING

Production planning is an interdisciplinary task in economics, mechanical engineering, production engineering, data analytics and computer science. It deals with strategic planning, structural planning, systems planning and operations planning of all processes in the upcoming factory [58]. Holistic production planning thus affects all domains of "product, technology, organization, tooling, personnel and finances" [58, p. 18] which are required to produce products and goods.

Production planning comprises all measures in designing a manufacturing system and production processes. In the Dictionary of Production Engineering it is defined as "a function that defines the totality of activities to put into place in order to meet the objectives of the production program, broken down into primary needs planning, material management and time management" [35]. Stecca defines production planning as "the process of translating customer orders to jobs for the manufacturing plant with attached due dates" [59]. In detail, "manufacturing process planning" specifies all required work steps to execute the customer's product demands while optimizing the production system with respect to multiple criteria within the given constraints. Therefore production planning is a systematic, goal-oriented process in consecutive phases using specific tools and methods in order to plan a factory from the first idea to SOP (Walla [27, p. 15] based on Grundig [60]). For example, production planning sets up manufacturing or assembly systems, e.g. stamping, mulling, turning, assembly and many more. "Process planning can be defined as the task which determines how a part should be manufactured according to the design specifications" [61].

The production planning process follows a **top-down approach**, from high-level planning to an in-depth process specification [62]. El-Maraghy et al. specify four steps: First generic planning is carried out, where conceptual plans specify the required technologies, such as the overall throughput of the production system. This is followed by macro-level planning for product sequencing and multi-domain optimizations. Subsequently, detailed planning focuses on single optimization domains with detailed plans, e.g. tools and resources in assembly. Finally, micro-level planning is carried out, optimizing certain parameters of the production process for optimal conditions, such as MTM time analyses, work task sheets and alphanumeric work task descriptions [62].

When executing a new projects, production planning can have differing **starting conditions**. Sometimes no factory exists and a completely new production system can be designed. All production tools,

locations, structures and personnel must be planned from scratch. This case is called "green-field" production planning. By contrast, "brown-field" production planning implies that only parts of existing production systems are redesigned and others must be reused. Depending on the strategy, existing factories must be extended, reduced, renewed, restructured, relocated or outsourced in order to integrate novel products, change production quantities, change the organization or update the production structure [27].

3.2.1   *Assembly planning*

As this doctoral thesis specifically focuses on automotive assembly, assembly process planning is described in detail. In the "CIRP Encyclopedia of Production Engineering" [63, p. 827], Riggs distinguishes two levels of "planning assembly operations": process level planning and operation level planning:

- **Process level planning** considers transport routes and handling tasks between assembly operations. The work content and time required for a standard execution of each assembly operation is planned to enable line balancing to achieve the highest possible utilization for each work cell. The distance of movement and mass of transported parts between operations is also recorded. Maximum efficiency assembly processes must be found.

- **Operation level planning** is carried out for each workstation. Manual assembly sub-tasks are analyzed by operation method studies, namely motion studies and time studies. In these studies, the layouts of the workstations are analyzed and the placement of the components to be assembled is described. Motion studies concentrate on recording and analyzing the motion elements' types and respective magnitudes of motions. The optimization goal is to minimize movement. Time studies are used to determine the average time required for optimized movements.

For maximum assembly efficiency, all operations in the assembly process must be well planned to avoid unnecessary movements and amount of time spent on tasks [63].

Another generalized procedure for the **systematic planning of assembly systems** is described by Lotter [64]. Due to its generalizability and broad acceptance in many manufacturing industries, this has become a reference work. Lotter originally presented 11 steps for systematic production planning and Hartel and Lotter extend and revise this systematic planning in their book "Montage in der industriellen Produktion" [65]: First **system requirements** (1) with product amounts, amount flexibility, maximum usage time, designated shift

model, overall throughput and amortization times must be calculated. Next, a **product analysis** (2) counts the amount of required parts, assesses joining tools and quality requirements. Then, the **assembly sequence** (3) is defined, followed by a **functional analysis** (4) where larger work contents are decomposed in singular consecutive basic task components. Determining **cycle times** (5), creating **workstation layouts** (6) and calculating **personnel requirements** (7) for assembly are additional tasks. After that, **availability checks** (8) of the latter points are carried out to determine whether parts quality, workstation count, structure of tooling and personnel training are sufficient. The final steps are the creation of **technical specifications** (9), the **invest calculus** (10) of the assembly system and the assessment and **comparison of costs** (11). This generalized procedure represents the process steps in automotive assembly planning.

Overall, once a proper plan of the detailed assembly processes is in place, line balancing of mixed-model lines is enabled by this in-depth knowledge (see lean production approaches [38]) for efficient production. The continuous re-planning of detailed processes throughout the PDP can incrementally reduce planning vagueness [66].

3.2.2 *Tools in assembly planning*

In the early years of mass production - before having systematic planning of assembly tasks - a great deal relied on previous experience and incremental adjustments of detailed processes. Planning quality depended heavily on the skills and knowledge of the planner himself. Such planning methods are not reproducible, can be time consuming, error prone and do not guarantee high quality results for model-mix production lines due to overall complexity [61].

Industrial practice shows that non-digital tools are still state of the art in many different application areas of assembly process planning. For example, pen and paper based tools can be found in the creation of process charts, walk path studies, ergonomics studies, and time studies for operation level planning and documentation. Ergonomics assessments are still carried out using standardized assessment worksheets for each workstation [67]. Walk path studies use pen and paper methods to draw so called "spaghetti diagrams" on a piece of paper [68]. Another way of simulating production processes in a non-digital manner is to use physical mock-ups or physical prototypes. For factory resources, cardboard mock-ups of material zones are typical tools for spatial understanding. For the physical assembly of the product, physical mock-ups are used.

Therefore, the literature presents multiple tools to compensate the above mentioned disadvantages of manual approaches and pen and paper based methods. Computer aided production planning (CAPP) is an algorithmic approach that partly automates the task of gener-

ating the best process plans (see [62]). It supports planners by bridging the gap between Computer Aided Design (CAD) and computer aided manufacturing (CAM) [61]. Gülesjn distinguishes variant approaches and generative approaches in CAPP. Whereas variant process planning is an assisted extension of manual process planning, generative process planning uses artificial intelligence to automatically derive and produce process plans [61]. Today, for CAPP there are multiple commercial tools, such as Delmia by Dassault, IPO.Log by IPO.Plan, Process Simulate by Siemens and many more. Depending on the scope of the CAPP tool, multiple optimization scopes can be followed. Typical goals are to optimize the production sequence of variants, achieve line balancing by changing the assembly sequence, and to describe the holistic production process by combining the relationships between Product, Process and Resource (PPR). Other generative tools try to automatically derive process relevant information out of CAD data [69], such as fixture concepts or work task descriptions and automatically assess assemblability using CAD product information [70].

3.2.3 *Automotive Mock-ups*

Automotive mock-ups are a central element for the collaborative discussion, testing and validation of the product's features and properties. Design departments build clay models to validate a product's overall appeal and to gather user feedback on novel designs. Research and development utilizes automotive mock-ups for crash simulations, weather resistance, driving properties, sound design and many more. Similarly, production planning uses mock-ups for assembly validation, disassembly validation and training. "While virtual vehicle mock-ups represent a geometrically ideal world of rigid components, physical mock-ups allow investigations that include real-world physical effects such as plastic or elastic material deformation or production tolerances." [25, p.290].

In the following, different types of mock-ups are described in detail, from partial build-ups to full physical mock-ups up as well as digital representations, known as digital mock-ups. Figure 3.7 shows the usage of both digital and physical mock-ups throughout the generalized PDP. All types share the common goal of simulation and validation.

Physical Mock-Up (PMU) In the context of automotive product development, the term physical mock-up is used for a wide variety of hardware dummies, such as total vehicle mock-ups, component mock-ups or conceptual mock-ups [43].

Physical mock-ups are hardware-based representations of the upcoming product. They are typically built in true-to-scale and undistorted resemblance for the analysis, testing and visualization of the product's properties [71]. "At the beginning of each prototype build
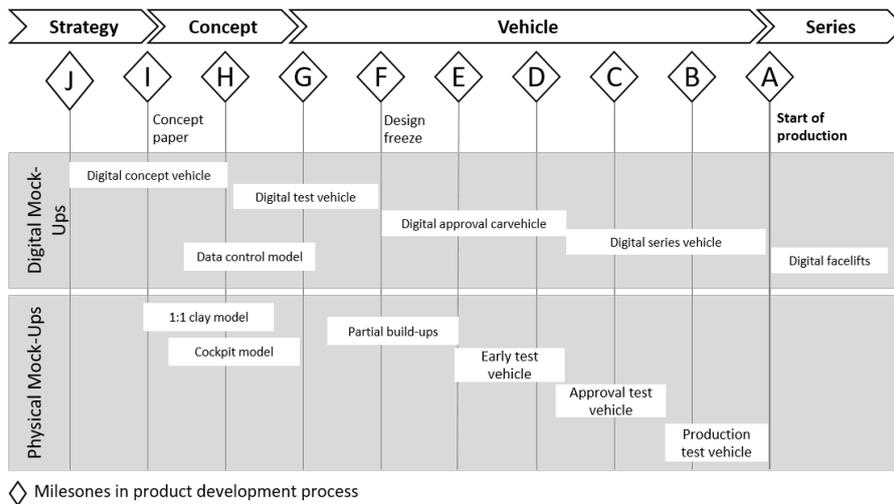
Figure 3.7: Types of digital and physical mock-ups throughout the PDP (in accordance with Walla [27] and Geißel [43])

phase, laminated bodies and other rapid-prototyping techniques are used to investigate the behavior of deformable, elastic or labile parts prior to the actual build" [25, p. 290].
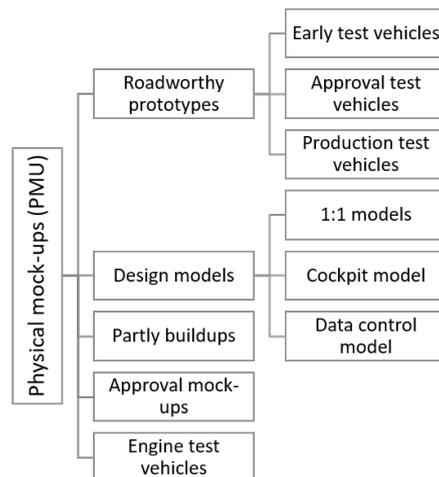
Figure 3.8: Classification of physical mock-ups in the automotive industry (in accordance with Geissel [43])

Each automotive development project uses a wide variety of mock-ups (see Figure 3.8). Geißel [43] describes the following PMUs:

- **Data control models** are milled reference models, based on CAD data of car chassis, interior and exterior parts. Its goal is to achieve a final commitment on the design of surfaces of the upcoming product.

- **Approval mock-ups** are true-to-scale models of the chassis. These mock-ups are built in the early development phase using rapid

prototyping methods. Typically, approval mock-ups are not road-worthy and do not have any functionalities besides geometric validation. Typical use cases are to validate the wiring harness or to check the installation space for supplier products. Depending on the use-case, they are built of wood, clay, carbon, polymers, steel or aluminum. [72, p. 291]

There are mainly three different stages of roadworthy prototypes:

- **Early test vehicles** are the first roadworthy prototypes of a new product. All parts are intended to be used in the novel product, making this a non-partial build-up. These early total vehicle prototypes contain all experimentation components brought together for the first time in a new product. The interaction between individual, novel components are tested, parametrized and validated. Based on this, approval of concepts for machinery and plants is granted. Only few products are built, since early prototypes are highly cost intensive.

- **Approval test vehicles** are more mature test vehicles. Design must be finished and these prototypes are built using final production tools and instruments. This prototype is built in order to verify the product specification sheet of functionalities and goals. Compared to early prototypes, larger quantities are built.

- **Production test vehicles** are already built in the designated plant and on the designated production line. They are used to demonstrate process compatibility and the overall manufacturability of the car under real cycle time conditions. With these prototypes, a smooth ramp-up during SOP is achieved. Additionally, production test vehicles are used to train the workers on the upcoming product to prepare them for SOP.

Digital Mock-Up (DMU) In contrast to PMUs, digital mock-ups (DMUs) are defined as realistic computer-generated digital models or as-well simulations, which in terms of appearance and functionalities resemble the original product under development as closely as possible [73, p. 67]. DMUs are also referred to as virtual and digital prototypes.

The use of DMUs allows "designers, manufacturing planners and management" to work on virtual products in order to make decisions related to the "concept, design and all downsteam processes" [74]. For example, research and development simulates drive trains, "noise, vibration and harshness" of car chassis, crash simulations, stiffness of single parts, aerodynamics, thermal design and ride handling (see [43]). In production planning, DMUs allow engineers to design, configurate and validate products without the need to build a physical model. "A Digital Mock-Up (DMU) is used for simulating assembly and dismantling, in addition to testing for collisions

and buildability. For this purpose, all geometry data of all vehicle parts are brought together in their planned installation positions" [5]. "Because of their flexibility and cost-efficiency, virtual mock-ups are the preferred means for assessing assembly processes, not only during concept phase. Container, fixtures, tools and man-models can be integrated into the virtual scenes – which allows a realistic evaluation of transport and handling, fixing processes, and ergonomics" [25, p. 289].

With the broad availability of advanced 3D computer graphics techniques, light-weight, standardized data structures, direct PDM interfaces and real-time rendering, DMUs are becoming more important than PMUs. They offer the ability to load models, measure, analyze, simulate, and redesign instantaneously. "Virtual cars have become the central communication platform for the co-operative vehicle development process. They are the substitute for hardware prototypes to an ever growing extent" [25]. The continuously growing possibilities of digital methods in Product Lifecycle Management (PLM) systems in product development leads to a higher penetration of DMUs instead of physical ones.

Building PMUs is costly and time-consuming [75]. By using of DMUs in the PDP, costs can be significantly reduced compared to PMUs [76, p. 43]. Moreover, qualitative aspects such as earlier and interdisciplinary usage of digital mock-ups for geometric, functional and production-oriented validation is an even bigger benefit [73, p. 68], e.g. crash simulations or assembly simulations can be carried out without hardware models. The long-term goal is to be able to assess all validation tasks using DMUs with virtual assembly and simulation of 3D geometry. Therefore, products would be of higher quality [77] and have a higher degree of maturity in shorter development times and with reduced overall costs (see [76, pp. 42-43]).

3.3 PRODUCTION VALIDATION WORKSHOPS

PV workshops are held to validate planning results with respect to multi-objective optimizations. Such PV workshops are common in industrial practice and are held in various manufacturing companies, especially in the automotive sector. Examples of published PV workshop processes and experiences can be found at Ford [42], Daimler [78], BMW [25] and Volkswagen [79]. Even though they have slightly different names, they share the same assessment scopes. Such multi-objective optimizations processes involve the validation and verification of upcoming production processes. PV processes are interdisciplinary and require various expertise, background and roles of the participants.

PV workshop stakeholders resemble the main customers of the framework presented in this doctoral thesis. Therefore in-depth insights of

validation and verification processes are given in the following section. This section presents the validation principles, required inputs, organizational aspects, participants, roles, goals and assessment criteria in the following.

> **Publication**
>
> Parts of this section were published in the conference paper CIRP CATS 2016 "Dual Reality for Production Verification Workshops" by Otto et al. [80]. This section has been extended and revised for this doctoral thesis.

3.3.1 *Goals*

The overall **goal in production** is to generate products and services of perfect quality with optimized costs and maximum flexibility [54] while ensuring profitability even when unexpected changes occur, e.g. a drop in production volume. This holds for planned and unplanned changes of the production system [25, p. 288]. Therefore, the production system must be optimized to react to production process "disturbances", such as design changes, model year measures, the launch of a new model, variants and options, but also unplanned events such as parts quality issues or failures in production equipment. Quality improvements, customer satisfaction, reducing manufacturing costs, and lead time reductions to bring new products to market are all in the scope of PV workshops.

Once all relevant aspects of the upcoming production system and its detailed work contents (see Lotter [65] and Section 3.2.1) have been pre-planned, PV workshops aim to improve the pre-generated planning quality of the various planning departments in order to reach a bullet-proof, mistake-free production ramp-up in the actual plant by iteratively simulating and optimizing future production plans. Consequently, the **overall goal of PV** is to ensure such smooth ramp-ups and to reach the maximum throughput capacity of the production system.

In practice, PV assesses all aspects of the upcoming factory processes, such as work plans, product precedence graphs, availability of tooling, ergonomics, cell layouts and many more. PV workshops focus on manual labor in final assembly in automotive flow line production, but not on fully automated processes. Hence, such an interdisciplinary task in production planning consists of collaborative work and must find trade-off solutions for potentially conflicting objectives. For some optimization goals, no optimal solution can be determined, which means a trade-off has to be found (compare [15, 54, 81, 82]).

The **Kaizen management philosophy** is a theoretical basis for reaching goals in PV workshops. It proposes a constant optimization of production and translates from Japanese as "change towards the better."

It aims for continuous, step-by-step improvement of all internal processes in an enterprise through involvement of all affected employees. Its objectives are to eliminate all sorts of redundancy, to concentrate attention on the place where value adding happens and to create new standards of productivity and quality [35].

3.3.2   *Verification and validation principles*

The automotive development process (see also Section 3.1.1) is an iterative process with steps back and forth and recursive loops. Changes throughout the PDP can be induced by stakeholders, market demand, sales, logistics or production itself.

The "International Standardization Organization" differentiates **verification and validation** in ISO 9000:2015 [83]. Verification is defined as the "confirmation, through the provision of objective evidence, that specified requirements have been fulfilled" whereas validation is the "confirmation, through the provision of objective evidence, that the requirements for a specific intended use or application have been fulfilled" [83, p. 3.8.13]. This subtle difference can be explained as follows: In verification the question is addressed whether "we are building the X right", whereas in validation the question is whether "we are building the right X" [84].

During PDP, production engineers consistently check the vehicle's compliance with production requirements using both virtual and physical mock-ups [25]. Therefore validation serves to control the current development and planning status during PDP (see [27, p. 13], [85]). As depicted in Figure 3.9 and as stated by Meißner and Walla, validation takes place in three steps:

- **Analysis of the current development and planning status:** By tearing down, separating and examining the features of a product's elements and dependencies, information is gathered.

- **Examination of the analysis results:** The steps required in examination are data authoring, interpretation of results and assessment. Data preparation is carried out by selecting, sorting, formatting and visually presenting in an interpretable manner. In order to simplify interpretation, data must be presented in a way that directly links causes and results.

- **Decision for further measures:** Countermeasures must be decided if the real development status deviates from the expectations after analyzing the results. For example this can lead to changes in the product, production plans or novel manufacturing principles.

Walla and Meißner propose to carry out change procedures in analogy to the VDI-2221 development process in four steps:
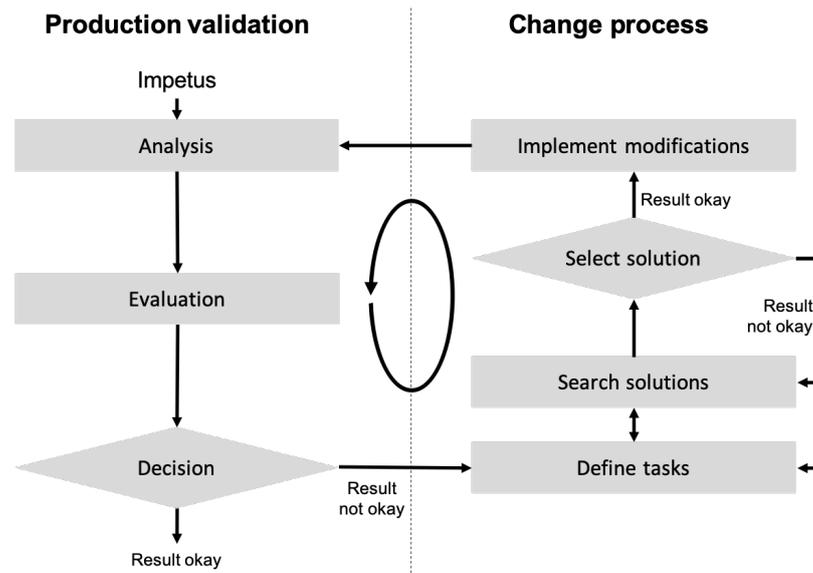
Figure 3.9: Iterative cycle for verification/validation and its corresponding change process (Meißner 2010 [85])

- Definition of task and description of the causes and reasons for the required change process

- Finding multiple possible solutions

- Narrowing down solution set and decision for one realizable solution, which meets the technical, organizational and economic criteria.

- Realization of the changes

This optimization loop is carried out until the results of validation and verification meet the requirements. When the product features are finally in accordance with the requirements, this iterative change cycle of PV ends. Especially for multi-criteria optimization, this cycle must be repeated multiple times in order to dissolve heterogeneous, contrary aims, e.g. initial investment for handling devices vs. ergonomics or introduction of cost intensive logistics vs. cluttered material zones in manual final assembly.

On a practical basis, this validation cycle is carried out at all automotive OEMs. PV workshops are summarized by Weber [25, p. 287] as follows: There are two main factors on how production can be optimized: "production environment, including equipment functionality and availability, workforce qualification, and process maturity" and that "parts' and vehicles' design is in compliance with the requirements of i.e. manufacturing, assembly and logistics." Additionally, Boothroyd [42] describes a Design for Assembly workshop situation at Ford with the following process steps:

- "Review the parts list and processes.

- Break up into teams.

- Analyse the existing design for manual assembly.

- Analyse the teams' redesigns for manual assembly.

- Teams present results of original design analysis versus redesign analysis.

- Prioritize redesign ideas: A, B, C etc.

- Incorporate all the A and B ideas into one analysis.

- Assign responsibilities and timing"

This process reveals strong similarities to PV workshops, whereas DfA workshops mainly focus on product optimizations with respect to assembly. PV workshops follow a broader approach as more verification criteria are taken into account.

3.3.3 *Organization and timeline*

During the entire PDP, multiple PV workshops are held in order to fulfill differing verification tasks and to assure product and process quality at certain milestones. The timeline of PV workshops depends on the project timeline and is backdated from SOP for every new model or derivative vehicle. Figure 3.10 depicts a generalized timeline, showing multiple workshops taking place throughout the development stages. Depending on the intensity of the development project, PV workshops typically take from a few days to several weeks. For a complete new car model, three weeks are typical, whereas for derivatives solely changing contents are validated in three days.

During PV workshops the new product is being assembled part by part. Each mating part is added to the vehicle in the sequence of the plant. Each part and process step is evaluated according to multiple criteria (see Section 3.3.5). Assessment criteria change throughout the PV workshops:

1. Early PV workshops focus on **product quality, product optimizations and a rough manufacturing sequence.** In detail, producibility, collision freeness, standardization of fixtures, buildability and accessibility of tools and operators' limbs.

2. Subsequent PV workshops handle **product-related processes**. At this stage, there is no focus on the production plant.

3. After that, **logistics, carriers, racks** in combination with walk-paths and overall operation level optimizations are evaluated with respect to the general model-mix.
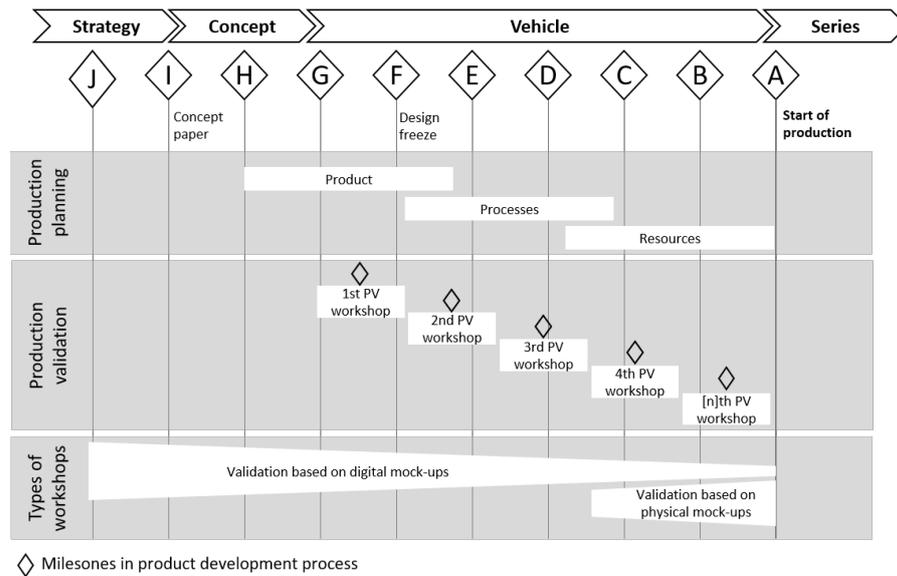
Figure 3.10: Generalized timeframe for production validation workshops (partly based on Weber [25, p. 291])

4. Then the overall **factory-related processes** with factory geometry and tooling are validated.

5. Final PV workshops focus on the **sign-off** of the overall process regarding iteratively optimized products, product-related processes, plant-related processes, cell layouts, logistics layouts and manufacturing sequences.

In practice, there are two types of PV workshops: **Purely digital PV workshops** and **purely PMU-based PV workshops**. In purely hardware-based PMU workshop situations, all assembly parts and resources have to be physically present. No mixed reality build-ups are currently used in industrial practice, as proposed in the literature by Arteaga et al. [15].

Both PMU and DMU-based workshops share the same verification criteria. Which kind of PV workshop is held depends on the current state of the PDP, and whether or not PMUs are already available. In general, DMU-based workshops are held in the early phases, whereas PMU based workshops are held closer to SOP (see Figure 3.10). In PMU-based hardware setups, only one specific car is built step by step. Due to the large variance of options and extra equipment, not all variants of options can be evaluated during PV workshops. This PMU reference product typically includes as much extra equipment as possible, in order to check as many details as possible. Nevertheless, only one configuration can be validated, but not all possible combinations. Neither can the effects of mixed-model assembly lines [86] be simulated in PMU based workshops, due to missing flow line production. Overall, the number of traditional PMU-based scenarios is decreasing. In

purely hardware-based workshop situations, all assembly parts and resources are physically present. However, both for hardware and virtual assessments, the assessment tasks remain the same.

Hence, PV workshops are not run for planning purposes. The input variables such as well documented process descriptions, cell layouts, manufacturing sequence, etc. have to be prepared prior to the workshop. Planning itself cannot be carried out during PV workshops. Some small, local optimizations can be carried out on-the-fly, while most optimization changes are only identified within PV workshops and implemented afterwards.

In the literature, PV workshops are closely related to the continuous improvement process [81]. Both share similar verification, validation and optimization goals, but whereas the continuous improvement process takes place after SOP, PV workshops take place before SOP. The duration of continuous improvement workshops are individually determined based on actual optimization needs. For example, they are carried out for planned and unplanned impacts during production, such as rapid changes in demand or for planned model updates.

| | | Verification criteria | | | | |
		Product-related	Human Related	Logistics-related	Process-related	Resource-related
Work place 1	Work task 1	●	83% ○	20% ●	97% ◗	74% ○	11%
	Work task 2	◗	73% ●	91% ◔	22% ◐	47% ◗	63%
	Work task 3	○	6% ○	9% ◗	75% ○	0% ◗	74%
Work place 2	Work task 4	◔	25% ◔	21% ●	87% ○	15% ○	8%
	Work task 5	◔	36% ◗	70% ◗	74% ●	79% ●	85%
	Work task 6	○	12% ○	7% ◐	44% ◔	23% ◔	23%
	Work task 7	◗	74% ●	89% ●	78% ◐	47% ●	96%

(left vertical label: Manufacturing sequence)

Figure 3.11: Documentation of the process maturity according to the verification goals in a matrix (in correspondence with [25, p. 292] and [34])

For each build-up step all criteria have to be assessed and documented. This leads to three main outcomes of PV workshops:

- On-the-fly optimizations are carried out in the planning systems if limited time allows for online documentation. Problems and derived countermeasures are collected in a list and assigned to a person who is in charge of resolving the issue. This systematic documentation is filled out using Excel spreadsheets or proprietary systematic documentation tools.

- Problems and derived countermeasures are collected in a to-do list and each one is assigned to a specific person who is in charge of resolving the issue. This systematic documentation is filled out using Sharepoint sheets or proprietary systematic documentation tools.

- Management dashboard documentation is filled out to track the overall product and process maturity on a quantitative basis, as shown in Figure 3.11.

PV Workshop Managers	Process planners	Product engineers	MTM time & time-measurement specialists	Workers
Planning engineer for assembly processes	Planning engineer for plants	R&D product optimization experts	Logistics experts	R&D responsible for product & packaging
Ergonomics experts	Plant automation engineer	Handling device experts	Standardization experts	Safety experts
Plant responsible	Tooling experts (screw technology)	Technical operators		

Figure 3.12: Roles of stakeholders taking part in PV workshops

3.3.4 *Participants & Stakeholders*

PV workshops consist of 10 to 30 interdisciplinary participants. Each person is assigned to a certain role, according to their expertise. Since PV workshops have to find trade-offs for multi-objective optimizations, there cannot be an optimal solution for all goals such as efficiency, throughput and flexibility [54]. These moderated workshops bring together all stakeholders of automotive manual final assembly. As the focus changes throughout PV workshops, participating roles change accordingly. Figure 3.12 depicts all roles joining PV workshops.

During all workshops, PV workshop managers act as moderators. Process and product engineers also always take part. Domain-specific experts, such as ergonomic specialists or MTM time specialists, are invited whenever the development project reaches a certain milestone and these verification criteria are discussed. In DMU-based PV workshops, a technical operator is present to provide support so that production engineers do not have to manipulate the virtual scene interactively while discussing the verification criteria.

The roles of the stakeholders benefiting from the framework and methods are shown in Figure 3.12.

3.3.5 *Verification goals*

In this section, the verification goals of PV workshops are presented and clustered. PV workshops follow the verification goals presented by UGS: "maintain alignment between evolving product and manufacturing process definitions," "accurately defining and documenting manufacturing processes for all product variants," "validating assembly processes" and "validating manufacturing facilities" [23]. For each verification criterion a success criterion is defined [25, p. 292] to en-

able transparent status tracking. In industrial practice, findings in PV workshops are manifold, as the following practical examples show:

- **Process flaws:** Alphanumeric work task descriptions tell the worker to clip in 5 times a part, whereas CAD data of the connected part has 7 clips points. The process or product has to be adapted.

- **Product flaws:** CAD data quality is not sufficient for the assessment of assemblability. CAD product documentation requires a higher degree of maturity.

- **Sequence flaws:** Mounting an assembly part in the specified sequence is not possible due to overlapping edges of previously installed parts. Manufacturing sequence must be adapted.

- **Tool accessibility:** The assembly part cannot be installed since the screws cannot be reached with a screwdriver. Manufacturing sequence or product mounting concept must be adapted.

- **Assembly part trajectory:** Assembly part cannot be installed due to limited and restricted clearance situations in the respective assembly state of the product. Manufacturing sequence or assembly part accessibility must be adapted.

- **Ergonomics issues:** Parts are too heavy for manual labor, mounting points are not visible or too much lateral bending needed. A handling device must be added to support the worker.

The following assessment criteria are clustered into five categories and success criteria are formulated for each entry: product, human, process, resource and logistics-related assessment criteria. Each cluster shows the geometric representations that are required to achieve the assessment goal. This geometry may contain product information, human information, factory information, station layout information, resource information and logistics information. Geometric information can be represented either physically or virtually. In early stages, fitness "for assembly of separate parts can be evaluated as part of the design process, the complete vehicle production sign-off requires a comprehensive assessment of all production steps in the correct order and according to different criteria"[25, p. 291].

3.3.5.1 *Product-related verification criteria*

In the literature, product-related verification criteria are also called "design review process" [34]. Weber proposes multiple product-related verification criteria:

- "Fitting: Provision of obvious and unambiguous fitting paths

Table 3.1: Product-related verification criteria

Task ID	Assessment Goal	Success criterion / KPI
P1	Static packaging	Assembly is collision free
P2	Dynamic assembly paths	Collision free assembly paths given
P3	Product maturity	Product data maturity reaches quality gate requirements
P4	First time fixture	Part fixture is given immediately
P5	Unambiguous mounting	Part can only be assembled in one way
P6	Integration of mounting elements	Assembly part contains mounting elements
P7	Auxiliary material	No additional auxiliary material needed
P8	Tolerance concept	Verified assembly tolerances
P9	Standardization of assembly	Standardized assembly concepts for fasteners & screws
P10	Confusion immunity (Poka-Yoke)	No parts can be confused

- Fixing: Tool accessibility, direct visual or acoustic feedback confirming correct fitment etc.

- Joining: Facilitate joints (bolts, clips, rivets, welding, adhesives etc.) by provision of the respective part geometries [25, p. 289]."

In accordance with the concepts of concurrent engineering (see Section 3.1.1), manufacturing planning departments simultaneously reveal optimization potentials within the product itself (see Table 3.1). For virtual assessments, the product geometry is required, DMUs of auxiliary material are helpful.

3.3.5.2 *Human-related verification criteria*

Falck et al. state that early identification of ergonomically poor workstations can result in enormous cost benefits in the automotive industry [87]. Weber describes the human-related verification criteria as the "Alignment of a parts weight and shape with manual handling requirements; prevention of possible injury, prevention of contact with toxic or noxious substances etc." [25, p. 289]. Additionally, besides the intrinsic motivation of ensuring employee health and well-being, legal requirements must be met in order to provide a safe and ergonomically proper workstation. Table 3.2 lists human-related verification criteria assessed in PV workshops. For virtual human-related

Table 3.2: Human-related verification criteria

Task ID	Assessment Goal	Success criterion / KPI
H1	Overall ergonomic process	Positive ergonomic score
H2	Overhead assembly	No overhead assembly needed
H3	Reachability	For 5% and 95% population positive
H4	Body forces	Avoidance of critical body forces
H5	Hand/finger forces	Avoidance of critical finger forces
H6	Ergonomic postures for picking and working	EAWS compliant postures, low repetitions of critical body postures (bending, asymmetric postures)
H7	Visibility of mounting points	All mounting points can be seen by the operator
H8	Walk paths	As few walk paths as possible
H9	Part weight and size	Part can be handled ergonomically

assessments, the product geometry and an animated Digital Human Model (DHM) is required, DMUs of auxiliary material and cell layouts are helpful.

3.3.5.3 *Process-related verification criteria*

As detailed process planning is an error-prone task without having a mock-up available, validation is crucial for avoiding mistakes. Table 3.3 lists process-related verification criteria discussed in PV workshops. Weber sums up process-related tasks with the alignment of a part's weight and shape with manual handling requirements [25]. For virtual process-related assessments, geometric representations of the product and cell layouts are required.

3.3.5.4 *Resource-related verification criteria*

Table 3.4 lists process-related verification criteria discussed in PV workshops. Mainly the completeness and effectiveness of resources at a workstation are assessed. For virtual assessments, mainly the product geometry and DMUs of the auxiliary material are required.

3.3.5.5 *Logistics-related verification criteria*

Table 3.5 lists logistics-related verification criteria assessed in PV workshops. Weber sums up these verification goals as checking the usability of existing transport equipment, prevention of transport damages, minimization of required transport space, etc. [25, p. 289]. For virtual

Table 3.3: Process-related verification criteria

Task ID	Assessment Goal	Success criterion / KPI
PR1	Engineered hours per vehicle	Maximized percentage of value adding tasks
PR2	Manufacturing value	Low manufacturing value, reduction of unnecessary tasks
PR3	Line balancing	High degree of capacity utilization
PR4	Integration of value adding tasks into walk path	High degree of parallelization
PR5	Completeness of material	Required parts are available
PR6	Single point provisioning	One part one picking point
PR7	Independent processes	No dependencies between operators
PR8	No intermediate material handling	Part is picked and assembled in the same process
PR9	Impact on product variance	Analysis has been carried out
PR10	Drift optimization	No interference of operators
PR11	Information for operators	Documentation provided
PR12	Two-handed processes	No need to use two hands
PR13	Material assigned	All parts are assigned to work tasks
PR14	Quality of work task descriptions	All PPR data is linked and work task descriptions are easy to understand

Table 3.4: Resource-related verification criteria

Task ID	Assessment Goal	Success criterion / KPI
R01	Equipment available	No additional equipment required: Assembly belt, -box, -trolley, handling devices, screw drivers, torque wrench or scanners
R02	Equipment handling	equipment is suitable for processes
R03	Accessibility for tooling	Tools can easily reach the handled parts

assessments of logistics-related verification criteria, geometric representations of the cell layouts and the auxiliary material are required.

Table 3.5: Logistics-related verification criteria

Task ID	Assessment Goal	Success criterion / KPI
L01	Cell layout & walking paths	Overall optimized
L02	Material supply	Material supply documented
L03	Shopping cart	Working procedures for kick in & -out and conveyor link
L04	Racks and shelves	Only standardized racks applied
L05	Cell equipment	All cell tools available, e.g. trash

4

EVALUATION OF STATE-OF-THE-ART VALIDATION METHODS

This chapter reviews state-of-the-art collaborative validation methods for manual assembly processes in the automotive industry. As described in the domain analysis in Chapter 3, validation methods and tools have limitations as they are not fully suitable for virtual production validation. Several approaches for efficient production validation can be found in the literature [27, 88], material provisioning in final assembly lines [78] and varying degrees of physical and DHM [15]. Only few papers mention drawbacks in current state-of-the-art PV workshops. The whitepaper "general assembly manufacturing" [23] published by Siemens UGS PLM summarizes the automotive companies' barriers for effective manufacturing planning as follows:

- "Inadequate synchronization between product and manufacturing engineering

- Wasted time and effort due to inadequate management of manufacturing data

- Insufficient ability to optimize and validate critical aspects of the manufacturing process prior to launch

- Inability to work collaboratively and manage change within a shared context"

These deficiencies are in accordance with Walla [27] and Seiffert & Rainer [89, p. 28], who also mention multiple limitations on the usage of virtual assessments. These include a lack of simulation techniques, non-standardized data pipelines, high effort for simulations and insufficient data quality.

A contextual inquiry study is described below. On-site attendance at PV workshops are carried out to substantiate literature findings with a focus on characteristics, drawbacks and optimization potentials. This also includes semi-structured expert interviews used to gather qualitative data on state-of-the-art in industrial practice during PV workshops. This contextual inquiry study follows the research questions presented in Chapter 2. Collecting qualitative data is relevant for understanding whether the drawbacks presented in the literature (e.g. UGS [23]) and Section 1.4 still apply in today's assembly planning.

The following chapter is structured as follows: First, the contextual inquiry study procedure is presented, based on qualitative data acquired from on-site attendances and subsequent interviews. Then, the

results are presented using clusters: Physical and virtual PV, methodological challenges, technological challenges and an outlook.

4.1   STUDY PROCEDURE

The contextual inquiry study was carried out by attending PV workshops at Daimler AG in Sindelfingen and Böblingen, Germany. Subsequent semi-structured one-on-one interviews were conducted in quiet separate meeting rooms.

Each participant planned enough time in advance so that the interviews could be conducted in a timely unrestricted manner. All interview partners were encouraged to talk about their observations on specific topics, and their personal notion and optimization potentials for PV workshops. Additionally, participants were informed that they could stop the interview at any time and that the interview was recorded, transcribed, coded, evaluated and published subsequently.

The interview started with a short introduction to this doctoral thesis and the research context (see Chapter 2). The participants were told to disregard organizational barriers within their personal organization (Daimler AG and its departments) and to answer the questions to the best of their knowledge about the current state of the art in manufacturing, but not with respect to company-specific processes. They were encouraged to think about novel use cases, future technologies, methods and processes in the context of PV. This contextual inquiry study was carried out in the timespan from August 2014 to October 2014.

4.2   PARTICIPANTS

These users came from a representative population for PV workshops, as they were actual key users and managers. Their roles were virtual technology experts, PV workshop managers and PV experts (see Participants and Stakeholders in Section 3.3.4). They intentionally came from multiple departments, different ages and heterogeneous Virtual Technology (VT) experience, but all of them had profound knowledge about the PV processes.

The participants were recruited via an e-mail or telephone invitation. They were all participants of PV workshops but with differing backgrounds. Production engineers represented the main stakeholders, as they were in need of the simulation results. Virtual technology experts had experience in authoring and carrying out virtual simulations. All participants took part on a voluntary basis and did not receive any extra reward.

At the beginning of each interview, the participants were asked to characterize themselves with respect to their self-proclaimed VT

Table 4.1: Interview participants and their respective background, characteristics and technology affinity

ID	Background	Organiz. unit	Gender	Age*	VT experience*	Duration [min]
T1	VT Expert	Production Planning	Male	38	4	0:29:47
P1	PV manager	Production Planning	Male	46	2	0:36:33
T2	VT Expert	Digital Factory	Male	51	5	0:28:40
T3	VT Expert	IT department	Male	31	4	0:15:55
T4	VT Expert	Prototype Building	Male	57	3	0:42:03
P2	Process expert	Production Planning	Male	31	5	0:43:26

*optional items, P1 is a manager

experience on a 5-point Likert scale, where high values describe high pre-knowledge on VT.

Table 4.1 shows an anonymized list of interview partners, their individual background, organizational unit, their management background, if any, gender, age, self-proclaimed experience in virtual technologies and the interview duration. Their IDs represented the category of experts, where P stood for production engineers and T for technology experts.

4.3 DEMOGRAPHICS

All interviews were condudcted in German. The results were translated and are presented in English here. The mean age of the participants was 42.33 years (SD = 9.83) and mean self-proclaimed VT experience was 3.83 (SD = 1.07) on a five point Likert scale, where 5 stands for maximum experience. All interview partners were male. Four VT technology experts could be acquired for the interviews, one PV workshop manager and one process expert. One out of six interview partners was in a management position.

4.4 STUDY RESULTS

The results were evaluated using qualitative research methodologies using the software MAXQDA Analytics PRO 12 made by VERBI. A total of 3:16 hours of audio were recorded with six interview partners. Partly colloquial answers by the participants were intentionally

included in the results section to properly reproduce the stated opinions. The coding process went through an iterative process of continuous optimization of the coding taxonomy. The codebook was discussed with experts afterwards for iterative optimization and clustering.

The coding process revealed three major topics the experts discussed: Current physical and virtual production validation processes, organizational challenges and technological challenges.

4.4.1 *Physical and virtual production validation process*

All experts (6) commented on current state-of-the-art production validation processes in manufacturing as a baseline.

General process descriptions: The validation process starts approximately two years ahead of SOP and lasts until six months before SOP (P1). In the early phases, DMU-based assessments are carried out as parts are altered on a daily basis and assembly concepts are highly volatile. Later on, mostly physical workshops are carried out (P1). PV workshops are carried out at certain milestones of different planning phases during the PDP, where all data come together. For each part or process, only a restricted time frame is allocated (T4). A typical workshop duration is limited to three weeks for assessing several thousands of assembly parts or processes (T4). This only allows a quick validation at the end of the workshops, but does not leave time for planning (T4). Workshop participants expect to get a general idea of the upcoming product and processes and have to find countermeasures for product and process flaws in a collaborative manner (T2). The PV workshops are often the first time that all experts, such as logistics experts, ergonomic experts, time experts, planners and product development engineers, meet (P2). The hands-on experience of workers is highly valuable, since they detect many effects that production engineers may fail to see (T4).

(T4) requests an **overall strategy** for DMU and PMU-based reference cars. Reference cars maintain the same extra equipment configurations throughout the PDP. According to (T4), reference cars should be matched between DMU and PMU validation phases so that the same configurations can be re-assessed. People get used to the configuration of repeatedly used reference cars, and this enhances the re-use of digital simulation contents (T4).

Product related assessments: Static DMU product geometry is assessed on large screens or small power-wall projections, in order to optimize the product, manufacturing sequence and processes (T3). (T4) reports that DMU product geometry is sometimes not available, which causes misunderstandings for workshop participants. He points out that when people see the parts either in the physical or the virtual world, they have a common discussion basis. (T1) mentions that

digital collision checks, static assembly and disassembly assessments are already partly assessed in the virtual domain. Standardization of mounting and assembly concepts are assessed in buildability as well (T4). (T2) adds that DMUs have to be as new as possible to satisfy the demands of PV participants. (P2) sees the huge advantage that the DMU simulation scenes can be altered quickly, whilst this wouldn't be possible using PMUs.

(T4) wants **closer cooperation** between manufacturing planning, packaging, prototype building, R&D and product engineers even prior to the PV workshops. This would simplify the workshops at the end of each phase in PV. (T4) regrets that within buildability assessment workshops, many errors are revealed that could easily be detected earlier in the development process. He wonders if there is too little time to create high quality DMU and if there would be a higher reliability character if physical prototypes would have to be built, since there are higher costs involved (T4). (T4) assumes that the time and cost involved tend to be underestimated for DMUs.

Process related assessments: In the early phases of DMU assessments, no work task descriptions are available. A rough manufacturing sequence is created for workshops (T4). Only when alphanumeric work task descriptions or rough sequences are available is process validation (P1). Even though the central purpose of PV workshops is the collaborative validation of existing planning data, people sometimes start to plan during the PV workshop sessions (P2 and T4). Processes are refined on-the-fly during the assessment of the product, and additional process plans are written or changed (P2).

Human related assessments: (T3) argues that ergonomic assessments can be carried out in virtual space already, but with a lot of effort. (P1) supports this by adding that tracking systems are sporadically used for capturing assembly and disassembly trajectories of parts to check clearances and accessibility in the manufacturing state. Feasibility studies are carried out, such as clearance assessments between product and human or between product and tools (T3). (T4) remarks that the assessment verification criterion "clearances" is important for assembly, e.g. whether there is sufficient space to access the product by hand or with tooling. Mostly human-related movements are evaluated for value adding tasks (P2). An example of an reachability assessment is if a 5%-sized worker from a representative population database can reach a screw point ergonomically.

For non-value adding tasks, physical validation methods are typically used (P2). For example, walk paths have never been assessed with other methods than by putting duct tape on the floor for a "spaghetti diagram" or pen and paper methods (P2). Even though it would be possible to **determine times** in virtual space, this is not carried out at all. Predetermined Motion Time System (PMTS) must be used in European automotive industry for task planning (see also

[90], such as MTM-1, MTM-UAS predominantly used in Europe or MODAPTS in the US), instead of time measurements (T4) (see REFA [91]).

Cell layout-related assessments: The factory layout is valuable for visualizing the context for all workshop participants either in the digital or physical domain (T4). (T4) points out that there have been "infinitely long" and controversial discussions on the work heights in relation to the car, if there is no visual reference. This leads to the requirement to visualize and simulate work heights for each cell. This would narrow down the discussion to an objective simulation.

Logistics related assessments: (T1) states that logistics-related topics are still carried out in physical space, especially when human motion is involved.

4.4.2 *Organizational challenges*

Overall, all experts (6) mention challenges in the current methodology and processes of production validation.

Hindrance factors for digital validation: One expert remarks that production engineers have heterogeneous technological backgrounds. Some are **technophile** whereas others are not accustomed to virtual methods and constantly try to avoid them (T1). He suggests making virtual methods look as real as possible to convince production planners to lose their fear of virtual methods. (T4) sees greater technophilia for younger colleagues, whereas older production planners tend to use more PMU-based validations (T1). Another hindrance factor can be found in the overall **costs for PDM systems**. Production planners do not apply PDM systems widely, due to the high costs (P1). He mentions that only 20% of production engineers have installed the native PDM software. In addition, PDM systems require a **high expertise** in handling virtual models (P1). (T4) states that production planners often **lack access to DMU representations** of their parts. Even though PV workshops are intended for final assembly validation, there some production planners who see their parts as a DMU for the first time. This clearly shows the importance of **reducing the entry barriers**. (T4) adds that the general **availability of 3D hardware** must be ensured, besides the broad availability of software and training.

(P2) points out the advantage of DMU-based workshops in comparison with PMU-based workshops, which is that **process variants** and **product variants** can be altered instantaneously, whereas using PMUs, people always lose time retrieving some parts from the warehouse. "The speed in which solutions are proposed or process variations are simulated, is of course higher in the virtual space, but all people require a common understanding of the virtual scene" (P2).

Five experts commented on the topic of **entry barriers**. (T3) distinguishes two major work steps where entry barriers apply: Authoring and live-simulation.

For the **authoring of 3D scenes**, all interview partners regard this work as an expert's task, even with optimizations in the near future. (T1) is displeased as he sees the preparation of virtual scenes as unnecessarily time consuming. (T2) mentions the idea of a One-Click export functionality for 3D scenes, such as a 3D PDF exporter. (P2) remarks on the data provisioning process and introduces the problem that fusing heterogeneous data out of the different planning databases or systems into one holistic virtual scene is challenging by now. Even though it is obviously needed, there is no tool to view 3D layouts integrating all DMU parts, DMU tools, DMU handling devices and DMU material zones.

According to all experts, for **live simulation in PV workshops** the entry barriers can be reduced so that technical experts are no longer required to use the system. (T3) wants to optimize the assessment software User Interface (UI) in a way that unskilled production planners are able to utilize the software. (P1) agrees with (T3) and states that previously prepared live simulation during PV workshops could be carried out by production planners on their own. In contrast, (T2) and (T4) state that both authoring of virtual scenes and interactive simulation during PV workshops is an additional qualification and should not be conducted by production engineers even in the near future.

(T4) remarks there are currently only a few people who can execute such interactive simulations. He wants to simplify the usability of these PDM and production validation systems. (P2) does not want the planners to have to attend a "2 weeks training" in order to use the systems. (P1) and (P2) agree that conducting live simulations should remain the technical operators' job, so that the process experts can focus on their native task, namely validating the product and processes. Similarly, (T2) argues that VT systems can be used more efficiently by VT-experts, and production planners should stick to their original work. (T4) and (P1) agree that even the common PDM system has so many functionalities that it continues to be an expert system which can be used by some people but not all. (T1) remarks that the planners have to use so many heterogeneous planning tools, issue tracking tools and simulation systems that they are often unwilling to learn and utilize an additional and even more complex VT simulation system.

Variance: As described in Section 1.3.1, the increasing product variance and extra equipment is a potential source of errors. For purely DMU-based assessments (T4) points out that theoretically multiple variants can be assessed simultaneously, but this is not currently carried out. Supporting this, (P1) remarks that for hardware-based val-

idations it is hard to simulate multiple variants. For example, in a powertrain there are hundreds of combinations as the tooling frame must be built for various axles, all-wheel-drive systems and various drive shafts. This process includes so many variants that even the hardware equipment and tooling cannot be validated properly.

Ergonomic assessments: (P2) remarks that the capability of full-body motion capture is there but is not utilized frequently, because it is too cumbersome. Typically, this marker-based tracking system is used in combination with DELMIA V5. He sees great opportunities for virtual ergonomics assessments if entry barriers are reduced. If motion capture technology is too cumbersome, many ergonomic assessments are postponed until a PMU is available (P2). (P1) also wants to utilize motion capturing more frequently for ergonomics assessments to obtain early validation results. But he remarks that "authoring is so work extensive, that there is only the possibility to validate 2-3 worker processes a day." "This is simply too few, because in regular PV workshops 300 to 400 worker processes are validated a day." "This is why the method [marker-based motion capturing] is sufficient for the validation of very few critical processes, but not generally throughout the validation process for every work task" (P1).

Documentation: For documentation purposes in PV workshops, multiple tools are used. Checklists and action lists are generated in tools such as Microsoft Excel, Microsoft SharePoint (P1) and proprietary status tracking databases (P2), (T2), (T4). For additional media content, clarification screenshots of the rendering are merged in Microsoft PowerPoint slides. (P1) comments that images lack the ability to be modified afterwards in order to re-check contours, clashes, or add tools. Videos, 3D trajectories and motions are currently not recorded.

Collaboration between PV participants: The experts distinguish collaboration in PMU workshops and in DMU-based workshops. (P2) explains that the "collaboration in PMU-based workshops is really good" as everybody is able to participate actively in the creation process because there is no technical barrier (P2). Workshop participants stand next to the PMU and build the product collaboratively. Everybody can participate, they can talk to each other and show the others "do it like this, do it differently, look here" (P2). In DMU-based PV workshops, where all participants sit in the same room, there is "too little movement," which (T2) supposes leads to passive involvement in PV workshops. Continuous bustling of people is helpful and represents a key factor for workshop success. What applies to all types of assessment workshops is that collaboration must be supported by generating a common understanding of the current verification criteria (P2). (P2) points out that each workshop participant has his own verification scope, such as logistics or assembly planner, but all should work together on a robust and overall solution.

4.4.3 *Technological challenges*

Multiple codings regarding technological challenges were mentioned:

Interaction: There is no direct interaction between the DMU and the production engineers as they have to ask the operator to move or manipulate the virtual objects (T1). "When somebody wants to see something different or from a different view, they have to ask the technical operator to change it for them" (T1). "This might reduce the workflow a little" (T1). (T1) suggests that the planners should be able to manipulate the scene on their own to support collaboration between planners. "For example, when they talk about a fixture in a product, it is much harder to verbally describe it, than to show it in the actual CAD model" (T1) They should take the tracked part, manipulate it, rotate it and show the specific property on their own.

Immersive assessments: Regarding immersive visualizations, the use of VR and AR is seen as helpful as production engineers may get a better idea of the product and it "is highly welcome and helpful" (T1). (T4) wants general availability of VR components for all production planners' work desks, so that everybody has the possibility to immersively visualize and interact with the DMU components (T4). (T1) remarks that immersive assessments using Head-Mounted Display (HMD)s have to overcome the barriers of embarrassment and expert barriers. They are afraid to use systems they are not accustomed to so as not to embarrass themselves. As production planners have a different affinity to virtual technology, he suggests visualizing 3D environments as close to reality as possible. In contrast, VR could simplify interaction with the 3D scene since interaction metaphors are closer to reality than when using regular HIDs (T1). (T4) adds the idea of extending user interfaces to include gestures so that the users can manipulate the scene with hand gestures. (P1) proposes the use of AR to support the validation of multiple product variants in hardware-based workshops. (P1) additionally points out that a closer connection between PMUs and DMUs would support PV workshops "enormously." (P2) describes a real-life problem which could be solved using AR technology. PV workshops are held even though some physical assembly parts are missing. Nevertheless, because of the missing parts, some processes cannot be assessed. With the use of AR, at least these missing parts could have been augmented.

Data related challenges: (P1) states that data is available for product CAD data, racks and carriers but not for cell layouts. There is still no standardized, interoperable exchange format for material zones, factory layout data, rack layouts etc. Additionally, there is no central database for this kind of data and no tool to collect all data in one place (P2). Outdated or non-synchronous data maturity leads to errors in validation results (P2). (P1) argues that the overall goal is not only the photo-realistic rendering of 3D geometry but more emphasis

should be placed on the manufacturing context. (P1) wishes to have meta information, which cannot be seen in the sole 3D geometry. For example when highlighting a screw, the torque and type could be automatically be visualized using text.

Flexibilization and physics: All (6) interview partners mentioned that for future assembly simulation, flexibilization and physics are important topics. All experts need physics simulations, such as collision detection and avoidance, gliding objects and gravity. (P1) sees the absence of flexibilization techniques and haptic simulations as a huge hindrance factor for broader use of virtual assessment simulations.

Visualization of CAD models: (T4) remarks that the huge product variance requires PV workshops to offer the possibility to show multiple visualizations in parallel to compare manufacturing states. Switching CAD-data models has to work instantaneously so as to not lose the mental model of different variants. "It makes a huge difference whether rendering takes half a second or three minutes for switching the product variants. The longer switching takes, the more likely people lose the impression of difference" (T4). This matches the requirements of (P1) who complains about the long initialization times when loading a new car model in "10 to 15 minutes", which is "no longer acceptable nowadays." The experts disagree with regard to stereoscopic visualization. While expert (T2) explicitly says that it is not helpful in workshop situations, others see it as a chance for better spatial understanding (T1). "The perceived depth is helpful for the workshop participants to estimate distances of interactively manipulated objects" (T1).

4.5   DISCUSSION, STUDY LIMITATIONS AND SUMMARY

Regarding the results presented in the context of the literature review, large agreements can be found. The white paper on "general assembly manufacturing" [23] by Siemens UGS PLM and the deficiencies presented by Walla [27] match the contextual inquiry study results:

- Synchronization between product and manufacturing engineering is still an issue because of missing digital simulation capabilities, technological restrictions, heterogeneous data sources and differing quality gates in PDP.

- Handling manufacturing data continues to waste time and means more work for manufacturing engineers.

- Optimization opportunities for the "insufficient ability to optimize critical aspects of manufacturing process prior to launch"

- Optimization potential for collaboration in shared contexts

- Optimization potential in time planning and line balancing which offers enormous optimization potential for reliable estimates of manufacturing costs.

The findings of this contextual inquiry study have to be regarded in the context of time and cannot be generalized as they summarize the state of the art for one company in 2014. Nevertheless, the deficiencies found in the literature are confirmed and are used as a baseline scenario for this doctoral thesis.

In this chapter, expert interviews were presented. The results revealed methodological and technological limitations in PV workshops. Moreover, the experts indicated optimization potentials.

In the following chapter, these optimization potentials are implemented with a holistic framework for virtual assembly validation to support collaborative PV workshops.

5

THE VIRTUAL MANUFACTURING STATION FRAMEWORK

This chapter presents a framework for co-located, collaborative assessments of manual assembly processes. This framework integrates several interactive components in a methodology for professional application. In the following, this framework is referred to as the VMS. Hence, this comprehensive set of virtual, mixed and augmented reality methods is intended to be used for interactive, real-time assessments of manual assembly processes in PV workshops. Therefore, the VMS must represent a collaborative workshop environment. It unifies basic research interaction concepts with large scale high-resolution output technologies and simulation software components from a variety of domains.

The research contribution of this chapter is the determination of gaps in the literature for collaborative, co-located workshops. Publications in the areas of computer supported collaborative work, production engineering and human computer interaction are systematically analyzed and put into the context of real-life automotive production validation. A framework for the VMS objectives as well as required further research are defined in applied and basic research domains. Based on the literature review, a holistic concept for the VMS is created. In Chapter 4, the requirement analysis revealed deficiencies in current virtual and physical PV workshops such as improvable collaboration between workshop participants and missing interactive features. These issues are addressed in the VMS's development plan in accordance with the current literature.

> **Related publications**
>
> Research results presented in this chapter have been developed for this doctoral thesis. Nevertheless, multiple parts of this chapter were presented in the conference paper "Dual Reality for Production Verification Workshops" [80] by Otto et al. at the CIRP CATS 2016 conference. The goals and concepts presented have been extended and revised in comparison to the peer-reviewed publication. In addition, parts of this chapter were presented in "Motion Capturing" in the 2017 Springer book: "Web-basierte Anwendungen Virtueller Techniken" [92].

This chapter is structured as follows: First, the objectives for the VMS are defined. Since all objectives are interlinked, they are clustered into three different layers of objectives and their interdependencies are discussed. Secondly, the objectives lead to the VMS's key

properties, which the novel Collaborative Virtual Environment (CVE) intends to reach. All developments in upcoming chapters will contribute to the key properties presented as the VMS concepts also focus on practical aspects such as workflows in CVEs for workshop environments. These VMS key properties and goals are described in the context of relevant publications in the literature. The proposed VMS system setup is presented subsequently, including specifications for hardware and required simulation software components. Finally, the general evaluation possibilities of the overall concepts are discussed using an evaluation matrix.

5.1 OBJECTIVES

The VMS is based on a complex, multi-layered objectives environment. Three different categories of objectives have been deduced. All of them are interlinked to one another and therefore form a utility chain: Company-scope objectives, objectives of digital simulation methods and objectives of the VMS. Kunst et al. present a model for calculating the profitability of VR systems [93] and summarize such utility chains accordingly as discussed in Table 5.1. Kunst et al. state that strategic objectives are difficult to quantify, have a long-term effect horizon and are difficult to assess, whereas in the operative organizational unit, utility can be quantified directly and utility can be revealed on a short-term basis. Table 5.1 has been adapted to the following VMS objectives [23].

Table 5.1: Analysis of utility categories (adapted from Kunst [93, p. 11])

Utility category	Company scope utility	Digital simulation utility	VMS's utility
Company scope	Strategic	Tactical	Operative
Quantifyability	Not directly quantifyable	Calculable	Directly quantifyable
Time horizon	Long-term	Medium-term	Short-term

Based on Kunst et al., three categories of objectives are depicted in Figure 5.1. In the following, each category of objectives is presented in the respective order.

5.1.1 *Company scope objectives*

One of several company goals is to produce maximum quality products at a minimum total amount of costs. For production planning, this means that the total amount of costs and production time must be reduced despite higher changeability, flexibility and quality in the

Company-scope objectives	Objectives of digital simulation methods	Objectives of VMS
Reduction of total amount of costs	Less problem-solving cycles	Multi-user visualization and interaction
Reduction of total amount of time	Higher data quality	Seamless integration of PPR data
Higher changability	Less PMUs required	Workshop collaboration performance
Higher flexibility	Enabling late product changes	Efficient processes and methods
Higher quality	Reduction of overall PDP timespan	Flexible degree of virtualization

Figure 5.1: Interconnected objectives: From company-scope objectives to objectives of the VMS

production systems (see Figure 5.1). Production systems have to become more changeable (e.g. faster changes in production facility) and more flexible (e.g. output quantity). Product related costs caused by errors can be reduced via higher planning quality. Such company-scope objectives of automotive OEMs have been broadly discussed in the literature (see [54, 94]).

Advanced virtual simulation methods promise to reach these goals earlier at a higher quality and thus lower costs. We present a comprehensive set of virtual and augmented reality methods for real-time assessments of manual assembly tasks during interdisciplinary PV workshops.

5.1.2 Objectives of digital simulation methods

Reducing the amount of PMUs: Building and testing physical prototypes of new cars is one major cost factor in vehicle development [25, 78]. Automotive OEMs thus generally try to reduce this cost-intensive development stage of building prototype cars. "Physical prototype building and testing is one of the major cost factors in vehicle development. This cost increases with the number of models." [95]. Enhancing virtual simulation methods for usage in virtual assessments would allow for building fewer PMUs. Even though, PMUs are still required in the very late stages, the overall number of physical prototypes can be reduced. Late prototypes can already be produced using final production rigs, making it possible to disestablish early iterations of development rigs. In early 2017, the automotive OEM Tesla announced that they want to completely skip the hardware production line verification stage and only build a small amount of hardware prototypes in order to reduce costs as well as speed up the production process by virtually simulating all production phases [96]: "Tesla, however, is skipping that preliminary step and ordering permanent, more expensive equipment as it races to launch its Model 3 sedan by a self-imposed volume production deadline of September [2017]." These virtual assessment methods are labeled "advanced analytical

techniques" by Tesla. "Therefore, reducing the number of physical prototypes is a key aspect of future competitiveness" [95].

Reducing the overall timespan of the PDP: Shorter development cycles in the automotive market are necessary to achieve greater flexibility and mutability in the demand context of mass-customization, quicker product updates and changing market demands (see Weber [25]). That is why it is necessary to achieve a higher quality for both the product and production planning during the early stages of the PDP. This process is called frontloading. As a side effect of shorter development cycles, the overall development costs drop.

Higher data maturity: To obtain shorter development cycles and fewer product related errors, the DMU data quality must improve. This problem has also been widely discussed in the literature (see Manns et al. [66]). In the early stages, DMU data quality often is not sufficient for the efficient verification of production criteria. Manufacturing information must be created from DMU data in order to get PPR information updated synchronously. For example, DMU data already geometrically models product screws properly. The corresponding process information often is still built manually because only little manufacturing information is linked to the DMU data. Ideally, changes in geometric DMU data would also cause automatic updates in process data if this information is linked properly.

Enabling late changes: Even if the product development cycle cannot be accelerated for certain products, the use of virtual validation methods does offer the possibility to enable late changes at lower costs. The later physical prototypes are built, the less investment is needed and the more flexibility a product has [25, p. 89].

Smooth ramp-up: On-time launches and smooth ramp-up processes can only be guaranteed if processes are robust from the first day of production. The ramp-up process aims to achieve the expected throughput of the new production line as quickly as possible.

Mixed model line: As the market demands more and more mass customization and diversification of products, many automotive OEMs do offer more models, variants and derivatives with additional extra equipment. "Mercedes-Benz went from offering nine models in 1993 to expecting to offer 32 in 2015" [95]. This rising diversification places additional requirements on production planning in order to integrate multiple products in the same production line while also ensuring high profitability. The greater the time spread of a production cycle time and the more diverse work contents is located at a single workstation, the more complex planning becomes. Even when integrating the succeeding product in an existing production line, line balancing must be optimized since the old and new product must be produced efficiently at the same time. This means that the work contents and materials required can change considerably. This complexity in work

contents cannot be assessed without digital simulation and validation methods.

All of these market demands, company demands and demands for digital simulation and validation tools put pressure on production planning to find novel possibilities to meet these requirements. This is why the VMS framework is presented below to fulfill some of the aforementioned objectives.

5.1.3 *Objectives of the VMS framework*

The consistent implementation of virtual assessment methods and processes is a promising approach to fulfill the aforementioned requirements. The systematic use of more digital assessments and validations can meet the requirements for a smoother ramp-up, efficient mixed model line planning, fewer PMUs and shorter development cycles. Therefore, the overall objective of the VMS is to provide a shared virtual environment for collaborative PV workshop situations:

Reach verification criteria using DMUs: Currently, PMU-based production validation for manual assembly processes is still state of the art with sporadic, static DMU build-ups. Therefore, one objective of the VMS is to integrate new technologies and methods to assess more validation criteria in the virtual domain.

Supporting workshop situations: Validation and verification tasks are executed in workshops. The VMS aims to efficiently present the simulation contents to multiple workshop participants. For shared content, using either PMUs or DMUs, participants have to create a common understanding of all planning aspects: Product, processes, workstation and logistics layouts, tooling, etc. Workshop participants should be enabled to interact easily with the DMUs without additional lead times and without training in virtual environments. If this prerequisite is met, collaborative decision making is possible and fewer problem solving cycles are needed.

Integration of PPR data: PV workshops use many heterogeneous data sources and simulation systems to achieve their workshop goals. Product data is displayed in a dedicated 3D visualization system, whereas process data is visualized in alphanumeric planning systems and workstation layout data is displayed separately in regular office systems. The VMS aims to integrate these data sources and simulation systems in one tool which also comprises process models of real PDM data.

Applicability in all phases of the PDP: It must be possible to use the VMS framework in all assembly validation relevant phases of the PDP. Due to the fact that PMUs should be built as late as possible, fewer PMUs are available during the early phases of planning, even though the PV assessments still have to be conducted. However, in the late phases of the PDP, PMUs are available, which is why operators

and participants must be enabled to use a flexible degree of virtual technologies, such as virtual reality, mixed reality and augmented reality.

Reduction of preparation effort: A reduction in the time required for preparation has two aspects: A reduction in authoring effort and a reduction in technical preparation effort during virtual workshop situations. Due to the system complexity in authoring, PV workshops must be prepared by experts a long time before the workshops begin. An efficient and consistent data provisioning process for the heterogenous data sources helps make authoring easier as well. This lead time must be reduced. When lead times and data freeze times can be set later, data maturity increases.

Reduction of entry barriers: Interactive simulation currently requires expert knowledge. However, the VMS's intended customers do not have vast background knowledge on virtual assessments. This is why complexity during the usage of virtual assessments must be reduced. By lowering technical entry barriers for interactive virtual assessments, production planners can carry out assessments on their own. The need for a technical operator would be eliminated, reducing total costs.

All objectives presented aim to form a dynamic, interactive, virtual replica of future workstations for the assessment of manual assembly tasks - a "Virtual Manufacturing Station."

5.2 KEY PROPERTIES OF THE VIRTUAL MANUFACTURING STATION

Multiple key properties are introduced for the implementation of the VMS framework's objectives. For each key property, basic research is presented and results are integrated and adapted for the VMS framework. The VMS framework covers integrated hardware, software and a methodology for PV workshops.

For the realization of the VMS framework as a collaborative workshop environment, multiple key properties must be met. This applies to the entire automotive validation process, from authoring to holding PV workshops. In the context of CSCW, Szalavári et al. present an architecture for multi-user collaborative virtual environments for AR content [97]. They discuss key properties of such a collaborative workshop environment and call it an "augmented laboratory" that they intend to use for "visualization, presentation and education." In accordance with the "Studierstube" approach, the VMS framework's key properties are presented and discussed in the following chapters with respect to the requirements of automotive production validation.

Figure 5.2 depicts the key properties of the VMS, which summarize the attributes of the VMS framework.
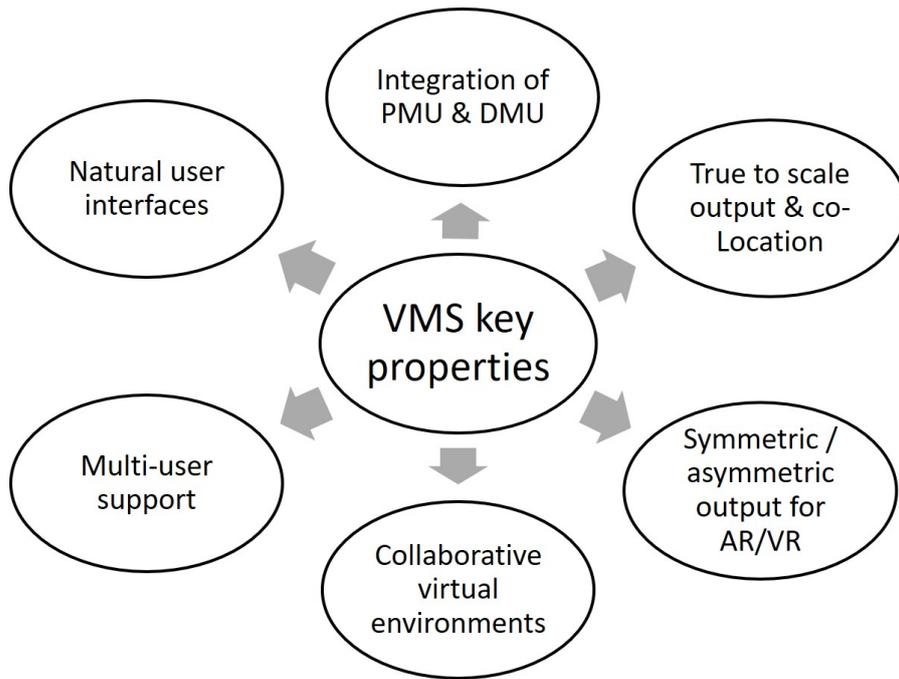
Figure 5.2: Key properties of the Virtual Manufacturing Station

5.2.1 *Integration of PMU and DMU*

As defined in Chapter 4, state-of-the-art PV workshops are held either solely with the assistance of hardware-based prototypes or solely in virtual space but not yet incrementally in the virtuality continuum as depicted in Figure 5.3.
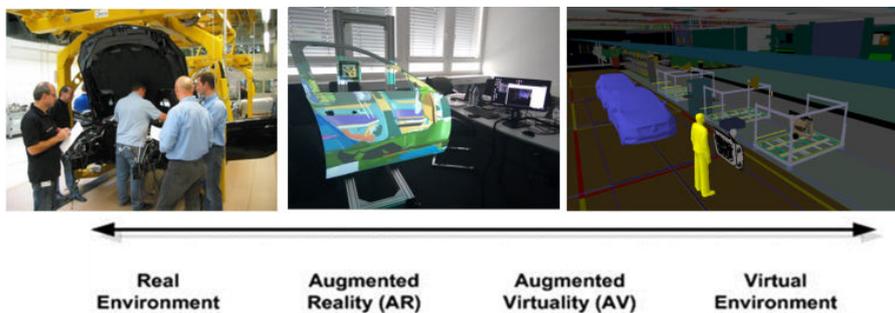
Figure 5.3: Virtuality continuum for production verification workshops according to Milgram and Kishino [98]

The VMS framework offers multiple output variants in the so called "virtuality continuum" [98]. Several increments ranging from the physical domain to mixed and augmented reality to virtual reality are provided. Depending on the state of the PDP, the workshop managers or technical operators can choose the proper input and output modalities for the respective validation task.

For example, in the late phases of the PDP most parts of the car are already physically available. Due to simulation restrictions, the

flexible behavior of assembly parts can only be assessed physically. DMUs are still required because not all extra equipment combinations can be cross-checked due to their enormous variance. Szalavari describes this need for augmentation as follows: "Real-world objects can be augmented with spatially aligned information. This allows smooth extension of real objects with virtual properties in design processes, like variations of new parts for an existing system. Superimposed information can also incorporate enhancing elements for real objects, like descriptions or guidance in training or education situations" [97]. This is why during the PV assessments DMUs are superimposed to switch between all variants of the product. Therefore, a co-existence of the DMU and PMU parts is compulsory for checking variant rich products.

Lifton and Paradiso have extended the concepts of such mixed reality scenarios to the concept of **dual reality** which is defined "as an environment resulting from the interplay between the real world and the virtual world" [99] that inherits the capability to "mutually reflect, influence and merge into one another." Transferring this idea to the manufacturing industry, this leads to various workshop constellations which are held between the hardware-based and the digital domain. These concepts proposed by Lifton and Paradiso [99] share several characteristics and objectives with research on 'cyber-physical equivalence' (compare Stork et al. [100] and Otto et al. [80]). Both research areas have influenced the concept of the proposed VMS framework methods.

For PMU&DMU registration in PV workshops, Arteaga et al. have presented a reference framework for manual assembly systems [15], extending the concepts of Bordegoni et al. [101]. In this framework, Arteaga et al. introduce a concept for three levels in the virtual continuum, namely for workers, objects and interaction. In manual assembly simulation, both workers and objects can be present as real, mixed or digital representations and interaction has two different manifestations: "Visual interaction" and "visual and haptic interaction." This results in multiple assessment variants involving varying degrees of physical parts and humans. For example, they propose a range of assessment types such as "Motion Capture (MoCap) based simulations", "AR based simulations", "purely DHM based simulations", "purely physical based simulations" and "enriched haptics based simulation." The VMS applies these reference framework's concepts in its key properties. Table 5.2 shows an example for choosing a mixed reality simulation with a physical base part, virtual assembly part, physical resources and a simulated worker.

In general, PMUs are increasingly available during PDP (see Section 3.3). PMUs could, for instance, be carry-over parts from predecessor models, 3d printed models or other related products. Therefore, similar PMUs could resemble the planned part in terms of geometry,

Assessment item	Physical representation	Digital representation
Base part (e.g. car body)	X physical chassis	
Assembly parts		X virtual part
Resources	X real screw driver	
Human	X real human	

Table 5.2: Example for mixture possibilities in virtual assessments. The operator can choose from among various purely physical, virtual and augmented reality assessments

weight and mounting. Since the number of combinatory possibilities of assessment elements with digital and physical models is too high, it is not possible to determine a priori which method will best suit the requirements of the specific verification task (see [15], [101]). This gap can be bridged by using the VMS framework which enables workshop managers to employ all techniques instantaneously. Consequently, depending on the assessment scope and availability of PMUs and DMUs, there are multiple possibilities to match verification tasks in workshop situations with physical and virtual assessment methods (see Table 5.2).

Overall, PV workshop participants are more familiar with the use of PMUs. To maintain a familiar work setup, digital content can be superimposed onto PMUs in the VMS without any body-worn AR devices. "Manipulation of the real world models (e.g. its orientation) is more intuitive to support than a purely virtual environment." [97] This enables production engineers to walk around a physical car body, grasp assembly parts and feel the weight of tools.

5.2.2 *True to scale visualization & co-location*

True-to-scale visualization is essential for direct user feedback. With an an orthographic view-independent rendering of the virtual scene in the VMS, experts get a good impression of distances, sizes and speeds. If PMU and DMU are registered to tracking systems at the same time, a 1:1 movement of both the trackable or the human representation is visualized in the digital domain as a digital twin or "cyberphysical equivalence." Stork describes the interaction and the challenges in visual computing of advanced manufacturing [100] for cyberphysical production systems. The VMS can be seen as a realization of the proposed conceptual model of visual computing. True-

to-scale visualization is one aspect of 'equivalence.' Virtual spatial relations match physical spatial representations, which allows for a precise registration of PMU and DMU. Human beings are able to estimate lengths, spatial relations and speeds linearly and precisely (see Chapter 8). Starting the simulation, the participant can easily step into the augmented workspace and verify planned assembly process steps with regard to planned times, position, dependencies, effective work time and the ergonomic aspects of the assembly cell with less effort of interpretation and mental scaling.

5.2.3 *Collaborative Virtual Environments*

The VMS framework supports the collaboration of all participants within the virtual domain. Ishii states in his paper on Tangible Bits that "Interactive surfaces are another promising approach to support collaborative design and simulation that has been explored by many researchers in the past years to support a variety of spatial applications" [102]. The VMS framework enables PV workshop participants to collaborate with the simulation scene. The literature offers multiple recognized definitions of CVEs. According to the "Encyclopedia of Multimedia," Furht defines CVEs as follows:

"A collaborative virtual environment is a shared virtual world that allows its users to collaborate in the synthetic world, performing shared object manipulation and other collaborative tasks" [103].

Furht further describes four axiomatic enablers for collaboration between workshop participants [103]:

1. **Connectedness:** Ensure necessary connectivity between participants.

2. **Awareness:** Facilitate awareness (and discovery) of other participants.

3. **Sharing:** Facilitate sharing and exchange of information between participants. Snowdon et al. support this aspect by stating: "Information sharing is central to collaborative work" (see also [104]).

4. **Communication:** Facilitate dialogue and interaction between participants (above and beyond information exchange). "Artefacts and embodiments also are an essential aspect of communication and this is how the representation of people and artefacts such as documents and tools within CVEs can facilitate communication" (see also [104]).

In order to find overall optimal solutions and trade-offs between PV workshop participants' areas of expertise (e.g. ergonomics vs. time optimizations vs. line balancing), they must share information and

communicate about the virtual assessments being carried out. PV workshop managers are expected to moderate the complex validation tasks and to guide the collaborative efforts in the VE, the same as in the physical domain. Such typical tasks are presented in accordance with the CVE "Studierstube" by Szalavári et al., as they intend to use the CVE for "viewing and examining [of] objects that are not accessible directly or that do not exist in the real world" [97]. Multiple digital and physical information sources must be fused in PV workshops to obtain a holistic understanding of the workstation under assessment. Alphanumeric work task process descriptions, CAD datasets, task documentation lists, process plans, material zone data and motion trajectories are only a few examples of data that are shared and collaboratively interacted on in the VMS simultaneously. "Size, complexity, physical properties are just parameters in a simulation, no longer are they constraints for the analysis" [97].

Whereas Furht focuses on collaboration in the digital domain only, the VMS extends the CVEs to include collaboration in the physical world, since both domains can be blended. Collaboration between workshop participants can take place equally in the physical and virtual domain. For example, in a co-located workshop two users can simultaneously interact with the virtual CAD model while discussing verbally and non-verbally (gestures and facial expressions) in the physical domain. Similarly, Billinghurst et al. present "SharedSpace" as an augmented reality interface for co-located, collaborative computing [105]: "Co-located users can see each other's facial expressions, gestures and body language thus supporting natural face-to-face communication cues. Thus the "SharedSpace" interface allows multiple users in the same location to simultaneously work in both the real and virtual world."

5.2.4 *Symmetric and asymmetric output for AR/VR*

Sharing both the physical and virtual content in a collaborative workshop session, participants are led to the question which output technology is best in finding optimal solutions for their assessment criteria. In a baseline scenario, all participants look at the same output device. In the literature, this is called symmetric output, since all participants consume the same, view-independent content from one screen. Symmetric situations are given, when "the visual information and the possibility to interact with the virtual content are the same for [the] collaborators" [106].

Another **symmetric output** arrangement uses multiple VR-HMDs for all participants. Langbehn et al. propose a way to deal with multiple people in such a setup [107] in a physically shared space. They propose shadow-avatars, which are virtual representations of physical users, in order to show their spatial distance in VR. Users interact

at a physical distance to avoid collision in physical space and thus prevent accidents. They propose semi-transparent silhouettes of virtual human representations. Evaluations have shown that temporarily visible shadow-avatars generate significantly more collisions compared to permanently visible humans. For this reason, the physical and virtual space stay permanently coupled implying that single user teleportation in the virtual domain is prohibited. Virtual spatial relations match physical spatial representations between users.

Nevertheless, the literature also provides insights into **asymmetric output** and the combination of multiple individual output devices, such as large high resolution displays, AR Optical See-Through (OST)-HMDs, window to the world, projection spatial augmented reality and many more (see Benko et al. [108]). Depending on the desired professional goal and effect (e.g. immersion and presence), there are several possible combinations to display the above-mentioned contents.

Gugenheimer et al. present ShareVR as a proof of concept for an **asymmetric, co-located VR environment** where both HMD and Non-HMD users are engaged at the same time in a living room environment [109]. They argue that different users want different experiences and different levels of engagement while gaming. Therefore, they enable users to interact in the same physical and virtual space using different (asymmetric) visualization hardware. This study is limited to living room environment use cases in the private space with gaming as the main focus. This is why they measured engagement and enjoyment. Similarly to the VMS key properties presented by Otto et al. [110], they use a floor projection unit, portable "window to the world displays" and HMDs as visualization hardware, all registered to a common coordinate space by using a tracking system. In contrast to the work of Gugenheimer et al., PV workshops do not focus on enjoyment during the assessment workshops but on solving manual assembly issues. In addition, the audience of PV workshops is larger than the typical asymmetric interaction proposed by ShareVR. While PV workshops typically have 10 to 20 people interacting with the system (see Section 3.3), the use cases presented by Gugenheimer et al. are limited to two persons interacting in the same physical space.

The publication RoomAlive [111] by Jones et al. uses SAR experiences to transform the entire living room into a gaming experience by registering multiple projectors with multiple Kinect V2 cameras to generate an immersive SAR environment. They investigate the design space of such SAR environments for gaming purposes. Interaction is limited to a symmetric, co-located approach because all users in the living room use the same projections and input channels for interaction.

Even if there was the technical possibility to synchronize virtual contents throughout massive amounts of VR head-mounted-displays in PV workshops, this would not be an efficient method of collabo-

ration due to limited physical space. Therefore, the VMS framework offers a way to synchronize up to three VR HMDs in the assessment environment; meanwhile the other participants are consuming an exoperspective on a large high-resolution display. The VMS framework is thus mainly designed for use as an asymmetric output for collaborative situations.

The technical operator of the PV workshop must switch the available output devices according to the interactive assessment goal. For example, if a PMU is available, in-situ projections could be suitable to reach the assessment goal by superimposing the wireframe of an assembly part onto the physical chassis. Nevertheless, only some use cases are suitable for SAR (as discussed in Chapter 9). This means that a combination of asymmetric output devices is required as well. Multiple variants of visualization, tracking, and interaction can be chosen at the same time.

5.2.5 *Multi-user support*

Similar to the collaborative features, multi-user support is also a key property of the VMS to generate shared virtual worlds: "A situation where multiple users congregate to discuss, design, or perform other types of joint work is generally categorized as CSCW (computer supported cooperate work)" [97]. As described in Section 3.3.4, PV workshops consist of up to 20 participants. That is why the VMS's input and output devices must be designed for multi-user scenarios.

Output components must be able to be viewed simultaneously by up to 20 persons for collaboration. This is why LHRDs are most promising for view-independent rendering and data visualization. In contrast, view-dependent renderings are generated by one or multiple VR devices. Non-VR users share simulation contents by either looking at public displays or using additional VR headsets. Synchronized asymmetric visualization devices share the same simulation environment. All participants must be able to consume the shared digital contents and perceive the same virtual space appropriately, even if they use different output modalities.

Input components, such as tracking devices must be able to simultaneously track multiple people and objects for multi-user support. For instance, for ergonomic assessments in the VMS, fast switching between participants is enabled and multiple persons can be tracked simultaneously. Workers depend on and interact with each other in reality and with the simulation model. The same holds for object tracking. Multiple tools such as screw drivers can be involved simultaneously in work tasks, i.e. for mounting a front module of a car.

Sharing content is crucial for PV workshops. "Investigated objects are in general shared among users, in the sense of visibility, this means that all participants can see the same coherent model, con-

sistent in its state over time" [97]. In contrast to shared data, private data must remain invisible for other workshop participants. Private output possibilities are available for everyone because participants bring their own notebooks to the VMS. No additional focus is set on data security in the VMS as all participants are authorized to see all data, however, this data is of no interest to others as each user works on their own field of expertise (see also Seifert et al. [112], Langner et al. [113] and Winkler et al. [114] for public and private data) Therefore, this VMS's key properties focus on sharing public content.

In multi-user environments, the literature also discusses **social communication** between participants to keep communication channels as unaffected as possible. Argelaguet et al. propose a multi-user system to "support spatial understanding" by extending the capabilities of reality. "Users can talk to each other, gesture and point into the virtual scenery as if it were real. As in reality, referring to objects by pointing, results often in a situation where objects are occluded from the other users' viewpoints" [115]. Argelaguet et al. propose a rendering technique so that the participants' view-dependent rendering shows occluding virtual scene objects as transparent if they interfere with the other person's line of sight. Such systems help support social communication between multi-users [115].

5.2.6 *Natural User Interfaces*

Besides conventional human interfaces devices (HID), such as keyboard and mouse, interaction, exploration and manipulation of VEs can be realized using a new form of interfaces, namely Natural User Interfaces (NUI). Steve Ballmer sees those interfaces as "touch, speech, gestures, handwriting and vision" [116] that enable "contextual and environmental awareness, immersive 3D experiences and anticipatory computing"[116]. Following Wigdor and Wixon in their book "Brave NUI world", the term "natural" does not imply a mimicry of the "real world", but describes NUI as a "design philosophy and a source for metrics" [117, p. 9]. Even though mostly direct touch and gesture interaction is described, multiple modalities can enable the construction of a natural user interface. NUIs must create an experience that can "feel like an extension of their body" and feels just as natural to a novice as it does to an expert user." It "does not try to mimic the real world" and considers the "context, including the rich metaphors, visual indications, feedback, and input/output methods for the context"[117, p. 13]. Other sources add aspects of an invisible interface, direct interaction with content, fast learning curves of increasingly complex interactions, and usage of natural human behaviors to their definitions of NUIs. In consensus with all sources, interfaces are not natural if they "exploit skills that we have acquired through a lifetime of living in the world" [118].
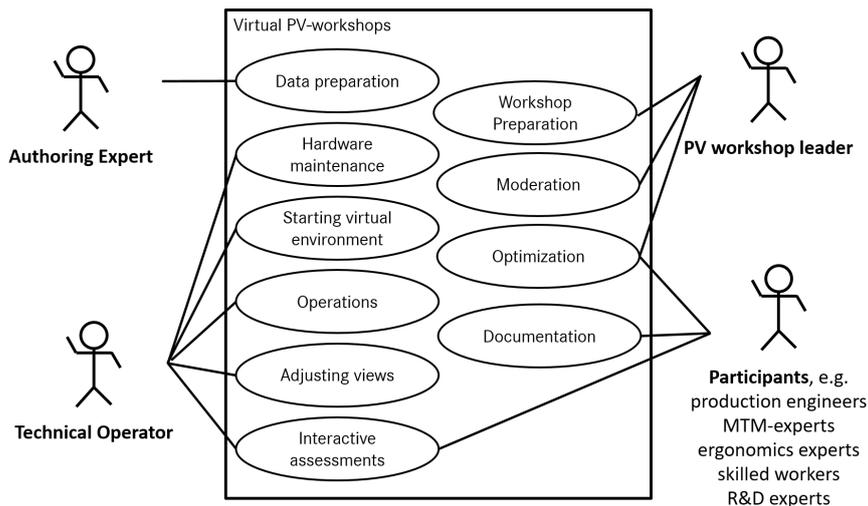
Figure 5.4: Four types of personas are involved in virtual PV workshops: Authoring experts, technical operators, PV workshop leaders and participants

As early as 2001, Zachmann and Rettig described a natural user interface for virtual assembly simulation scenarios [119]. They propose techniques and methods for multi-modal input techniques including speech input and gesture recognition for controlling the assembly system. Specifically, they present a natural grasping algorithm for collision-free assembly paths. As in the VMS framework, interactive assessments can be carried out using NUIs.

Following these design considerations for NUI interfaces, the VMS is expected to leverage new technologies and sensors. Space mice enable operators to manipulate CAD data easily with six degrees of freedom (see Preim and Dachselt [120, p. 289]). Markerless tracking with non-intrusive 3D measurements for full body interaction (see Preim and Dachselt [120, p. 304]) are also integrated in the VMS framework for DHM manipulation.

5.3   PERSONAS

In user-centered design processes, personas are useful to consider the workshop participants' desires, goals and limitations [121]. The entire interaction design of the VMS is adapted to the following personas. The personas are derived from the roles of all participants in the collaborative workshops (see Figure 5.4).

Even though there are many stakeholders for the VMS, this framework focuses on the main customers, namely production validation engineers. In interdisciplinary PV workshops, each specialist takes on a different role. Manufacturing specialists, including ergonomic experts, time experts, logistics experts or product experts, benefit from the VMS by utilizing interactive methods such as DHM manip-

ulation and immersive assembly simulations (for an exhaustive list of workshop roles see Section 3.3.4). Managers of interdisciplinary PV workshops hope to use VMS to achieve shorter execution times, fewer problem solving cycles and simplified documentation. Technical operators provide technological support for manufacturing specialists in the preparation and running of workshops. When using the VMS, technical operators provide supported by reducing system complexity during authoring but also during the workshops. They are in charge of manipulating the simulation as well as operating interactive components. In addition, adjacent stakeholders are all organizational units which deliver inputs or profit from outcomes of the PV workshops. For example, if PV workshops reveal optimization potentials in product parts, "research and development" becomes a stakeholder as they must optimize the product for manufacturing (see Design for Manufacturing and Assembly [40]). In terms of change management, stakeholders welcome developments that simplify their tasks or make them more efficient. Nevertheless, for strategic or tactical goals it is often necessary to put additional effort in one organizational unit to optimize the overall process. This potentially leads to stakeholders opposing the VMS's novel methods. Four types of personas using the VMS framework can be derived:

- **Authoring expert:** The authoring expert's main task is to prepare data beforehand the workshops. He/She is in charge of setting up the virtual simulation scenes in advance, which is called "authoring process." He/She must download all relevant CAD models, resource models, process data and fuse these solitary data sources to a holistic batch-operable virtual assessment environment. Multiple product variants with heterogeneous features must be configured and prepared for each workshop.

- **Technical operator:** The main task of the technical operator is to set up and maintain the entire assessment environment including hardware start-up, calibrating and instantiating the interactive components such as virtual reality and tracking systems. He/she loads the virtual environment prepared by the authoring expert. The technical operator is present during the entire workshop and steers the 3D views as desired by the specialists. He or she changes the manufacturing state accordingly to the validation process and compares all product variants with one another and aims at having the best usability as he/she is responsible for using all technical devices.

- **Workshop leader:** He or she is responsible for running PV workshops efficiently and assessing all verification tasks.

- **Workshop participants:** The participants are specialists in their area of expertise. They are responsible for finding optimal solutions and trade-offs in collaboration with the other specialists.
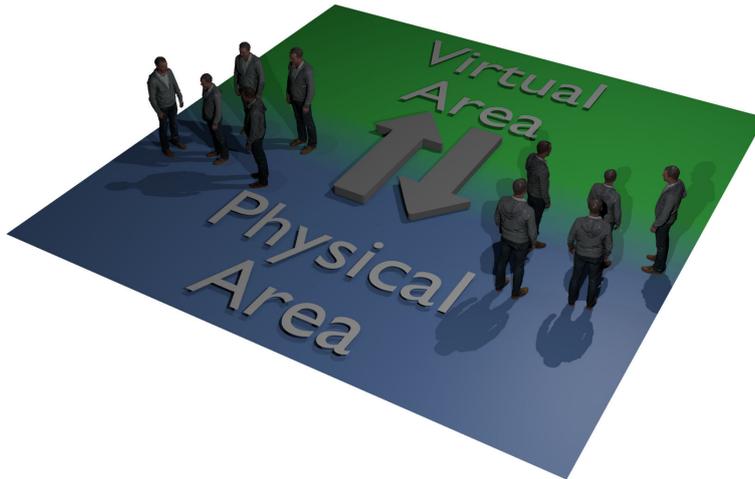
Figure 5.5: Merging the physical and virtual workshop areas

During longer workshops, participants may change over time due to their limited area of responsibility. For example, production engineers are responsible for a specific number of workstations.

5.4 HARDWARE COMPONENTS

Based on the key properties, a hardware setup consisting of multiple input and output devices and infrastructure is proposed.

5.4.1 *Room size and arrangement*

The VMS environment is a stationary workshop environment where people meet to run PV workshops. As described in Section 3.3.4, these PV workshops can consist of up to 20 participants, so the workshop area must be large enough for the respective number of people. In addition to this required space, one key property of the VMS is true-to scale visualization of digital contents (see Figure 5.6). Arteaga et al. describe such a workshop environment as an "area where assembly takes place includes the space required for equipment and workers, as well as the space required for the storage of components and finished products" [15]. Therefore, a real assembly cell must fit in the VMS's workshop area.

Since passenger cars are typically up to 6 meters long and 2.5 meters wide, a manufacturing cell with additional worker areas and material zones should be at least 9 meters by 6 meters for a realistic validation of the future assembly environment. Hence the VMS must be at least this size.

5.4.2 *Output devices*

For a holistic VMS workshop environment, multiple output components are combined:

- Laser and video-based projection systems for Spatial Augmented Reality

- Powerwalls for large high resolution displays

- Floor visualization for interactive manipulation of the simulation environment

- VR headsets for immersive 3D experiences

These approaches are illustrated in Figure 5.6, showing a rendering of all components of the intended VMS framework. Figure 5.7 shows a rendering of a large high-resolution, wall-sized visualization device as a vision for the VMS's multi-display environment.
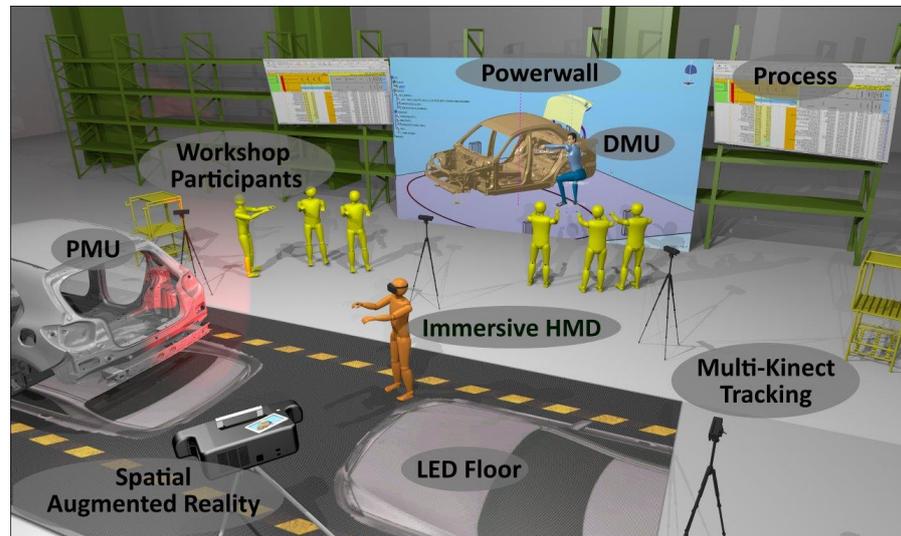
Figure 5.6: Rendering of the vision for an integrated virtual manufacturing station environment as a collaborative workshop environment for interactive assessments of manual assembly tasks

5.4.3 *Input devices*

Enabling natural user interfaces also requires input technologies. Besides classical interfaces such as keyboard, mouse and 6D-spacemouse, marker-based (ARTracking) and markerless full-body tracking and object tracking sensors for an arbitrarily large tracking area are integrated as well as direct floor interaction capabilities. For markerless tracking, an array of multi depth cameras (Microsoft Kinect v2) is
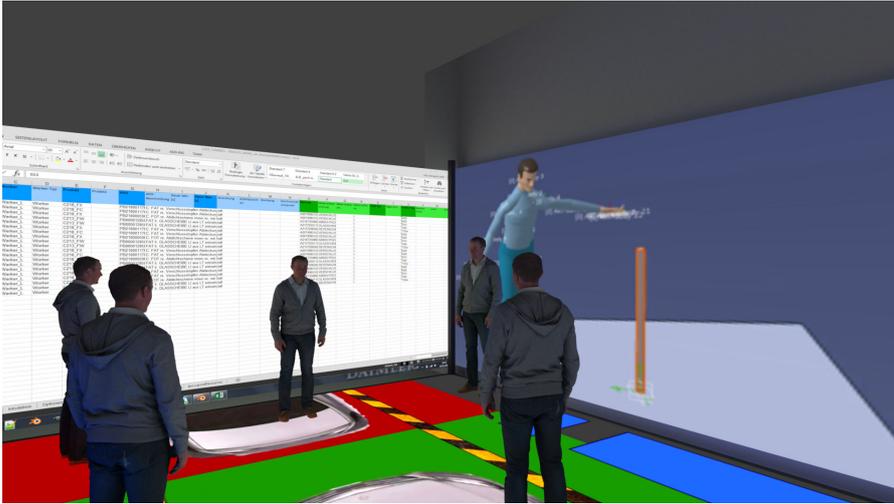
Figure 5.7: Rendering of large multiple wall-sized large high-resolution visualization devices with a LED floor.

installed. For VR interaction, base stations are integrated as input devices for precise 6 Degrees of Freedom (DoF) tracking (HTC Vive base stations).

5.5 SIMULATION SOFTWARE COMPONENTS

As the hardware components must be driven by a simulation environment, one central assembly assessment program is presented as a dedicated batch assessment environment for manual assembly processes. Nevertheless, multiple tools have been generated to drive the entire VMS's framework. The central software allows the holistic simulation of entire workstations in the upcoming factory. Therefore, heterogeneous data sources are fused and integrated in the VE. To enable the usage of the proposed hardware setups, the simulation software must also support the tracking of objects, manipulation of DHMs, true to scale rendering and orthographic bird's eye view for augmented floor visualizations. Spatial augmented reality is attached to standardized interfaces.

The participants' experience is a batch evaluation of a complete factory assessment. The software allows to build up the product virtually step by step, e.g. a passenger car. During these assessments, the batch simulation environment also allows for collaborative manipulation of the VE with immersive environments.

5.6 RESEARCH AREAS RELATED TO THE KEY PROPERTIES

Summarizing the key properties and the research agenda for the VMS, all studies in the following chapters are carried out to contribute to the general VMS's application objectives and key properties. All as-

pects contribute to the overall system but the respective studies contribute to specific key properties. At least one study contributes to the mentioned key properties and exemplifies how the respective key properties are enabled by research presented in the respective upcoming chapters. These interlinks between proposed key properties and research areas are depicted in Figure 5.8.

Chapter / Key Property	VAR assessment environment	Markerless Full Body Tracking	Spatial Augmented Reality	Large High Resolution Displays
Collaborative virtual environment	✔ VR assessment benchmark	✔	✔ Collaboration performance	✔
True to scale output	✔		✔	✔ Size perception study
Multi-user support	✔	✔ Ergonomic assessments	✔	✔
Symmetric / asymmetric output for AR/VR	✔		✔ Comparison to OST-HMD	✔ LED floor stencil study
Integration of PMU & DMU	✔		✔ Design space analysis	✔
Natural user interfaces	✔	✔ Tracking performance	✔	✔

Figure 5.8: Research matrix of the studies related to the key properties of the VMS

The following research studies are discussed:

1. VR2A: By presenting research on an open standardized benchmark for virtual reality assembly assessment, assembly simulation environment limitations and properties can be revealed. This enables more sophisticated assembly simulations and widens the application areas for CVEs.

2. For spatial augmented reality, the collaboration performance is evaluated in an abstract workshop situation with divided knowledge. This research generates insights into the task completion time, using different ways of computer-mediated communication including spatial augmented reality.

3. For the true-to-scale key property, a size perception research study is presented, testing the accuracy and precision perceived sizes when using an LED floor.

4. In the context of multi-user support, ergonomic assessments are presented using a scalable markerless tracking system.

5. For the asymmetric output key property, an augmented floor surface is used as a virtual stencil. An evaluation shows that task completion times for setting up typical cardboard layouts can be reduced by using LED floors.

6. A design space evaluation is carried out for spatial augmented reality to demonstrate the benefits and limitations of using spatial augmented reality in PMU and DMU registrated setups.

7. For the key property, natural user interfaces research is carried out on a novel scalable, markerless full-body tracking system is presented and analyzed with respect to its tracking performance.

Additionally, further research contributions are the state of the art analyses and the demonstration of a VMS's framework implementation.

VR/AR ASSEMBLY VALIDATION

Assembly validation software enables production planning departments to simulate mid-level to low-level production processes in the virtual domain. In their literature review, Cecil and Kanachanapiboon [122] distinguish between two major categories in **virtual prototyping**, namely virtual product design and manufacturing validation. For virtual manufacturing validation, the sub-categories of factory-level simulation, lower-level machining processes and virtual prototyping of assembly processes can be clustered. This chapter focuses on the latter sub-category, by presenting deficiencies in the literature and an implementation of this type of software for the automotive sector.

Ideally, planning and validation software should be integrated in a way that eliminates the need for data transfer and manual preparation when using the validation software components. Commercial CAD-based validation software is offered for various domains in the manufacturing industry, such as logistics simulation, line balancing simulation and robot simulation as well as manual final assembly simulation. This leads to a versatile environment of multiple commercially available validation tools, each having differing main scopes with overlapping capabilities.

This chapter provides insights into human-computer interaction and production engineering research aspects concerning assembly validation tools in the context of the VMS. The VMS framework is the central component for driving virtual and mixed reality virtual prototyping build-ups of upcoming products. The VMS framework's key properties (see Section 5.2) must be enabled by the assembly validation software in order to achieve the overall VMS requirements. In this chapter the following research questions are answered:

- What requirements exist for virtual validation of manual assembly tasks?

- What are the deficiencies of current commercial products?

- What does an assembly simulation tool look like that supports the key properties of the VMS?

- What data is required to generate a holistic virtual manual assembly line?

- How does VR support simulation capabilities and how can manual assembly processes be validated using VR?

- How can operators determine and quantify their overall spatio-temporal limitations of their VR assembly simulation?

For this doctoral thesis, the baseline has been defined based on the contextual inquiry study (see Chapter 4). Current automotive PV workshops are carried out in two ways: either entirely hardware-based build-ups or sporadic DMU build-ups in static virtual environments. PMU-based workshops cannot have all product and part variants physically available as there are too many combinatory possibilities of upcoming products. Physical assessments are increasingly replaced by virtual assessments as fewer or no PMUs are available in production preparation phases. DMU-based assessments currently use a PDM viewer only. This means that no interaction is carried out with the virtual prototype besides changing the visibility of parts in the manufacturing sequence. Moreover, no additional components besides the product itself are included, such as factory environments, tooling, interactive DHM, screwdrivers and others (compare intention of virtual prototyping in [122]).

This chapter first describes objectives of assembly validation software in the context of the VMS with a focus on VR assembly assessments. Next, the deficiencies of commercial state of the art validation software are discussed, followed by implementation details of a novel validation software program. This VR/AR assembly simulation software represents the core simulation software for the VMS. First and foremost, its purpose is to enable virtual prototyping and interactive production validation of manual assembly processes. In contrast to planning tools, the software limits the scope to the assessment of manual assembly processes that support the VMS key properties and uses all its hardware capabilities. It is thus a decision-making tool, not a planning tool for generating additional production planning contents. Batch assessments, fast rendering capabilities, standardized interfaces for AR/VR and fast data preparation must be enabled. At the end of this chapter, a VR assembly assessment score (VR2A) is proposed and evaluated to create a standardized, open benchmark for VR assembly assessments. Its goal is to quantify the overall system's performance and limitations when carrying out assembly tasks in VR.

> **Research and Development context**
>
> The validation software and simulation framework presented were developed as part of the publicly funded German BMBF research project "ARVIDA" and were then integrated into a real PV application. The virtual prototyping software discussed here has been implemented by "Daimler Protics." Personal contributions in this doctoral thesis are technical concepts, technical specifications, research exploitation and dissemination and product owner responsibilities for the development of a user-centric, productive simulation software program.

6.1 SIMULATION SOFTWARE OBJECTIVES

For automotive PV workshops, interactive virtual prototyping software must allow production engineers to assess as many mid- and low-level verification aspects as possible (see Section 3.3.5). Mid-level manufacturing virtual prototypes comprise "exploring process design issues within a work cell, [. . .], human operators, material handling devices [and] conveyors" [122] in order to virtually study "the interactions of these various components" [122], so that "process alternatives can be compared and modified" [122]. Low-level validations comprise optimizations related to detailed analyses, such as "design of fixtures" [122]. Therefore, four main purposes are derived as follows:

First, the simulation software must be able to perform a holistic **simulation of mid-level manual assembly workstations including all product, process and resource relevant items**. The virtual car assembly simulation must inherit as many details as possible, including factory data, resource data, human data and many more. Users in PV workshops must be provided with all details required to assess their own verification criteria (see Section 3.3.5), e.g. process quality, product data quality and economic aspects such as reduction of waste and time, process efficiency at all workstations and overall assembly line balancing. In order to reach all these aspects virtually, multiple data sources must be fused in the assembly validation scene.

Second, the **key properties of the VMS's collaborative virtual environment** must be enabled by the assessment software (see Section 5.2). All hardware components proposed in the context of the VMS must be supported natively. Following that, collaborative assessments in workshop situations are a central requirement for the validation software. Standardized tracking components must also be integrated in the software as well as true to scale visualization capabilities on large-scale output devices.

Third, **efficient simulation model** generation must be made possible. State-of-the-art commercial tools require high authoring effort (e.g. see Manns and Nestor for DHM animation effort [123]) for gen-

erating such simulation models, especially for the simulation of an entire assembly line. Choi et al. call for "techniques that enable handling the whole factory at an instance, analyze it rapidly, and provide speedy feedback to the shop floor" [124]. Most commercially available simulation software components focus on the assessment of single workstations with critical contents, whereas in PV workshops **batch simulation** of entire assembly lines must be made possible. For example, in PV workshops not only one ergonomically critical workstation must be assessed but in one PV workshop hundreds of workstations containing all work contents must be assessed with as little data preparation effort as possible. Even though this issue has been the focus of many studies (compare Graf et al. [125]), it remains to be solved.

Finally, **intuitive user interfaces** are required to reduce entry barriers for operators and attendants. Operators must currently undergo rigorous training so that they can build the VEs, navigate in the virtual space, assemble the vehicle virtually, meet the assessment needs of the workshop participants and enable VR assessments. All these required skills can be simplified by means of user-centric assembly validation software.

6.2 STATE-OF-THE-ART COMMERCIAL VALIDATION TOOLS

This section discusses drawbacks of current, commercially available validation software in accordance with the findings of the preliminary expert interviews (see Chapter 4). This state-of-the-art analysis has been carried out in 2016.

The market offers a variety of manufacturing planning and validation tools for industrial assembly tasks. They vary greatly in their scope but have overlapping capabilities as well. Commercial CAD-based validation software is offered for various domains in the manufacturing industry, such as intra-logistics simulation, line balancing simulation, path planning simulation, robot simulation, ergonomics simulation process optimization as well as manual final assembly simulation.

State-of-the-art commercial tools can be clustered in production planning tools and production validation tools. Often, the scope between planning and validation tools is differentiated even though both aspects are integrated. Typical representative planning tools are "IPO.plan" by IPO.Log, "Delmia Process Engineer" by Dassault Systems, "Process Designer" by Siemens, "AssembySuite" by Taktiq and others. They are natively generated for the initial planning of detailed production processes and for balancing the workload in continuous flow line production. Dedicated assembly assessment simulation tools are "IC.IDO" by ESI, "Process Simulate" by Siemens and

"Editor Menschlicher Arbeit" by IMK. "Delmia" by Dassault Systems can be used for both tasks planning and validation.

However, multiple clusters of general deficiencies reveal why commercially available tools can hardly be applied in automotive PV workshops and the VMS. The following clusters describe the most important challenges:

Deficiencies in visualization systems: Even though graphics cards and computing power are continuously improving for desktop computing systems, real-time rendering of massive data sets continues to be a problem for desktop validation tools. In 2015 Choi et al. stated that the "development of 3D expression technology for manufacturing 'big data' is required" [124]. Volume independent rendering of large datasets is a basic requirement for instantaneous switching between virtual prototypes. Since production engineers are responsible for manual assembly tasks, they must interact with 3D representations of both the product and the resource. For this reason, huge data volumes must be visualized at interactive frame rates. A minimum refresh rate of 30 Hz must be reached for powerwall visualization, and ideally a 90 Hz rendering refresh rate for VR applications, for fluent rendering and reduction of cybersickness (compare [126]). Typically, the tessellated 3D geometry of a complete summary automotive product car geometry is 30 to 50 gigabytes of data with 10 to 100 million vertices and faces. This is an area of ongoing research as documented in the "Visualization as a Service" approach [127] by Karl and Behr in 2017.

Novel techniques for the lossless rendering of arbitrarily large 3D data sets are visibility-guided rendering technologies (see 3DInteractive [128]) and other volume independent render systems (see NetAllied [129]). Large 3D product geometry must be visualized in realtime where ideally "render performance is essentially independent of data volume". High performance renderers allow the visualization of large-scale models with over 100 gigabytes of data on standard notebooks and PCs. Interaction can be implemented for these renderers, such as the picking of objects, position, animation or change of appearance. Scene-graphs can be built on engineering data structure or can be rearranged in object tree browsers with millions of objects. Most recent approaches allow for efficient server-sided rendering such as "visualization as a service" by Fraunhofer IGD, also called webVis/Instant 3D Hub [130]. These approaches are now increasingly used in the automotive industry [127] and allow the efficient streaming of pre-rendered, lossy models, but currently do not comprise any assessment features and interactive features required for PV workshops.

From a practical point of view, PV workshops require that all CAD data of a product be shown simultaneously – the so called summary product – in order to assess all variants of a product at the same time. In a slightly different way, the instantaneous switch in rendering be-

tween multiple product variants (e.g. Coupé to Hatchback) must be possible without any loading times. Lossy or reduced rendering of models is not acceptable for PV workshops as participants need to see the parts exactly as they are modeled. Since DMU geometry inherits multiple "layers," it must be possible to switch the visibility of each layer on and off. Data conversion steps must be avoided or at least be carried out on-the-fly during the run-time of the assembly validation software. This is not the case with most commercially available software tools.

Deficiencies in batch assessment features: Efficient PV workshops must be enabled by the assembly assessment software through the facilitation of batch assessments. Commercial products focus on single workstation simulations not optimized for the batch assessment of workstations. This leads to the requirement that the build-up of the entire product can be achieved with one pre-generated VE showing the build-up as close to reality as possible.

Deficiencies in interactive features: The overall goal is to visualize the entire production process in a dynamic environment. Such dynamic product build-ups include kinematics of the product and resources, e.g. rotation axes of the doors, hood or translatory kinematics for the marriage resources, which are applied from beneath the car body. When attaching this generic tracking information to products, DHMs or resources, many new use cases can be enabled, such as full body tracking, hand tracking and object tracking. With these capabilities, the simulation software can be used for the interactive simulation of screw driver insertion paths, attachment of sub-assemblies to trackers, visibility analyses using tracked views and many more. VR support is still limited in the field of assembly simulation software, even though research has on potential solutions and their benefits has been conducted for several decades [26, 75].

Deficiencies in digital human modelling: In addition to general interactive deficiencies in commercial validation software, digital human models and their animation are sometimes either completely unavailable or only have limited manipulation features such as proprietary protocols. Therefore, digital human modeling must be addressed in a standardized manner to allow the integration of new tracking devices. Forward and inverse retargeting algorithms must be integrated and the number of joints in the skeleton must be precise enough for full body tracking, including hands tracking. Avatars and skeletal parameters must be changeable during run-time. In addition, population databases for workers would be beneficial. Both online and offline animation must be enabled so that the simulation can be replayed after a workshop ends. Sometimes only parts of the DHM must be animated. Therefore, dedicated body parts must be driven separately, such as lower body and upper body. For example, while the lower body is in a static position, the upper body moves interac-

tively. Another example is setting a pre-defined hand posture, while the rest of the DHM is tracked online. These features are not yet commercially available in "out of the box" virtual prototyping tools.

Deficiencies in automotive DMU product build-up: Because assessment simulation software can be used generically in the manufacturing industry, "out of the box" they lack specific features for automotive manual final assembly: Data integration of DMU and textual manufacturing process plans, realization of pre-assembly cells in precedence tree, showing the car manufacturing state at a given workstation with one click, geometric mapping of moving product in a DMU assembly line, and the matching of tasks and products on different hierarchical levels. Additionally, for automotive manual final assembly easy adjustments of work heights, rotation of products in C-hangers, assessment of walking paths and the usage of parametric cell layouts are usually not included in commercially available assembly validation software.

Deficiencies in standardized interfaces and resource databases: Research into standardized exchange formats for manual assembly task processes, CAD data and resource data carried out by research institutes and OEMs on the "AutomationML" standardization board (see [131]) is ongoing. In addition, support of open protocols is provided by means of assembly paths, online tracking devices and kinematics. OEMs have compiled comprehensiv resource databases that they intend to use in virtual environments. However, generic assembly validation tools often do not have import functionalities for these databases since they do not have any previous knowledge on this proprietary information. Using the commercial tools, all resources must be imported separately as a singular resource item. Batch assessments must be able to import entire resource databases at once, such as a screw-driver database, tooling database, assist devices database, point cloud scans of factories and part annotation databases, such as screws, clips, nuts and bolt databases.

Deficiencies in the VMS key properties: None of the commercial assessment validation tools offer ways to natively drive the VMS's hardware. For enabling the VMS's key properties, the software needs to render multiple asymmetric views (Exo & Ego perspectives, VR, AR and powerwall) as well as true to scale orthographic bird's eye views for floor visualization, true to scale orthographic side views for powerwall visualization and must have standardized interfaces for augmented reality hardware.

Overall, this leads to the conclusion that a novel virtual prototyping software program is required which eliminates the drawbacks of commercially available tools mentioned above. In 2015 Choi et al. stated in their literature review on virtual reality applications in industrial settings: "Now, it is important to connect and integrate these VR technologies and efficiently implement them in the manufacturing area"

[124]. None of these systems inherit the ability to support all VMS framework's key properties for collaborative workshop environments. Therefore, as part of the publicly funded research project "ARVIDA" a novel virtual assembly validation simulation environment has been developed. In the context of this doctoral thesis, multiple parts of this assembly validation software were presented in the Springer book "Webbasierte Anwendungen Virtueller Techniken" by Schreiber et al. [132].

6.3 IMPLEMENTATION OF A BATCH CVE ASSESSMENT SIMULATION ENVIRONMENT

The features of the so called "veo" virtual prototyping tool are described in the following sections: Authoring and data provisioning for batch validation, rendering and visualization, general assessment features, output devices, interactive components and a digital human model. Figure 6.1 depicts the user interface of the assessment software with its central features.
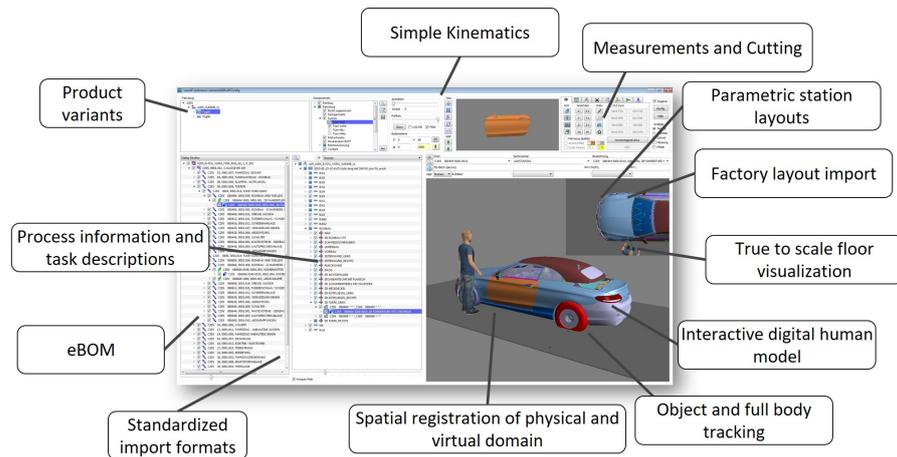
Figure 6.1: User interface of the assembly simulation software veo. (Left) Product variants, manufacturing sequence and eBOM can be switched. (right) 3D visualization of the virtual scene and bird's eye view of the parametric cell layout. (upper right) Manipulating the interactive assembly simulation. (top) Active single part visualization

6.3.1 *Authoring and data provisioning*

Generation of a holistic VE including PPR requires to integrate many data sources (see also Mbang [33] for template-based modeling of PPR). These data sources generate alternative structures for the VE and must be fused for enabling batch assembly validations in PV work-

shops. Data provisioning in veo fuses PPR structures semi-automatically in order to reduce preparation time for PV workshops.
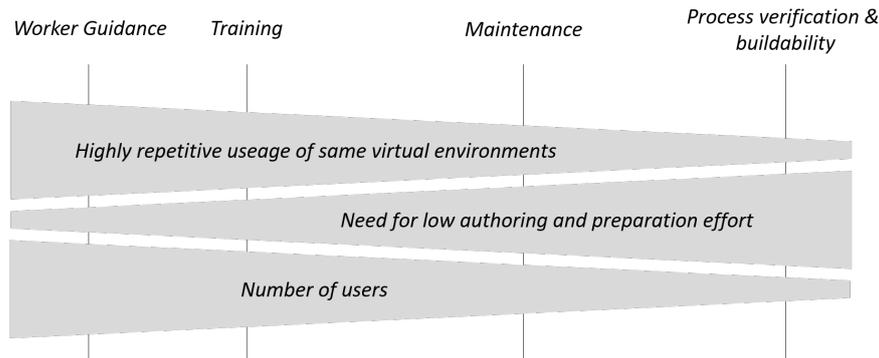
Figure 6.2: Need for low authoring and preparation effort of virtual environments in process validation

As depicted in Figure 6.2, using virtual environments for worker guidance and training, the VE can be reused multiple times as many products are produced or many workers are trained using the same VE. In contrast, production verification VEs cannot be reused multiple times. Having assessed all verification criteria, the VE's purpose is achieved and no further knowledge can be generated by using this version of VE. Therefore, the overall goal is to reduce authoring effort to a minimum or even to a fully automated data provisioning process. Veo features multiple mechanisms to achieve this goal.

The **product geometry** is exported from the native PDM systems to veo with its respective hierarchical product structure and bill of materials (BOM). Engineering BOM structures, as depicted in Figure 6.3 for complex products typically contain multilevel structures such as "Product -> Main Assembly -> Assembly -> Sub-Assembly -> Position -> Position Variant -> Assembly part" (see Figure 6.4 and "Infor" Engineering BOM [133]). Various manufacturing-related attributes are attached to meta-data of product parts, such as definition of screws, corresponding torque, weights, bounding boxes. This information is saved in a standardized, interoperable exchange format called AutomationML [131]. Geometry information (Spline-based and tessellated) and kinematics are saved in the ISO-standardized Jupiter Tesselation (JT) files [134] for lightweight product geometry files and its associated PLMXML files for structure information. One summary product model can inherit up to 100.000 separate parts and 100 gigabytes total data volume including all variants. A single buildable car representation typically sums up to data volumes between 10 to 30 gigabytes.

Besides product data, **process information** must be imported into the software. Process data consists of reports sourcing from the native production planning system and inherit information on all detailed work plans, namely the manufacturing sequence for manual assem-

Top-level product structure
"main assembly"

```
» Car
    » Body
    » Interior
    » Seats
    » Accessories
    » Electrical
    » Suspention
    » Exhaust
    » Engine
    » Transmission System
    » Brakes Mechanism
    » Air System
    » Clutch System
    » Heating System
    » Air Conditioning System
    » Steering System
    » Cooling System – Water
    » Cooling System – Oil
    » Safety System
    » ...
```

expand →

Multi-level product structure expand
"assembly" and "sub-assemblies"

```
» Interior
    » Instrument Panel
        » Dashboard
        » Cross Car Beam
        » Instrumentation
            » Cluster System
            » Audio System
            » Screen System
            » Column Switch System
            » Electronics Ignition System
            » Heads Up Display System
                » Heads Up Display
                » Cover
                » Insulation
        » Dashboard Covers
        » Storage Compartments
        » Air Vent
    » ...
```

Figure 6.3: Example for a multilevel product structure of a car eBOM. (Right) Multiple levels "main assembly", "assembly", "sub assembly", "Position" and "Sub-Position" is depicted for the example "Heads Up Display System" (product structure according to a2mac1.com [135])

bly operations. A connection between the product eBOM and manufacturing sequence hierarchy must be established (see Figure 6.4). Planning data is represented as a spreadsheet file, which contains alphanumeric information on the work task descriptions. They are listed in the execution sequence of the manufacturing plant with its respective organizational structure (Factory -> Assembly Line -> Cell -> Workstation -> Work Tasks). Additional attributes in the manufacturing sequence information are used to interlink product geometry with the manufacturing sequence. Additional information on the process is imported, such as product codes, variance, building frequency, building times, building grades, and safety relevant information. The interconnection between the product geometry and the process can be carried out by matching each part of the product to the process. For example, the "assemble rear mirror" work task has a "product geometry reference" attribute for matching the correct product geometry in the DMU files. Once all processes are connected with each single part of the product, the entire car is restructured in an alternative hierarchy which is then called the **manufacturing sequence** hierarchy. This process already represents the logical structure of the future production plant, where each part is assembled into the car at a certain workstation. All this process information is fused in the assessment software.

The product data and the alphanumeric process information are fused automatically. In this way, the respective manufacturing state
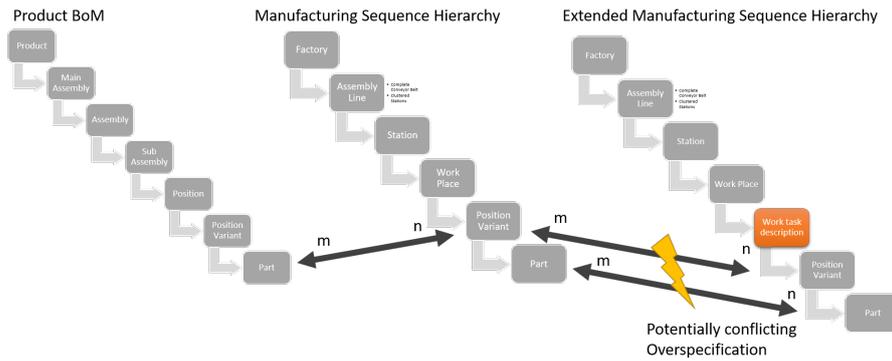
Figure 6.4: Interconnection between product BOM, manufacturing sequence hierarchy and work task descriptions

at each workstation can be deduced automatically as well as the assembly line instances. Pre-assembly workstations are positioned in the assembly tree, where the pre-assembled parts are delivered to the main assembly line according to the "herringbone principle."
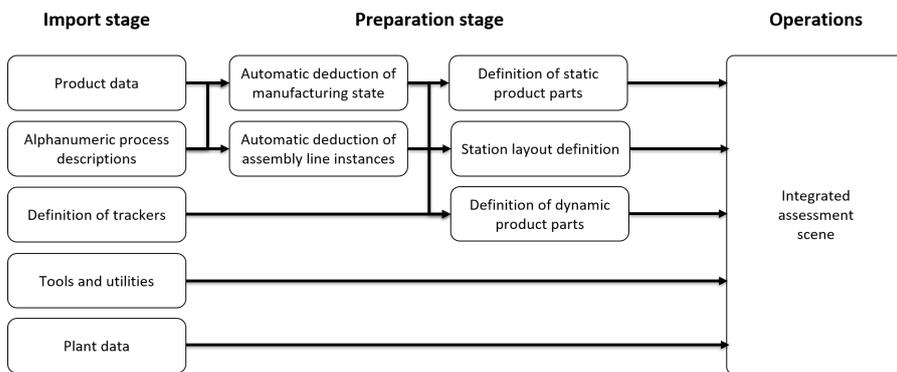
Figure 6.5: Schematic import and data fusion process to crate a scene for the virtual assessment software

To obtain a holistic view of the workstation, DMUs of **additional tools and utilities** are required, such as racks, carriers, garbage boxes, screwdrivers databases, specialized tools and handling devices. These databases are imported and labeled as such, as shown in Figure 6.5.

Furthermore, if additional DMUs of the factory are available, entire plant or factory layouts can be imported. Factory 3D data can be represented as point cloud scans [136] or meshed 3D factory data. This data is integrated in the VE to obtain a better idea of what the workstation will look like in the future. A typical example for a validation task using this plant data is checking restricted areas such as evacuation routes that might collide with the future workstation. Recent research focuses on the integration of **large-volume point cloud scans** instead of CAD models for the same kind of use cases (see Gong et al. [136], Lindskog [137], Shellshear, Berlin & Carlson [138]). In contrast to static CAD geometry or static scans, standardized **paramet-**

ric cell layouts can be imported as well. Parametric cell layouts are used to keep the VE more flexible so that quick changes can be made throughout the validation workshop. For example, racks shelves can be modified in terms of height, angle, amount and type, whereas in a CAD geometry this is not easily possible. Such parametric layout data also inherit information on cell size, workstation size, walking paths, no-go areas, racks, carriers, cartridges, picking zones and most importantly the interconnection with the product structure. The position of a product in a parametric cell layout can thus be visualized in a rack (beginning position) and after assembly in the car (end position). The data sources must be geometrically aligned in the virtual plant by determining the transformation. Especially for VR assessments a plausible cell layout as a background is required to generate a feeling of presence in the scene. Standardized 3D import geometry formats are JT, OBJ or STL.

6.3.2 *Rendering and visualization*

Rendering speed is a crucial aspect for virtual prototyping with DMU. A minimum refresh rate of 30 Hertz is required, and for immersive output visualizations a rate of at least 90 Hertz is necessary for both eyes to achieve a convincing interactive manipulation of the environment (compare [126]).

DMU models with up to 100 gigabytes of data must be visualized at those interactive speeds. In contrast to the gaming industry, where objects are reduced and optimized for the hardware platforms, industrial applications cannot manually optimize the product to lower polygon counts. The applied real-time renderer [139] is called "visibility guided renderer"(VGR) for large datasets made by 3Dinteractive. "Typically for extremely huge CAD data (>100 million polygons) more than 99% of data is excluded from rendering" [140]. This rendering platform is intended "for interactive handling of massive data sets." It was initially presented to face the "data explosion" problem in CAD data for "automotive, aerospace and construction models" [141]. They offer volume independent rendering capabilities so that the entire car with all variant geometries and all additional information can be rendered with interactive speeds. Therefore the volume independent rendering offers various culling techniques, such as occlusion culling [142] (only visualize the top-most visible layers) and frustum culling techniques. Visibility guided rendering identifies non-visible objects during run-time to exclude them from the rasterization process and therefore achieves constantly high rendering rates. When index structures are applied, the time required for calculation "rises significantly less than linearly with the data volume" and "for certain model sizes the calculation effort is almost independent from the object data volume" [140].

When the DMU is imported into the software, the material is set to non-specular material and diffuse material color is applied from the standardized color table. For this reason, the product does not look photo realistic but the parts are clearly distinguishable. Veo allows for the visualization of all product variants (summary product), engines and gearbox variations at the same time, including the entire factory environment.

Hence, in multi-user environments stakeholders want to use individual output components while sharing the same physical space and the same virtual environment. Distributed rendering thus ensures the scalability of the collaborative virtual environment for asymmetric views. Each output, for instance each VR output and high resolution powerwall, is rendered on a dedicated workstation. The scene graph is synchronized via a network between multiple veo instances.

Standard usage of veo has two high resolution rendering views. The main rendering view shows the product in its respective manufacturing state with the complete surroundings. A second rendering view displays the currently selected, single assembly part visualization. According to the actual work tasks in the batch assessment, the details of this single product can be assessed. For example, this view can be used to cross validate whether the corresponding work task description matches the geometrical fixtures.

6.3.3 *Assessment-related features*

Veo allows the visualization of entire **summary products and code-derived buildable products**. Summary products are not buildable in the physical domain but enable planners to compare all DMU variants simultaneously. For instance, 10 different front bumpers can be visualized at the same time to get an understanding of the same fixture concepts. So called code rules derive single buildable products from a summary DMU. Instant switching between preprocessed summary products and single buildable products is an essential feature in veo.

In PV workshops, a specific product is built up in the respective **manufacturing sequence** of the upcoming factory. Veo follows this approach in the virtual domain as well and enables the visualization of each product's manufacturing state for the respective workstations with one click. Validating such a manufacturing sequence of a single buildable car allows problems such as sequence issues to be resolved by altering the process information. An example here is an already installed part blocking the assembly of a subsequent one. In this case, the manufacturing sequence would have to be changed.

Comparing multiple buildable products can be cumbersome since the combinatorics of assembly parts are high. Therefore the **product comparison feature** allows colors to be assigned to each product. In this way, production engineers can easily distinguish product differ-

ences in the manufacturing sequence structure, work task description list and of course in the DMU rendering view. For example, products with 12V components are coded red and products with 48V components are coded green.

The product structure itself does not resemble a geometric neighborhood but represents a logical structure. That is why an additional structure element is available in veo, namely **components.** Components are used to create an alternative structure for geometric neighborhood elements. For instance, all parts belonging to the left front door should be clustered logically when kinematics are defined for the left front door. Respective parts of this left front door are spread all across the eBOM structure. Electrical parts, chassis parts and cover parts are all added to one component called "Front door left" in the component alternative hierarchy. These components are used to instantly switch the visibility of the whole component, to set components to semi-transparent and to define the kinematics of the component.

These **kinematics** are assigned to component structures. In veo simple kinematics are practical approximations of what are often more complex kinematics. Linear and rotary translations can be attached to each component so that the doors, motor hood, rear door and additional machinery can be animated. For example, when simulating a marriage of the drivetrain components, the mech frame can be easily animated by means of linear animation.

Alphanumeric work task descriptions are filtered and displayed in a list view according to the current respective manufacturing state or eBOM selection. One summary car contains up to 50,000 work task descriptions, whereas a single work cell can have up to 50 alphanumeric work tasks. In the authoring stage, product and process structures become interconnected. Workshop participants can display all corresponding work tasks that come with a selected part or a selected workstation. The list of work task descriptions is filtered accordingly.

Cross-highlighting between PPR structures is one of the most frequently applied features. Selections made in one of the PPR structures will lead to a (multi-)selection in both other corresponding structures. For example, when selecting a work task description, the linked assembly parts are highlighted in the overall car manufacturing state, or when selecting a whole cell in manufacturing sequence, all assembly parts that belong to this cell are visualized separately and highlighted. Meanwhile, the selected assembly part is also visualized in the single parts rendering window. A cross-highlighting functionality shows this specific connection. The same holds for selections in the eBOM structure and the manufacturing sequence structure.

Another essential feature is the **animation of product and assembly part trajectories** in relation to the factory environment. The height of the product can be transformed relative to the floor and the angle

of the floor in order to permit reachability and visibility studies. Different types of skids (work height differences) can be simulated easily using this feature. Given a certain cycle time and length for each cell, the dynamic visualization of product structures is made possible. In the case of a parametric cell layout, start transformation (assembly part in carrier) and end transformation (manufactured state) of an assembly part are defined. A linear transformation between both points is visualized for a better understanding of the processes. The assembly parts thus fly into the car "step by step" just as the manufacturing sequence describes the process. However, this animation does not include collision avoidance and path planning.

Search functionalities are given for all PPR structures. With thousands of objects in the summary PPR structure, a full text search makes it possible to find items. Results are clustered by object type and selected in the hierarchy they are found in.

For detailed assessments, DMUs can be **sliced and cut**. Especially in cluttered environments, detail assessments are important for product optimizations. The same is true for **measurement possibly**. Point, linear, circular and volumetric measurements are possible using veo.

Having assessed all verification criteria for each workstation, multiple outputs are generated such as tracking lists, verification KPIs and screenshots of revealed issues. A screenshot functionality automatically fuses alphanumeric information about the selected assembly parts with the manufacturing state rendering window.

6.3.4 *3D rendering features*

Three different rendering windows are integrated for the visualization of the same VE: Integrated assessment view, assembly part detail view and orthographic bird's eye view. All render views can be rendered on custom screen resolutions and on distributed workstations. In the integrated assessment view, the selected assembly part can be centered using the **"fly-to functionality"**.

The **integrated assessment rendering** visualizes all imported components: Product, processes, factory layouts, tooling and machines. There is global ambient illumination combined with a headlight attached to the virtual camera with an offset of 50cm to the top right for better shading. The color of product parts is determined by the product itself. No textures are applied to the products and random colors are used to distinguish the parts.

The **assembly part rendering** displays the selected assembly parts or multiple selected assembly parts only. This helps to identify the part structure itself. It can be either selected by product structure parts or manufacturing sequence parts. Illumination, color and shading are the same as in the rendering window mentioned above.

The **bird's-eye view rendering** has been tailored for floor visualization devices supporting the augmented floor surface described in Chapter 8. This rendering window uses an orthographic virtual camera in order to show VE true to scale with no perspective issues due to camera focal length. The virtual camera is set at a height of 10m and can be altered in terms of its native width and height and its resulting aspect ratio, so that it corresponds to the applied hardware floor visualization output device. When the operator mode is enabled, the connection between the applied physical output hardware representation and the VE can be visualized (see Figure 6.6). When the "floor visualization output" is enabled, a virtual representation of the output hardware is visualized in the VE. Users can thus easily see the registration of hardware devices and the virtual scene and interactively change transformations between them.

Integrated assessment scene	*Bird's-eye view*
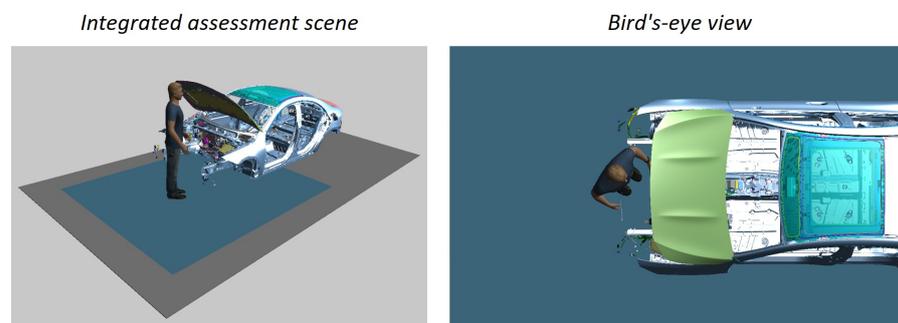
Figure 6.6: Blue floor plane represents the coupling / registration between floor visualization hardware and the VE. (left) Integrated assessment scene, (right) bird's eye view

6.3.5 *Interaction concepts*

Building a car virtually requires 3D (3 DoF and 6 DoF) interaction capabilities, such as natural user interfaces [116], [117], [120]. The virtual camera can be manipulated in three ways. Using regular HID, such as a mouse, the camera position (viewpoint) can be manipulated using a left click and orientation with center click press-and-hold. Zoom capability is realized using the mouse wheel. For more intuitive interaction, all 6 DoF can be manipulated at the same time using a space mouse controller with position, orientation and zoom features. Another way of intuitive view point and orientation manipulation is using Kinect (see Chapter 7) tracking capabilities. Kinect tracking can be attached to the viewpoint of the user's head position and orientation. In this way, CAD data can be explored interactively and intuitively. The user gets the impression that the virtual camera is directly attached to his head.

Furthermore, **manipulation of objects** is enabled using standardized protocols. When standardized tracker protocols are used, a wide

range of commercial peripherals can be employed in veo for object assembly simulation. So called Virtual-Reality Peripheral Network (VRPN) allows for a device-independent and network-transparent transmission of virtual reality peripherals [143] and is a standardized protocol for tracking data and both digital and analog controller inputs. Besides time synchronization and various predefined commercial trackers, VRPN also offers multiple simultaneous connections to devices. In addtion, ART protocol is a pseudo-standard for commercial multi-DoF tracking devices and controller inputs. Both protocols are implemented as an input interface for object tracking peripherals. These trackers can be attached to the dynamic assembly parts, also defining an offset to each of them. When a tracker is attached to the assembly part, tracking data can be used to join the assembly part and product in a realistic manner. Input trajectories of objects can be recorded easily and replayed in the VE. Interaction devices such as the HTC Vive controller can also be used as a VRPN tracker.

6.3.6 *Digital human model*

Similar to object tracking capabilities, having a DHM is crucial for the assessment of a manual assembly processes (for automotive assessments see Chaffin [144]). Ergonomic aspects, viewpoint evaluations, reachability studies and collision checks all rely on digital human models as depicted in the DHM interaction cycle in Figure 6.7. Besides the vast variety of DHMs in research and commercial products, in veo a Cal3D based digital human model is applied.
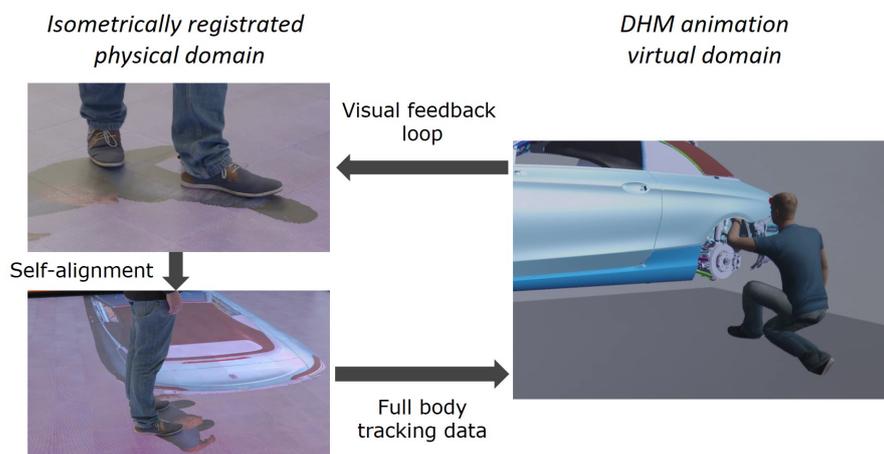
Figure 6.7: Interaction cycle for an isometrically registered augmented floor surface and DHM animation

Cal3D is a skeletal character animation library that can be used to apply different avatars to the DHM. Avatar appearances, called skins, are ported from avatar libraries such as "Rocketbox Libraries" or MakeHuman [145]. For Cal3D interactive animation of the skele-

tal rig, the same tracking protocols are used as described in the object tracking paragraph above, namely VRPN and ART. In contrast to single 6 DoF object tracking information for each trackable, for the transmission of human skeletal information each VRPN channel typically contains 18 to 77 data sets, each with 6 DoF tracking information. Each item of 6 DoF information contains a translation of a joint. This information is applied directly to the skeletal character. Using VRPN to stream all of the skeletal tracking information results in a fluent skeletal DHM animation.

H-Anim 200x or "Humanoid Animation" [146] is an ISO/IEC FCD 19774 standardized way of describing **humanoid animation**. This standard describes three levels of articulation and nominal body dimensions for a DHM skeleton. Cal3D uses "level of articulation" one resulting in 18 joints without manipulation of fingers, whereas when fusing finger tracking information, level two is applied with 71 articulated joints. Not necessarily all joints must be set by the tracking device as a subset of joints may be transmitted only. For example, Kinect V2 skeletal information delivers 21 positions - most of them containing orientation information (see Section 7.3.4). The mapping of tracking information onto a skeleton is called targeting, whereas mapping from one skeleton to another is called retargeting. This lossy registration process tries to reduce the overall error rate by adjusting lengths of the segments and mapping joint angles properly on the resulting skeleton.

With this DHM, two different full body motion capture systems are integrated, namely both ART Body and Kinect V2 skeletal tracking. In addition, leap motion finger tracking can be fused onto the full body trackers, which extends the full body skeletal tracking data to include finger tracking data.

Veo DHM allows the end effector of the human to be manipulated by selecting the hand and manipulating it using a space mouse in 3D. Inverse kinematics allows stretching and bending of the upper arm and upper body according to the required position. In contrast to more sophisticated inverse kinematics, this feature can be used for quick reachability assessments (compare other simplified inverse kinematic approaches [147].

For the assessment of **different anthropometrics**, the avatar can be altered using a population library. Body height and individual segment lengths of the DHM can be set, so that reachability studies with different character proportions can be assessed. Additionally, multi-worker processes can be interactively simulated by inserting up to six DHMs. Even though standardized tracking protocols allow the transmission of additional annotation data besides the skeletal tracking data, no re-identification of tracked persons is implemented (compare markerless user identification using body lengths by Hayashi et

al. [148]). Therefore, the DHM avatars are assigned to the tracking data in the order of appearance.

6.3.7 *Virtual reality implementation and user interface*

Virtual reality using head mounted displays has been researched for several decades [149]. VR implementations in the manufacturing industry must serve an overall purpose, such as higher quality validation results [80] or shorter task completion times. In general, VR allows users to be immersed [150] in the VE and can generate the sensation of presence in the virtual environment. Nevertheless, this is not a sufficient reason for utilizing VR in the manufacturing industry.

Veo enables virtual reality assembly validation and is optimized for usage with the "HTC Vive Pro" HMD hardware. Using VR, the overall goal is to achieve better planning quality, as the whole assembly process can be interactively validated. Typical examples of interactive VR optimization aspects are packaging, visibility, assemblability, production ergonomics, process quality, process efficiency, logistics, walk paths and many more.

In VR, the **operator performs the entire assembly task** according to the process plans, from picking the virtual assembly part out of the carrier, boxes and racks, carrying it to its geometric destination, assembling it in the product and, if required, using a virtual tool such as a screw-driver. The car is rendered in its respective manufacturing state. At the beginning of each cycle time, assembly parts of this cell are located in the racks and carriers. The car continuously moves at a constant speed on the virtual assembly line through the virtual cell. This provides operators with useful insights into the geometric circumstances in the workstation, the overall process flow and the upcoming product. For example, situations with bad visibility, bad ergonomics or too narrow clearances can be revealed.

The non-VR workshop participants can observe the VR user's activities from a third-person perspective on additional renderings. All participants are able to follow the assembly procedure. This represents a collaborative, co-located asymmetric output situation (also see Gugenheimer et al. with ShareVR [109]).

6.3.8 *VR user interface*

Just like the veo desktop UI, the entire assembly simulation operation can be controlled inside the VR environment as well. The operator sets VR user's height to the corresponding body height once at the beginning. The hands of the DHM are geometrically attached to the controllers to allow self embodiment of the VR users. When both DHM "full body tracking" and VR are activated at the same time, the VR con-

troller positions overwrite the left and right arm tracking information, due to the higher tracking accuracy of the controllers.

Assembly parts are defined as dynamic objects and can be grabbed by approaching the virtual object with a controller and then picking the object (see Figure 6.9 right). Both "press-and-hold" and "click-to-grab" have been implemented and these variants are assessed in a usability study below. When grasping and mounting the assembly parts, a vibro-tactile feedback is given by the controllers. When the user successfully finishes the current assembly step, the next work task is activated automatically with the corresponding dynamic objects.

For the direct assessment of different process variants in VR, a process switching possibility is included as a **non-diegetic overlay menu in 3D space**. Figure 6.8 (left) shows the 2D menu with its option for switching cells, workstations, work tasks and its corresponding dynamic assembly parts. In addition, the full text of the work task description is displayed in the upper right hand corner of the menu. When opening the menu, its orientation is orthogonal to the user's viewing direction and is placed at a distance of 4m in his viewing direction. Both fixed position and floating position menus are implemented. The advantage of fixed position menus is that the user is able to approach to the menu after opening it up.

As the VMS is intended to be used as an isometrically registered virtual environment, there are no infinite tracking frustum and movement space for VR usage. Therefore, a **virtual teleportation functionality** has been implemented (see Langbehn et al. [107] for VR teleportation effects). Figure 6.8 (center) depicts this "reposition mode" in VR space. The user enables the reposition mode on the controller. Rays point straight from the controller. When the user triggers the teleport functionality on the controllers, his/her new location corresponds with the intersection point of the ray and the floor plane. This allows the the user to move quickly within the VE - without moving physically. Changing the orientation around the up-axis can be achieved by using the controller's circular touchpad in the relevant direction.

Another VR feature for assembly assessments are **visual location indicator cues** for assembly instructions. Oftentimes geometrically similar assembly parts must be assembled. Figure 6.9 (right) shows a visual approach indicator of the dynamic assembly part. At the part destination a semi-transparent duplicate of the part is visualized and highlighted by blinking. The closer the VR user brings the object to the final position and orientation, the more the semi-transparent object changes its color from red to green. During the final approach of the assembly part within a radius of 1 cm and 5° difference, a snapping functionality helps to assemble the part in the final position, determined by the product geometry.
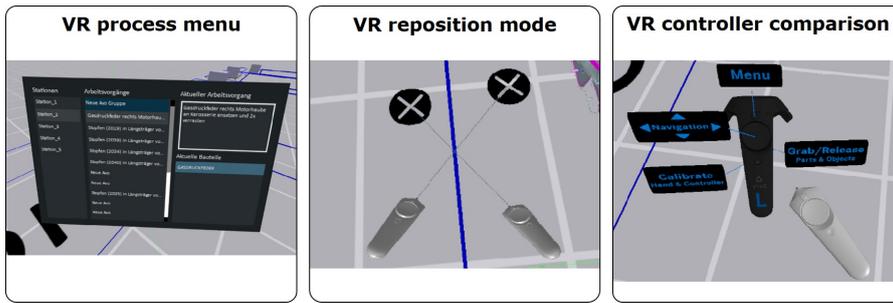
Figure 6.8: veo VR user interface components: (left) VR process menu, (center) VR reposition mode, (right) enhanced virtual controller representation

Figure 6.9 (left) depicts the **visual assembly part discovery cue** to determine the position of an assembly part or assembly goal in 3D space. Both controllers visualize a blue ray which points towards the goal, depending on the work step progress. Before picking an object, the rays point to the assembly part's position in 3D space (i.e. part in rack), whereas after picking a part, the ray points towards the final destination (i.e. trunk of the car). This allows even untrained operators to identify the correct assembly parts and find their destinations easily even if there are multiple similar parts.
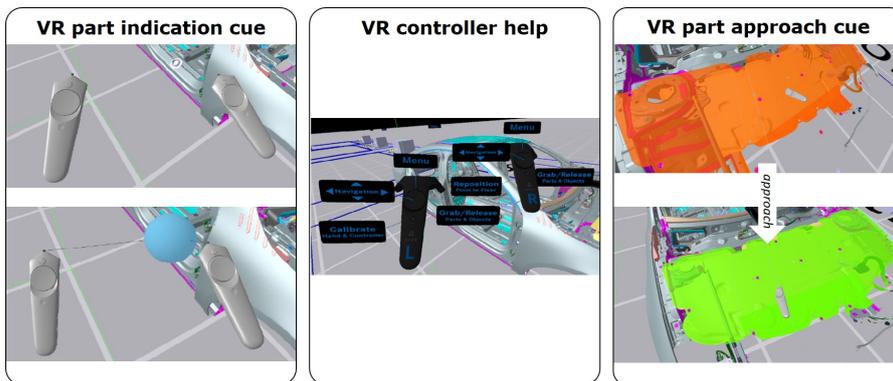
Figure 6.9: Optimization of veo VR user interface components: (left) visual part indication cue, (center) VR controller help, (right) assembly part approach cue

6.3.9 *Informal VR evaluation and optimizations*

After the veo VR capabilities had been shown to visitors of a public trade fair, 18 guests filled out the feedback questionnaires. All users were asked to write down their optimization ideas informally, as a type of "thinking out loud" exercise. The following optimization potentials were revealed and have been implemented:

As described in the features section, the DHM is registered to both the controllers and the full body tracking system. On the one hand,

this provides a good sense of embodiment in VR, whereas on the other hand full body tracking jitter caused some motion sickness. Thus a way of **hiding the DHM embodiment only in VR** has been implemented.

Furthermore, an **alternative representation of HTC vive controllers** has been modeled. Figure 6.8 (right) shows the standard HTC vive representation, which corresponds to the physical controller as well as the alternative model. This representation is helpful for a plausible VR experience but the virtual origin is not made clear enough. This leads to confusion when gripping small assembly parts, with dimensions smaller than the controller dimensions. As a solution, the vive controller has been remodeled so that the lower parts still correspond to the physical geometry, whereas the upper parts clearly indicate the virtual origin of the controller with an accurate point. The origin sphere has a diameter of 1mm.

Novice VR users asked for a **help functionality on the controllers** as the functionalities were explained to new users once but they were unable to remember all user interface components immediately. That is why virtual controller labeling has been added as a diegetic 3D controller description. All button functionalities are described as 3D labels next to the controllers, as can be seen in Figure 6.9 (center), such as "Grab" on the trigger button, "calibrate" on the side button, "Menu" on the top button and "Navigation" on the touchpad. Textures of the virtual controllers have been altered to help users distinguish between left and right controllers (see Figure 6.9 center).

Users also suggested optimizations for the **VR process menu interface**. Font size has been adapted to real data and the process menu placement has been reduced to the fixed position variant, since the floating position variant was opposed by all users due to limited HMD resolution and blurriness in the edges of HMD.

6.4 VIRTUAL REALITY ASSEMBLY ASSESSMENT BENCHMARK

When VR is implemented in veo for PV use cases, immersion is not an end in itself but must be beneficial to the overall assessment goal. In the literature, immersion is described as one of the main advantages of VR, but for professional use of VR, users expect to achieve their goals either with a higher quality or in a more efficient way. Multiple papers propose using VR for better immersion and better spatio-temporal understanding of the upcoming production process (compare Bowman et al. [150]).

In this chapter the "VR assembly assessment" (VR2A) benchmark is proposed as a unified experiment design, in order to quantify the practical VR system's performance without measuring the VR interaction cycle influence parameters. More precisely, VR2A measures the user's ability to visually assess the assemblability of the digital mock-

up (DMU) with respect to two independent variables: Assembly part sizes and clearances. The user represents both the operator and the product assessor at the same time, just as in real PV workshop situations. VR2A measures whether production engineers can achieve their assessment goals even with small parts and low clearances and how small both may be.

6.4.1 *State of the art*

Nowadays research presents many purpose driven VR applications, such as excitement in gaming [109], positive emotions for point of sale applications [151], novel rehabilitation methods [152] in medicine, more effective learning in schools [153], [154] and of course VR in automotive production. For the manufacturing industry, Zimmermann presents a brief overview of VR use cases in his survey [155] as well as Ottoson [156] throughout the product development process. Lawson et al. discuss future directions of VR for automotive manufacturers in a survey of 11 engineers, which shows further VR development needs [157]. Berg and Vance present an overview of the application scenarios in product design and manufacturing [158]. Multiple academic publications on VR in automotive production are presented in the following topics: Production verification and maintenance (see Gomes de Sá and Zachmann [26]), training use cases [159], [160], product design and packaging [158] and continuous improvement process [82].

All of these use cases share the same goal: They apply VR technology for a better spatio-temporal understanding and immersiveness for users. Basic VR research presents the effects immersion on behavior in VEs and its effectiveness. Immersion creates a feeling of presence in the VE or of "being there" and is often described as "the outcome of a good [gaming] experience" [161]. Jennett et al. have researched immersion experiences in games and found that immersion can be measured both subjectively using questionnaires and objectively by measuring task completion time or eye movements [161]. Interestingly, Ellis [162] doubts that presence might directly lead to better task performance, for instance when a more abstract view of an environment is required in flight control use cases, for achieving the goal. Beforehand, Witmer et al. present the widely known "presence questionnaire", which became a standard for measuring presence in VR [30] and is also applied in this study. Bowmann and McMahan ask, how much immersion is enough in VR [150] and give an overview on empirical studies which show that full immersion is not always necessary.

Overall, the literature does not include any uniform experiment design as a benchmark for VR assembly assessments for quantifying the VR system's limitations. Most closely, Funk et al. describe a uniform experiment design as a benchmark for evaluating interactive instruc-

tions using augmented reality for assembly tasks [163], which differs in the benchmark scope, since Funk et al. evaluate task completion times whereas VR2A is intended to quantify the geometric limitations. Therefore, research currently does not provide any answers on how to measure the limitations of such VR assembly assessment systems.

6.4.2   *Influence parameters on the overall VR purpose*

The VR interaction cycle consists of tracking devices, simulation software, rendering pipelines, hardware devices and of course the user. Each of those components inherits various sources of errors, unpredictable behavior and influence parameters. Figure 6.10 depicts a simplified VR interaction cycle including exemplary error influence parameters for each component. The following exemplary error sources limit the overall VR system's performance:
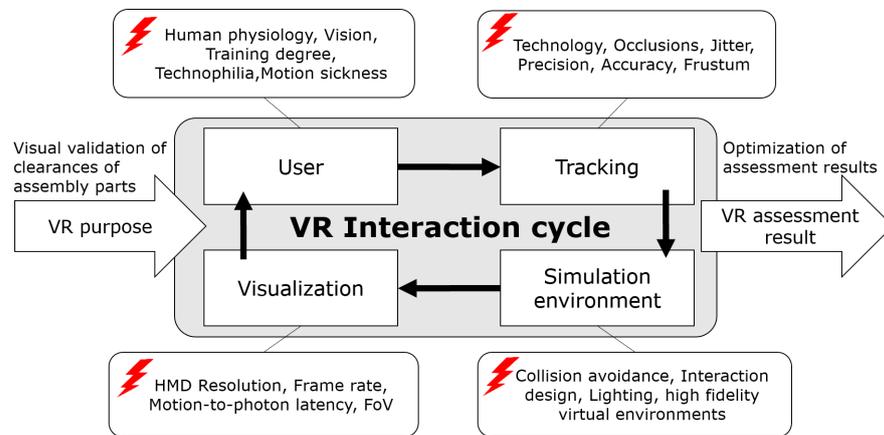
Figure 6.10: Block diagram of VR interaction cycle including error influence factors

- **Stable and precise tracking** is crucial for a good VR experience. All tracked components need precise 6 DoF tracking. Typical limitations of the tracking system are optical occlusions, limited spatial frustum and limited tracking precision, jitter and accuracy.

- The **simulation software** also introduces multiple sources of errors in the interaction cycle, such as unsuitable usability, rendering issues, scene lighting, simulation software properties and missing collision detection and avoidance.

- **VR visualization devices** such as HMDs have a limited field of view, limited motion-to-photon latency, limited framerate and resolution. This is why visualization additionally induces errors in the interaction cycle itself.

- Finally a major influencing factor on the overall system performance is the **user** himself/herself. For fulfilling the overall VR simulation purpose, he/she must be able to interact with the entire system, which means the respective training degree can be a potential source of errors. Additionally, limitations in physiology, vision and perception in general will influence the overall VR assessment results, such as human tremble or uncorrected vision.

As the above non-exhaustive list of errors shows, there are too many influencing parameters to control every single one of them. Nevertheless, users are not interested in quantifying these VR system properties but want to know whether they can reach their VR assessment goals efficiently. That is why, from a production engineer's perspective, each single error parameter presented in Figure 6.10 is less important than the overall VR system's performance. The respective error parameters in the interaction cycle can be regarded as a black box with an overall limitation for reaching the assessment task. Therefore, when using a VR2A benchmark, the system is tested for its applicability towards its native purpose.

6.4.3   *The VR2A benchmark*

VR2A is proposed as an open standardized experiment design for the evaluation of a VR system's overall geometric limitations for assembly assessment scenarios and is considered to be "quick and easy." The VR2A scene is publicly accessible here: `https://skfb.ly/6FQOV`

Two parameters are varied in an abstract assembly task: Clearance and assembly part sizes. By conducting the VR2A benchmark, the user gains quantified insights into how small the assembly parts and clearances can be to still obtain reliable assembly assessment results by production engineers. VR2A specifically abstracts all abovementioned influencing and error parameters within the interaction cycle and only focuses on the assessment results relevant for assembly: Assessment of clearances and part size limitations.

VR2A carries out an abstract assembly task, inspired by a kids' game called "shapes sorting toy" (see Figure 6.11). The virtual reality scene has been published to set VR2A as a standard benchmark. As depicted in Figure 6.12, within the virtual environment, there is a static table, six static discs each with five cavities on a wall. As depicted in Table 6.1, six dynamic, graspable cubes are placed on a table.

All six discs are placed on the wall, which are rotated horizontally and flipped at randomized angles. Each disc contains five cavities corresponding to the sizes of the cubes (see Figure 6.11). Each disc has five cavities at a size of 97%, 100%, 103%, 105% and 110% relatively to the corresponding cube size (see Figure 6.11 bottom). For example,
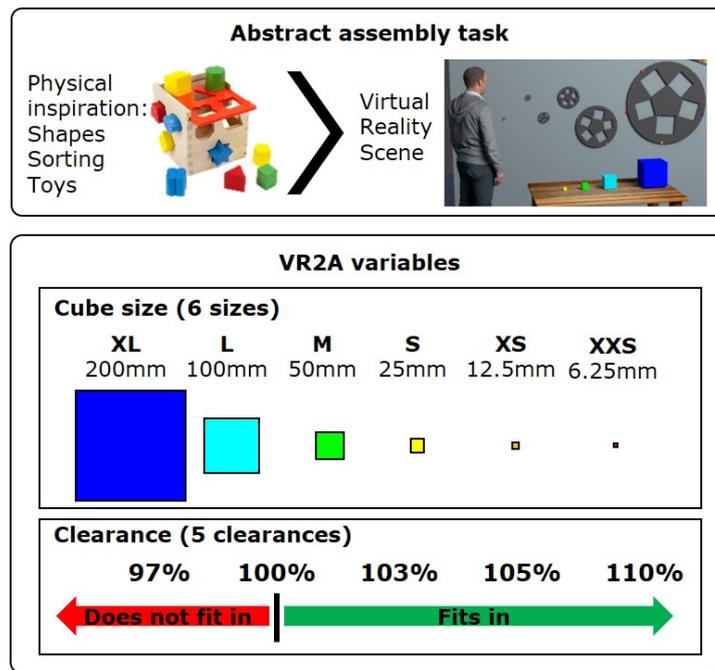
Figure 6.11: Overview on the standard experiment design of VR2A and the two independent variables: Size and clearance

Table 6.1: Description of cubes in VR2A benchmark

Cube	Size	Color
XXS	6.25 mm	red
XS	12.5 mm	orange
S	25 mm	yellow
M	50 mm	green
L	100 mm	cyan
XL	200 mm	blue

the XL disc's 100% cavity exactly matches the size of the XL cube. The L cube does not fit in the respective "97% L cavity", but the S cube does fit in the respective "103% S cavity."

The procedure of the benchmark has a straightforward design: Each participant inserts all six cubes in each of the five corresponding cavities of the matching disc size. For example, the L cube must be assembled in all five L disc's cavities in randomized order. Unlike the experimenter, the user does not know the correct answer. Possible answers are "Fits in," "Does not fit in" and "I can't assess it." The experimenter tells the participant that the goal is not to insert the cube without collision but to assess correctly whether it could be mounted that way – according to real production validation tasks. The scope of this task does not include task completion time.

The results are calculated as follows: Each of the three possible answers is sorted into matrices containing the relative frequency for each condition. These relative frequencies of the answers "Fit in" ($A_{Positive}$), "Does not fit in" ($A_{Negative}$) and "I'm unsure" ($A_{Neutral}$) are calculated. Equation 6.1 calculates the relative homogeneity of answers between the assessments. If $S_{Homogeneity}$ equals zero in the matrix, the value of 0% would indicate that the same number of people state "Fits in" and "Does not fit in." Therefore, the assembly assessment would not include any reliable results.

$$S_{homogeneity} = abs(A_{Positive} - A_{Negative}) \qquad (6.1)$$

The overall VR2A score S_{VR2A} additionally penalizes "I'm unsure" feedbacks by the participants (see Equation 6.2). Therefore, the VR2A score can be interpreted as the overall uncertainty for each variation of size and clearance.

$$S_{VR2A} = (abs(A_{Positive} - A_{Negative}) - A_{Neutral}) \qquad (6.2)$$

S_{VR2A} can therefore theoretically range from -100% to 100%. Using these results, the overall VR system limitations can be explored using VR2A. Setting an individual threshold of, say, 80% VR2A clearly illustrates how small assembly parts and clearances may be in order to achieve the personal VR assessment purpose.

6.4.4 *Setup, stimuli and design*

In this study we use the VR2A benchmark to evaluate the overall performance of a veo VR assembly simulation system applied in the automotive industry. This is used to carry out validations on assemblability. Even though automotive products and the resulting assembly paths can be more complex, this abstracted assembly task gives useful insights into the system's performance.

The hardware setup consists of a HTC Vive Business Edition (110° field of view, 2.160 x 1.200 resolution) attached to a high-performance Intel Core i7-8700k PC, 16GB RAM with a GTX 1080 TI graphics card. The tracking devices are calibrated in accordance with the technical specifications. The open VR2A scene is loaded in a proprietary assembly simulation software program called veo. This software natively supports the HTC Vive headset via OpenVR. Assembly parts (VR2A cubes) are set to dynamic objects. No physics, collision detection or gravity are turned on during the evaluation. Participants use the HTC Vive VR controller. Its virtual representation is visualized 1:1 but ends in a sharp cone as the root point to allow participants to perform grasping operations with the highest precision (see Figure 6.12 right).
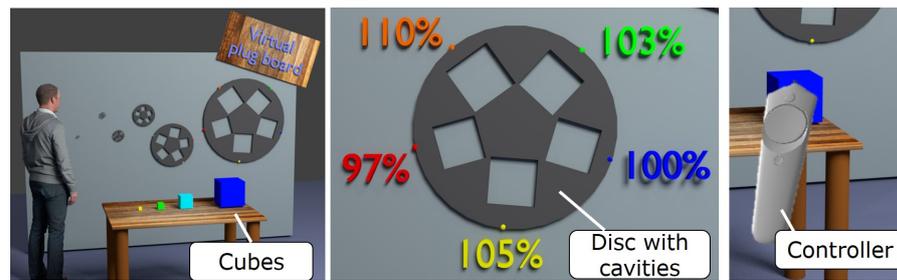
Figure 6.12: (Left) Rendering of the open virtual environment with six differently sized cubes. (Middle) Explanation of disc cavities relative to the corresponding cube sizes, which are not visible to the user. (Right) Controller with sharp grasping point

6.4.5 *Participants and procedure*

For this study, 32 production validation workshop participants were selected on a voluntary basis, such as research engineers, ergonomics experts, production engineers and students - all working for various planning departments in an automotive OEM company. This means that this study was carried out with the intended key users of the system. As users are an important performance factor in the VR interaction cycle, the overall population within the study should represent the population of users, for example for PV workshops. They did not receive any special rewards for taking part in this study. 24 male and 8 female participants took participated, ranging in age from 18 to 51 years (M=28.2, SD=6.7). All participants reported normal to corrected vision.

The experiment consists of two parts, the VR2A experiment and a final questionnaire. The experiment takes approximately 25 minutes per user, 10 minutes for the VR2A evaluation itself and 15 minutes to fill out the questionnaire. The experimenter describes the assembly task in a standardized way. The participants are asked to familiarize themselves with the VR environment, the controllers, the virtual scene and the dynamic handling of the cubes by playing around with them. When the respective participant feels confident with manipulating the virtual scene, he/she completes all 30 VR2A assembly tasks. Starting with the biggest cube (XL) and progressing to the the smallest (XXS), each cube must be inserted in all five corresponding cavities of each disc, but the experimenter randomizes the order of the cavities. For each cavity, the user verbally tells the experimenter the result of his/her visual assessment and whether the cube fits into the cavity without collision. If required by the VR user, the experimenter adjusts the vertical height so that the user always has a comfortable view of the discs. After finishing the assembly task, the participant fills out questionnaire, consisting of five non-standardized assembly experi-

ence questions and two standardized questionnaires, the "Prescence Questionnaire" and the "System Usability Scale."

6.4.6 *Results*

The VR2A benchmark gives insights into the limitations on size and clearance when performing a VR assembly assessment task. Figure 6.13 depicts the relative frequencies of the according answers "fits in", "does not fit in", and "uncertain." Hence, for clearances >100% the objectively correct answer is "fits in" whereas for <100% clearance scenario, the objectively correct answer is "does not fit in" since cubes overlap with the disc. For a 100% clearance scenario, the expected answer would be "uncertain" as the cube theoretically the cube fits, but practically in VR the cubes cannot be placed in a mathematically correct position without any overlap. Interestingly, for the "100% clearance scenario", an average of 63.02% of the participants answered "does not fit in" whereas only 26.56% answered "fits in." Only 10.42% answered "I don't know".
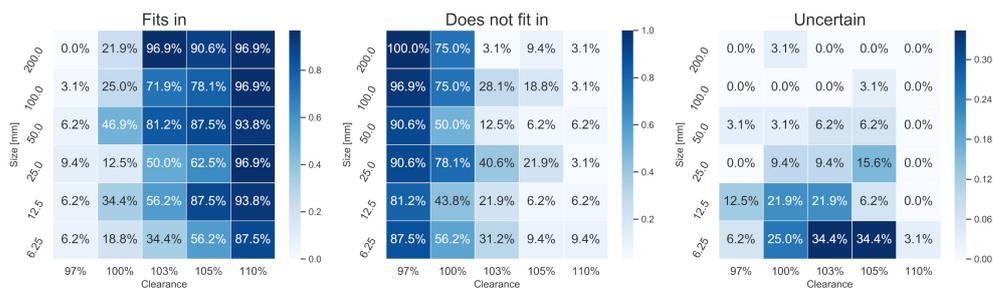
Figure 6.13: Relative frequencies of the participants' answers in VR2A benchmark over the different scenarios

The data presented in Figure 6.13 is the source data for calculating the VR2A score using Equation 6.2. Results are shown in Figure 6.14. Low scores indicate high uncertainty and inhomogeneity of answers. The lowest VR2A value can be found in scenario 6.25mm sized cube with 103% clearance with the value of -31.2%. Highest values have been found for the biggest cube in 97% scenario: All participants recognized correctly, that the 200% cube does not fit in.

Plotting the mean VR2A scores over one of the two independent variables provides interesting insights into the assessment performance of the participants. Figure 6.15 plots mean VR2A scores over the cube sizes in non-percentage values. One can clearly see that the VR2A positively correlates with the size of the cubes, as indicated by the 2nd polynomial regression. For a 6.25mm cube size the mean score is only 28.75% whereas the 200 mm cube averages at 83.75%.

Figure 6.16 plots the mean VR2A results over the absolute clearance scenarios. Low mean VR2A scores can be found for the scenarios 100% (26.04%), 103% (30.21%) and 105% (54.1%). For both the 97%

Overall VR2A score
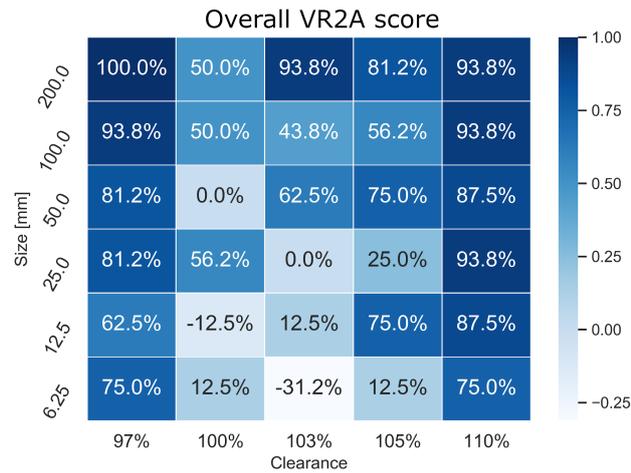
Figure 6.14: Results of the VR assembly assessment score. Low values indicate high uncertainty or inhomogeneity of answers

and the 110% scenarios, the scores are higher at 82.29% and 88.54% respectively.

In the open questionnaire, people responded to free questions using a 5-point likert scale from 0 (strongly disagree) to 4 (strongly agree). They tended to be able to carry out collision checks purely visually without the help of technical collision avoidance (M=2.46, STD=0.97). In accordance with the objectives results, subjectively the participants also saw an increasing assembly complexity with decreasing object sizes (M=3.25, STD=1.39). Participants also stated that complex insertion trajectories can be assessed in VR (M=2.84, STD=0.87) and they were able to understand manufacturing processes better than when using a conventional desktop PC (M=2.97, M= 0.95).

Additionally, users reported on their favorite grasping method using the VR controllers. 53% (17) of the participants preferred the "click and hold" grasping method, whereas 47% (15) preferred the "click-to-grasp and release" method. The standardized questionnaire "System Usability Scale" [28] scored 84.58 with 31 participants. According to [164], the usability of the VR assembly simulation system can be interpreted as "good."

6.4.7   *Discussion and practical conclusions using VR2A insights*

From a practical standpoint, the VR2A benchmark helps production engineers decide on how reliable their assessment must be. They can define their own personal threshold and can thus easily derive how small the parts and their clearances may be. For instance, they can set their required VR2A threshold to 80% and can get a rough estimate of whether the assembly part can be assessed correctly, e.g. at the 150mm level or whether positive clearances should be bigger than
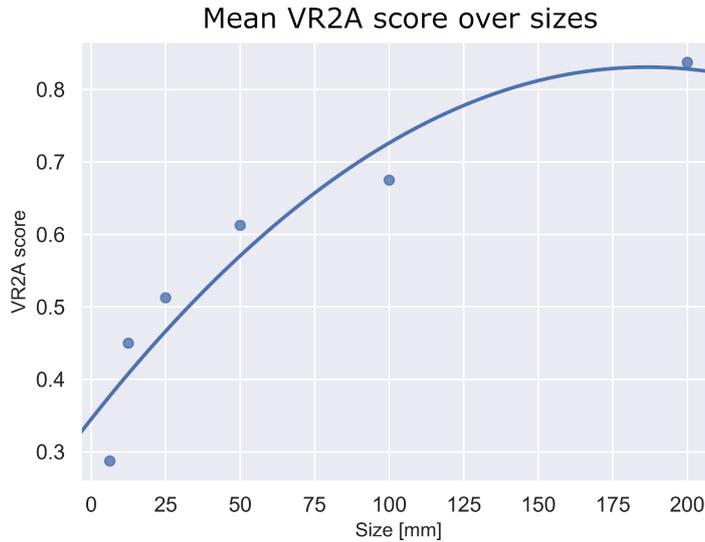
Figure 6.15: Mean VR2A score over size scenarios with the respective 2nd polynomial regression

110% (see Figure 6.15 and Figure 6.16). In contrast to robust parts with large clearances, the VR2A threshold can be set lower to 50%.

Results also indicate that the same negative clearance can be detected more easily compared to the same positive clearance. The mean VR assembly score for a 97% percent overlap performed a great deal better than the 103% clearances. Even when comparing 97% overlap to 110% clearance values, they performed almost identically in terms of mean VR assembly scores (see Figure 6.14). In general, the maximum uncertainty was expected at no tolerance scenarios (100% clearance), whereas the 103% clearance cavity led to the overall smallest VR2A values.

Results indicate that even though participants are encouraged to tell that "I can not assess it" is a valid answer, they still tend to give a judgmental answer such as "fits in" or "does not fit in," even though there is no clearance at all.

Subjective feedback from participants indicates potential reasons for these system limitations: Human tremble and resolution of VR HMD: For the cube sizes XS (12.5mm) and XXS (6.25 mm), the vast majority of participants started holding the VR controller in both hands to reduce human tremble. Tracking accuracy still seems to be more stable than human tremble for small cube sizes. Therefore, in this evaluation human tremble is currently the limiting factor for improving assessment performance (in comparison with HTC Vive precision and accuracy also see Niehorster et al. [165]). Additionally, for the smallest cube size (6.25 mm) all clearances are on a sub-millimeter scale. Even though participants could move their head as close to the discs as necessary, the VR HMD resolution was mentioned as the subjectively limiting factor.
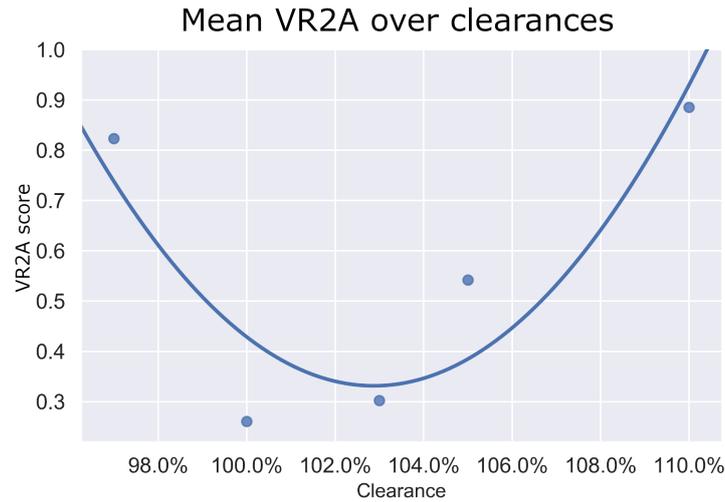
Figure 6.16: Mean VR2A score over clearance scenarios with the respective 2nd polynomial regression

On the other hand, four participants actively told the experimenter that assessing large cubes is harder than small cubes due to the necessary big head movement for assessing clearances.

6.4.8 VR2A summary

In the "Virtual Reality Assembly Assessment," the (VR2A) benchmark is a standardized open source experiment design to evaluate the overall VR system's assembly assessment performance and limitations. VR2A can be universally applied for different environments, simulation software and VR hardware devices. All experts are encouraged to assess their own assembly assessment system using the open source VR2A scene. This allows production engineers can gain practical insights into their next VR assembly assessment simulation. The evaluation showed that VR2A is a reliable benchmark for quantifying the overall assessment performance and for revealing its limitations in assembly. The use of VR2A in production validation in the automotive and manufacturing industry makes validation results more reliable. In the context of the VMS and veo as the central PV simulation software, in-depth insights can be obtained regarding how well the system performs in VR and where its limitations are.

6.5 CO-LOCATED PMUS AND DMUS IN VR ASSESSMENTS

As presented in Section 5.2.1, the VMS's key property "integration of PMU and DMU" can be seen as realized by using the PV simulation software "veo". Using tracking components in combination with the registration steps of the VE and physical world, a holistic VR cell assessment with combined PMU and DMU can be realized. Figure 6.17

depicts a realization example of the key properties' implementation within the VMS concept:

- **Asymmetric visualization:** PV workshop participants can choose whether they want to use the simulation as an immersive HMD-based first person view (Figure 6.17 upper left) or as an exo-perspective of the whole scene (Figure 6.17 right).

- **Markerless tracking:** Via standardized, stateless protocols (see ARVIDA project report [92]) markerless object tracking information is externally polled (Figure 6.17 lower left) by Fraunhofer IGD markerless object tracking. This 6 DoF tracking data is continuously updated in the simulation environment. Additionally, fused full body motion capture data from the multi-kinect system (see Chapter 7) is polled via the same ARVIDA protocol and brought into the same world coordinate frame via a least-squares estimation of transformation parameters [166].

- **PMU-DMU integration:** For example, when registering the world coordinate systems of the VE, the VR tracking system and markerless object tracking system together, participants see the DMU door in the HMD and can also feel haptic feedback from the PMU door at the same time. When the PMU door is moved, the DMU door in VR moves accordingly.
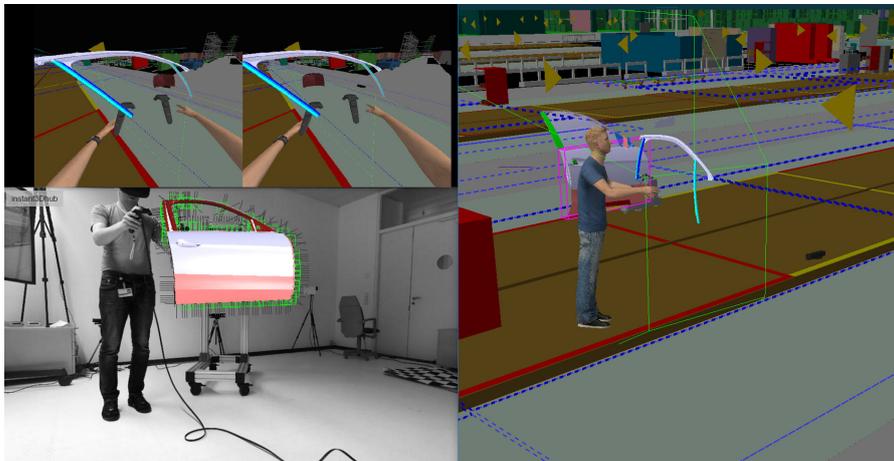
Figure 6.17: Evaluation of a co-located VR experience combining a tracked PMU, markerless full body motion capture and an immersive HMD view with 3D interaction devices referenced relatively to the digital human

The usage of true-to-scale PMUs in VR scenes has multiple advantages for users:

1. **Haptic and physics feedback:** As physical items build a barrier, users cannot grasp through the PMU or through the DMU.

Additionally, PMUs have a certain weight and therefore accelerations and velocities can be realistically transferred to the virtual domain.

2. Tracking of the PMU allows for an **intuitive manipulation** of the VE. Users can simply rearrange PMUs in the physical domain and do not have to manipulate/teleport DMU components in the virtual domain. True-to-scale PMU representations can be seen as the ultimate tangible user interface (compare Tangible Bits by Ishii [102]) for natural user interfaces.

3. Users can easily **align themselves within the virtual domain** since it matches the relations in the physical domain as well. For example, when walking one step towards the door, users expect the virtual domain to behave in the same manner (see key property true to scale in Section 5.2)

6.6 SUMMARY

In this chapter, multiple research interests have been considered. A description of the needs for batch assessment simulation software in PV workshops have been provided as well as a novel VR benchmark:

Insights have been presented on how the simulation of entire final assembly lines can be realized. Multiple different data sources must be fused, such as CAD product data including product variants, product kinematics, manufacturing sequences, task assembly descriptions, cell layouts, factory CAD data and of course resources, tools and auxiliaries. The outcome of the automated authoring process and the presented assessment features allow the instantaneous batch simulation of hundreds of workstations.

Especially the rendering features of orthographic bird's eye views generate the images needed for the LHRD floor visualization presented in Chapter 8. An immersive user interfaces has been presented, so that an efficient overall factory simulation including all processes is made possible in VR.

Since accessible external research did not give a clear answer on the overall limitations of VR systems to be applied in assembly assessment scenarios, a novel open source, standardized experiment design has been proposed, the so called "virtual reality assembly assessment" (VR2A). This benchmark is useful for understanding the overall VR system's limitations for assessment scenarios. Evaluation results using VR2A showed a correlation between product size and assessment quality. Hence, the evaluation also showed that for the same clearances, overlaps (negative clearance) can be detected more reliably than positive clearances. For the VR2A benchmark, additional research must be carried out on the effects of more complex assembly part geometries, for example balls, stars, triangles or screw-shaped ge-

ometries, and other parameters, such as task-completion time. Third-party researchers are encouraged to conduct the VR2A benchmark themselves.

MARKERLESS, SCALABLE FULL-BODY MOTION CAPTURE SYSTEM

As the VMS framework aims to establish an interactive collaborative virtual environment, input is at least as important as output. Therefore, the physical user's movements are intended to be tracked to enable animation of DHMs and to allow natural interaction in 3D scenes. DHMs are used to visualize production processes [15] and thus assess ergonomics [167], process plausibility, assembly part visibility and other geometric assessments. For example, in trim and final assembly of an automotive car production several hundred workstations are continuously optimized, and the ergonomically critical workstations are assessed using digital simulation tools for each new car derivative.

Offline motion generation tools are based on different motion synthesis approaches. Multiple simulation tools rely on key frame-based animation of DHM movement and interpolate the animation between these key frames such as Delmia V5. These manual animation methods are not used in a broad manner due to "time consuming modeling, inflexibility to changes on process and product, and to often occurring unnatural movements of the DHM if dynamic processes are modeled" [123]. Additionally, these methods require considerable CAD expertise as well as knowledge of processes and products: "It is necessary to know ergonomics but also to have CAD skills and to have a detailed knowledge of the various features of the product being designed/evaluated" [168]. Other methods try to address time consuming key-frame based movement generation by decomposing tasks into basic operations [123]. These tools have an integrated predefined library of actions (e.g. ema [169]) or deterministic motion synthesis algorithms (e.g. IMMA [170]). Typically, these movements are parametric and concatenated one by another, but generated movements do not appear highly realistic due to missing parallelized sub-actions and non-continuous movement flows.

Therefore, in this chapter an online, interactive motion generation approach is presented for the animation of DHMs, namely a scalable, markerless tracking system. Full body skeletal tracking information is captured in real-time and applied to interactively manipulate DHMs by retargeting skeletal tracking information onto the character model in the simulation environment. Conventional, commercially available marker-based tracking systems, such as A.R.Tracking, OptiTrack or VICON, require a great deal of preparation prior to a full body tracking session. Grimm et al. indicate that users might regard marker

suits as obtrusive which also influences the realism degree on how they interact with the virtual scene (see [171, p. 103-105]). As ranging cameras are becoming more affordable, markerless body tracking has become a feasible option for production validation simulation. Being an alternative to more expensive motion capture systems, depth cameras can also be used for gestural interaction, natural user interfaces and of course motion capture. Depth camera based systems have quickly become an appealing alternative for marker-based full-body motion capture in industrial applications, when overcoming the limitations and meeting the requirements. These markerless optical outside-in methods use commercial depth-camera systems, but their performance is limited with respect to tracking frustum size and accuracy.

Related publications

In the following sections, a set of methods for multiple depth-camera registration and heuristics-based sensor fusion using skeletal tracking is described, including its applications in an industrial environment. Parts of this chapter have already been included in following publications, such as "Journal of Virtual Reality and Broadcasting" (JVRB) [172], "EuroVR 2015" conference [173] and in "CIRP CMS 2015" production engineering conference [174]. The development of this approach has been carried out by the author in cooperation with several students' theses and in cooperation with the institute of media informatics at Ulm University. Furthermore, a publication on the real usability of the system was presented in the paper "Applicability Evaluation of Kinect for EAWS Ergonomic Assessments" at CIRP CMS 2019.

The remainder of the chapter is structured as follows: The chapter starts with the VMS's requirements in a motion capture system for full body skeletal tracking to be used for production validation simulations. After an in-depth literature review of state-of-the-art multi-depth-camera systems, an accuracy analysis of Kinect v2 skeletal tracking is provided in which a robot moves a mannequin for accurate, reproducible motion paths. Based on the results of this evaluation, a distributed hardware consisting of multiple Kinect v2 sensors and service-oriented software setup with a set of registration and fusion techniques to create a ready-to-use tracking system is presented for real-time interaction with virtual environments. The system's technical evaluation shows that the system presented helps to increase tracking areas, resolving occlusions and improving human posture analysis by analyzing the spatial accuracy of registration performance. The implemented system is further evaluated with respect to its applicability for carrying out standardized ergonomic assessments EAWS in production validation workshops. Finally, a set of op-
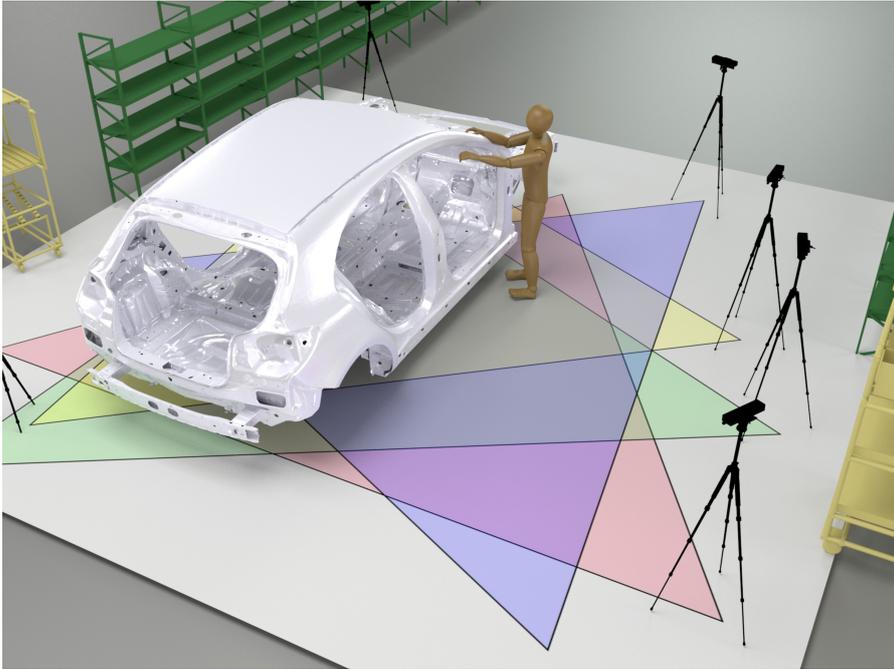
Figure 7.1: Workshop area in automotive car assembly with full body skeletal tracking, typically sized 6m x 4m. This also represents the required tracking area

tical light PMUs for low occlusion are presented, developed to be used with optical outside-in tracking systems for industrial workstation assessments.

7.1 APPLICATION SCENARIOS AND REQUIREMENTS OF SKELETAL TRACKING

When carrying out detailed assessments in PV workshops, DHMs are used for ergonomic assessments, visibility assessments, buildability assessments, population assessments and many more.

Users perform real movements of the pre-planned processes. PMUs can be included in the assessment but do not necessarily have to be present. Physical representations of the upcoming product are helpful in obtaining haptic feedback, even if product representations are outdated compared to the DMU version. Such true-to-scale PMUs are placed in the workshop area, for instance a car body (see Figure 7.1). Additionally, physical racks with shelves, material wagons and operation material may be present in the workshop environment. All parts that are physically available are intended to be integrated in the virtual simulation scenario [15].

In this environment, the tracked user must be represented as a DHM in the assessment simulation software, which robustly follows his movements throughout the process. Passive PV workshop participants, such as ergonomic experts, are not tracked at the same time,

but alter simulation variables, such as work height or process variants, to optimize the process at the respective upcoming workstation. Supporting PV assessments with novel methods, multiple requirements for a suitable motion capture system can be derived:

Room-scale tracking frustum: Various applications in industry and research require large-scale tracking systems, e.g. for interaction with virtual environments. Since the concepts of the VMS propose an isometric, registered VE, the tracked users must be able to move around in the entire tracking space. Therefore, the tracking system must be scalable so that the sensor arrangement can be adapted to local requirements: If a small workshop room must be tracked, fewer tracking sensors are applied. If a bigger room is available, technical operators must be able to integrate more sensors and track the entire room. The minimum tracked area for a standard workstation in final assembly should be at least 6 by 4 meters.

Multiple tracked users: In virtual assessments, real workstations have to be simulated. Workers cooperate and rely on each other for successfully finishing the process, especially when mounting large and heavy parts, such as the front module of a car. That is why the virtual assessment and the tracking system must support multiple tracked workers, at least 2 tracked workers at a time.

Instantaneous tracking: Conventional marker-based optical tracking systems rely on marker suits. These are cumbersome, time-intensive to put on and obtrusive [171, p. 103-105] and therefore might influence the user's movements. In practice, in order to reduce changeover times only one workshop participant used to be the tracked. A novel motion capture system must be able to track users instantaneously without any preparation process. Users must be able to step into the VE without putting on special suits.

Low-cost Tracking Components: For industrial use-cases, cost efficiency and short amortization periods are crucial. Tracking systems with marker based approaches involve considerable initial investment. Cost efficient tracking systems allow a broad roll-out in production engineering.

Interactive frame rate and latency: The motion capture system must deliver tracking data at interactive speeds. In virtual assembly scenarios, interactive speeds are defined as update rates of at least 30 Hz, whereas for immersive VR assessment environments using HMD visualization update rates of 60 - 90 Hertz are recommended [171, p. 223].

Level of DHM articulation: DHM animation for PV tasks requires accurate full body motion capture data. The workgroup for ISO 19774 Humanoid Animation [146] proposes three levels of articulation in terms of the number of bones and feature points. The maximum level of articulation proposes 109 skeletal joints. Depending on the validation task in the VMS, various levels of articulation are required. First

and foremost, body tracking of limbs and the head is required, finger tracking is optional. Face tracking is not required.

Accuracy and precision: The required accuracy and precision of the tracked user's body depends on the use cases. Reachability assessments must resemble size and overall body posture as accurately as possible whereas very small buildability checks in small and cluttered build volumes within the product require accurate, jitter-free hand tracking.

Standardized protocols: Multiple simulation tools are used in PV workshops. The tracking data must use standardized protocols, such as VRPN, ART and dTrack. Besides translation and orientation, interaction inputs must be able to transmit via the protocols, such as button clicks, analog input, inertial measurement units with gyroscopes, accelerometers and magnetometers or other means of sensory information.

7.2 OBJECTIVE OF MULTI-SENSOR ARRANGEMENTS

Using active ranging camera systems, such as Microsoft Kinect v2, Intel IntelliSens, ASTRA Orbbec, is a promising way to fulfill the abovementioned requirements for a tracking system in the VMS. However, **single** ranging sensors have multiple limitations:

1. Limited field of view (horizontal and vertical)

2. Depth ambiguities

3. Limited emitter intensity

4. Limited sensing range

5. Only one line of sight

6. Susceptible to reduced reflections caused by reflection parameters of materials

True-to-scale simulation of an entire car assembly workstation requires a spatially tracked area of at least 6m x 4m. This impedes the use of a **single depth camera** as the tracking frustum is too small. A limited sensing range, high susceptibility to self and external occlusions and a greatly varying sensing performance depending on the user's posture and position are some of the major drawbacks that must be faced if such systems are used in the VMS scenarios mentioned.

In order to overcome these limitations, a novel multi-depth-camera system is presented consisting of multiple "Microsoft Kinect v2" sensors. It extends the sensing range and improves overall tracking performance.

Therefore, a number of technical challenges must be addressed: First of all, a common coordinate frame for the cameras must be established by registering them to each other. Afterwards, the data coming from different cameras must be combined in a meaningful way to actually achieve improvements in tracking performance and range. Finally, fused skeletal data must be provided to the VR/AR systems via standardized tracking protocols, such as VRPN, trackD, dTrack or ARVIDA.

The overall system must prove whether it can be utilized for DHM process validation and of course for standardized EAWS ergonomic screenings, in accordance with PV verification tasks as described in Section 3.3.5.

7.3 STATE-OF-THE-ART IN FULL BODY HUMAN MOTION TRACKING

"Motion capture is a technique and a process that digitally record the movements of a live 'performer', such as human motion or the movement from an animal in a specified 3D environment" [175]. Zhou and Hu present a survey of various types of human motion tracking for rehabilitation, but focus on technical aspects of motion capture in general [176]. Motion capture systems are tracking systems which can be classified in multiple ways, such as precision, sensing distance, discrete or continuous events, frame rate, or more typically by their physical sensing medium [171, pp. 97-98]. Research has determined multiple ways of tracking objects and thus enabling full body motion capture. Bishop et al. [177] classify tracking systems by their physical medium, whereas Zhou and Hu differentiate between "Non-Visual", "Visual" and "Robot-aided" human motion tracking systems [176, p. 3]. Besides more theoretic approaches, such as mechanical, magnetic and acoustic measurement systems, two physical measurement methods are widely used, namely inertial and visual measurement systems:

7.3.1 *Inertial full body motion capture*

Pure **inertial full body motion capture suits** are based on accelerometers, magnetometers and gyroscopes [178] and can be used for ergonomic assessments in industrial environments [179]. Since position is not measured directly, fusion and filter algorithms must compensate drift effects with plausibilization and calibration routines. No absolute position can be measured without initial references.

Roetenberg et al. analyze a full body motion capture system called XSens [180]. These full body motion capture suits consist of 17 IMUs. Each IMU unit has an integrated 3 DoF accelerometer, magnetometer and gyroscopes. The calibration routines allow drift compensation

and plausibilization of noisy data. This system is commercially available and widely used in the manufacturing and entertainment industry. For industrial use cases, multiple side effects of this technology prohibit its use in the VMS: Cumbersome change of IMU suits, calibration routines and drift effects which are getting more intense when having large magnetic objects such as a car body close to the IMU sensors.

Hybrid systems integrate the benefits of inertial and vision tracking technologies [181]. Commercially available hybrid systems, such as the AR-Tracking "hybrid suit" [182], combine optical tracking with inertial tracking. If the line of sight is blocked, hybrid trackers deliver IMU sensor data and vice versa. Due to the geometric size of these units, no finger tracking possibilities are currently available.

7.3.2 *Optical full body motion capture*

Visual full body motion capture systems attempt to recover 3D articulated poses by utilizing image sensors. These sensors reconstruct the object's 3D position by triangulation using two or more cameras. A digital skeleton consists of concatenated tracking points fulfilling additional constraints in a digital human model (see H-Anim [146]).

Optical tracking systems for full body motion capture are typically set up as outside-in arrangements. **"Outside-in" sensor arrangements** consist of fixed sensor positions facing inside the tracking frustum, while the tracked subject moves relative to the sensor array. A line of sight between the sensor and trackable is required for optical tracking systems. Cameras must be calibrated and extrinsically registered to each other in order to create a common spatial reference.

As depicted in Table 7.1, two main differentiations can be made in optical full body motion capture measurement methods: The measurement method can either be active or passive and the users can be tracked either with or without markers:

Active camera sensors in combination with **markers** have proven to be reliable in industrial applications for a couple of decades. These systems use retro-reflective balls as markers that are illuminated by IR emitters or directly active LED-based markers. All systems require a fixed outside-in calibrated camera array to gather video streams and fuse tracking results to reconstruct the tracked 3D position of the markers. To track 6 DoF bodies, at least three markers must be joined rigidly so that both the position and orientation can be tracked uniquely. Companies such as PhaseSpace, VICON, Optitrack or ART [183] produce active camera systems which have synchronized flashes, continuously updating the marker positions at high frame rates of up to 960 Hz and high accuracy. Nevertheless, marker-based systems have a disadvantage immanent to the system, which is that users

Table 7.1: Classification of optical full body motion capture systems

Measurement method	Prerequisites	
	Marker-based	**Marker-less**
Active	**Sensors:** Active IR-camera	**Sensors:** RGB-D ToF Sensor RGB-D Structured Light Sensors
	Examples: VICON, Optitrack, PhaseSpace, A.R.Tracking [183]	**Examples:** Shotton [184], [185]
Passive	**Sensors:** Greyscale camera	**Sensors:** RGB camera Greyscale camera
	Example: Sementille [186]	**Examples:** Cao et al. [187], Hasler et al. [188], [189]

must put on marker suits. This process is time-consuming and may influence the user's motion.

Another hybrid optical- and IMU-based tracking technology is used by the tracking system of the HTC Vive VR System. They use static outside-in anchors which emit time modulated laser beams, and active trackable devices demodulate this signal for precise positional and orientational tracking inside the tracking frustum. The active tracking markers fuse this modulated signal with additional IMU data.

Passive outside-in MoCap systems using **markers** on the body are rarely used due to their limited practical applicability. Sementille et al. [186] present an outside-in tracking system attaching fiducial markers to the tracked subject with limited impact.

Passive, markerless human motion tracking methods would be the ultimate tracking solution for motion capture data if they worked reliably. The literature shows that passive sensing without any markers continues to be a research topic. These systems consist of commercial off-the-shelf cameras. In 2009 Hasler et al. presented a system for the statistical tracking of human poses and for estimating body shape for multiple unsynchronized cameras [188], [189]. This system is now commercially available as "TheCaptury." In 2016 Cao et al. presented a system called "OpenPose" for single-camera pose estimation using "part affinity fields" [187]. This open source project by CMU University gathers context information of the scene. Its real-time performance is invariant to the amount of people tracked in the single

video stream. The algorithm learns to associate body parts with individuals and estimate their body posture. "Algorithms/systems still need to be improved to compromise robustness and efficiency" [176], and having evaluated the tracking performance, the results are not as reliable as required for PV workshops.

The following realization focuses on a MoCap system **without markers** that uses **active sensors**. Most of these "commercial off-the-shelf hardware systems are ranging cameras. In comparison to the above mentioned passive or marker-based categories, ranging cameras are expected to meet the requirements of PV workshops in the VMS. Novel methods are developed to overcome technical limitations by concatenating several depth cameras and thus generating a scalable sensor network.

7.3.3 *Microsoft Kinect v2 depth-camera sensor*

Markerless tracking is enabled by both hardware and software components. Optical measurement systems generate reliable image information, either grey scale, color or depth images. Whereas regular cameras generate intensity images for a certain spectral bandwidths by accumulating light throughout a period of time per pixel, ranging cameras generate depth images, where each pixel represents the distance between the sensor and the respective objects [190]. These images are called 2.5D images because they transport three-dimensional information in a two-dimensional matrix.
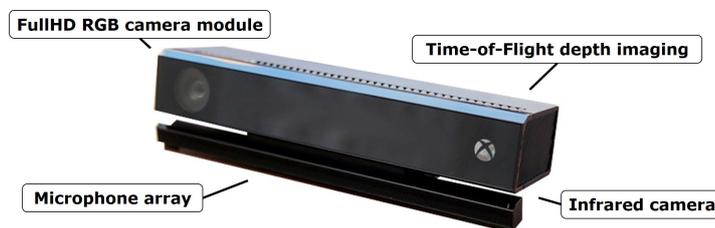
Figure 7.2: Microsoft Kinect v2 sensor with RGB-D sensor array and microphone array

One ranging camera technology is Time of Flight (ToF) measurement which permits the no-contact acquisition of depth information in a respective field of view by actively illuminating the scene with a continuous wave intensity modulation approach [191, p. 1]. "The time-of-flight system modulates a camera light source with a [square] wave. It uses phase detection to measure the time it takes light to travel from the light source to the object and back to the sensor, and calculates distance from the results." [192, p. 49] These cameras use intensity modulated emitters at the modulated frequency f_{Mod} and the scene objects reflect this light. "The phase reconstruction is performed using multiple phase images with different phase shifts which is equivalent

to sampling the inherent correlation function at different locations"
[191, p. 1]. This means that the distance d is calculated by evaluating
the phase shift $\Delta\emptyset$.

$$d = \frac{c\Delta\emptyset}{4\pi f_{Mod}} \tag{7.1}$$

Since the modulation function is typically symmetric, the phase
shift is limited to a range of $[0, 2\pi]$ for unambiguous results. Either am-
biguities must be resolved by software or the modulation frequency
must be adapted. Sell and Connor show that the overall depth reso-
lution is influenced by adjusting the modulation frequency [192]. The
overall result is a topological map of the scene, also called the depth
image, and an intensity image of the reflected light intensities.

Microsoft Kinect V2 [193] sensor was presented by Microsoft in
June 2014 (see Figure 7.2) and unites both RGB and depth (RGB-D)
sensors in one device. As described in detail by Sell and O'Connor,
Kinect V2 uses a square wave modulation signal. "A differential pixel
distinguishes the time of-flight sensor from a classic camera sensor"
[192, p. 50], which is able to detect phase shifts between the emitted
and attenuated received reflection in each pixel. Kinect V2 uses three
different carrier modulation frequencies at the same time (120 Hz, 80
MHz and 16 MHz) in order to simultaneously resolve both the prob-
lems of ambiguities in a large depth area and precision. The resolu-
tion of the ToF sensor is 512 x 424 pixels with a 70° horizontal and 60°
vertical FoV (field of view). The ToF sensor is capable providing a reli-
able reconstruction of depth information at distances of 0.5m to 4.5m,
whereas Sell and O'Connor see "smooth depth readings out to 16 m
[. . .] without wrapping" [192, p. 52]. The regular RGB camera sensor
has a resolution of 1920 x 1080. Both sensors have floating update
rates of approximately 30Hz. They are delivered pre-calibrated, so
that the correspondences between the depth image and color image
can be calculated automatically. Thus intrinsic and extrinsic camera
parameters have been determined already in the production process
[192].

7.3.4 *Microsoft Kinect v2 skeletal tracking*

Coming with this ToF hardware "Microsoft Kinect v2", "Kinect 2.0
Software Development Kit (SDK)" [193] offers full body tracking algo-
rithms including human body posture tracking, facial tracking and
hand tracking: "Tracking as many as six people and 25 body joints
per person, the tracked positions are more anatomically correct and
stable" [193] compared to Kinect v1.

This human motion tracking reconstructs fully articulated skele-
tons, namely three dimensional human motion (see Figure 7.3). Zhou
et al. [176] distinguish and reference model-based (stick figures mod-

els, volumetric models, deformable models) or feature-based (dependence graph based algorithms, local feature based, global feature based) tracking approaches [176]. Kinect v2 SDK motion tracking primarily relies on the publications of Shotton et al. [184], [185]. In 2012, Shotton et al. [184] presented two algorithms to estimate human body posture in super-real-time and independent of his clothing, body height or stature: "Body part classification" and "offset joint regression." Each algorithm independently outputs a list of confidence weighted 3D joint positions at each frame and for each skeletal joint. Both approaches use depth images as input data, and randomized decision forests for classification to deduce logical regions. Having derived the silhouette of the tracked person with image segments, body regions of the tracked subjects can enrich this with additional information. "Realistic, and highly varied synthetic set of training images" [184] is used in order to train the classificator with many different statures and of course a large variety of possible body postures. They varied 15 base characters, poses, rotation and translation, hair & clothing, weight & height, camera position & orientation and camera noise. For enabling a massive parallelization, per-pixel calculations and segmentation help to perform calculations on the graphic processing unit (GPU) so that the calculation can be accelerated. Using this generated data, a mean-shift procedure proposes skeletal joint angle proposals also with regard to temporal continuities and other skeletal restrictions for plausible human movements.

Using Kinect body tracking, no calibration routine is needed when the subject enters the tracking frustum. The complete calculation process can run per frame repeatedly at a speed of 200Hz on an Xbox 360 [194]. Kinect SDK delivers a fully articulated human skeleton with 25 joints. Each hand is only represented with three joints with a rough estimation of the thumb position. The remaining four fingers are merged to one position. Subsequent studies in 2014 by Qian et al. [195] presented algorithms to track hand and finger motions using a 26 DoF hand skeletal model from depth images. They approximate the hand posture with 48 spheres and define a fast cost function. In contrast to full body tracking, these algorithms have not been integrated in Kinect v2 SDK.

7.4 STATE OF THE ART IN MULTI-DEPTH CAMERA SYSTEMS

As the VMS's full body motion capture requirements cannot be met using one Kinect v2 camera only, the following section presents research in the field of multiple depth-camera systems. However, these publications mainly focus either on certain applications or specific technological aspects of such systems, thus leaving some of the integration and application challenges largely unaddressed. These aspects can be clustered into the following groups:

- System architecture

- Interference handling

- Registration

- Fusion algorithms

Depending on the application scenarios, some publications discuss their handling of application specific issues such user identification or world coordinates registration.

7.4.1 *System architecture*

Multi-camera approaches have already been proposed for the first generation of Kinect sensors. Each sensor is connected to a separate computer, thus simplifying the interfacing of each sensor. However, studies by Schönauer [196] and Martinez-Zarzuela et al. [197] also implement distributed systems, in which skeletal and depth data is gathered on camera nodes and sent to a central fusion unit. This component handles the creation of a common view of the tracking space. In addition, solutions have emerged in early states, which allow to stream Kinect data via network, e.g. by Wilson [198]. As distributed systems, those approaches greatly improve scalability, however at the cost of increased complexity. For Kinect v2 sensors, Rietzler et al. [199] propose a distributed tracking architecture for the so called "FusionKit." The system presented in the below also uses distributed approaches, although different interfaces. All information can be requested via service-oriented Representational State Transfer (REST) interfaces such as those proposed by Keppmann et al. [200] in order to handle additional complexity while maintaining scalability.

7.4.2 *Interference handling*

As depth cameras actively illuminate the scene, interference can occur as soon as tracking frustums overlap since any camera also receives light emitted from other cameras. There are two main approaches to interference handling that can be found in the literature, (1) optical multiplexing (e.g. presented by Butler [201] or Faion et al. [202]) and (2) post-processing algorithms e.g. hole-filling as in Maimone and Fuchs [203]. Often it is also possible to simply ignore interferences when using certain camera types and setups, especially in skeletal tracking applications where high frequency noise does not directly affect tracking performance. The system presented in the latter uses ToF depth cameras which generate only negligible interference noise due to their working principle and slightly different modulation, and no countermeasures against cross-camera-interference are implemented.

Direct sunlight exposure has proved to be another source of interference, presumably due to overexposure of the IR imaging system. As the PV workshops in the VMS take place indoors only, no countermeasures against interferences by direct sunlight exposure are taken into account.

7.4.3 *Registration*

One of the main challenges in multi-depth-camera systems lies in establishing a common coordinate frame by determining rotation and translation of the cameras to each other. Various approaches have been used for this, ranging from methods adopted from the 2D computer vision domain, horn-based methods such as those proposed by Wilson and Benko in [204] or checkerboard-based approaches such as those discussed by Berger et al. [205] or Zhang et al. [206], over Iterative Closest Point (ICP) (see [207]) approaches [208] to skeleton based (ICP-like) methods in more recent publications by Faion et al. [202], Asteriadis et al. [209], Baek and Kim [210] and Kaenchan et al. [211]. Most of these methods yield comparable results according to their evaluation, however, they vary greatly in the ease-of-use and setup time with different approaches. The approach proposed below focuses on reduced setup times and an easy setup procedure while maintaining high precision. Thus, a combination of multiple registration approaches from the studies above is applied for VMS multi-camera tracking.

7.4.4 *Fusion*

Having established a valid registration, skeletal tracking data from different cameras can be converted to a common coordinate space; however, body tracking skeletons remain individual and separate. To benefit from such a setup, data fusion methods can be employed to gather an improved view of the tracking space. Possible methods range from simple best-skeleton approaches, to joint-counting approaches [212], substitution approaches ([213], [214]), weighted averaging methods [202] and [211], to dedicated fusion algorithms e.g. by Yeung et al. [215] and Asteriadis et al. [209], which respect data quality and the specific tracking situation. This helps when dealing with occlusion and sensing limitations. Combining the advantages of each previous study mentioned, a set of novel fusion heuristics is discussed presented and analyzed in terms of its performance against a ground truth.

While covering many of the relevant aspects, most of the previous work leaves out important factors of a multiple depth-camera system for universal use. In general, registration and fusion approaches also lack end-user optimization for PV workshop use in the VMS as well

as a comprehensive evaluation of underlying assumptions, e.g. for factors influencing registration and fusion methods and quality. Currently missing insights are provided below, which have proven to be useful for a multiple, scalable depth-camera system.

As scalable, multi-depth camera approaches continue to be researched, multiple solutions have been presented over time. Studies by Otto et al. [172, 173] and Geiselhart et al. [174] have directly influenced many other publications and cooperations as can be seen by the citing publications. Brekelmans presents "Multi Sensor Alpha v3" for Kinect v2 sensors as a stand alone software [216], without research dissemination. Besides this, a close cooperation with the publicly available "FusionKit" by Rietzler et al. [199] yielded synergies in development, but with separate systems.

7.5 SKELETAL TRACKING PERFORMANCE MEASUREMENTS

As the proprietary sensor "Kinect v2" and the respective SDK must be regarded as a black box, it is crucial to learn more initially about the utilized hardware. For this, an in-depth analysis of tracking performance is conducted as a first step, similar to earlier work for different hardware or properties, e.g. by Wang et al. [217] or Yang et al. [218].

The proposed scalable multiple depth-camera system relies on the Kinect v2 sensor and its SDK, both supplied by Microsoft. Since the combination of hardware and SDK is proprietary and therefore a closed system, several features and properties of the Kinect v2 sensor cannot be determined or influenced directly. Skeletal tracking algorithms, for instance, cannot be influenced or extended due to proprietary training sets and algorithms [184]. These body tracking algorithms estimate a skeleton consisting of 25 3D-positions including 18 orientations of up to six users (see Figure 7.3) at a refresh rate of approximately 30Hz.

However, it is possible to analyze this closed system in order to derive tailored approaches for the multi-sensor tracking system. In 2012, Obdrzalek et al. [219] performed similar evaluations of the first Kinect hardware generation which is based on structured light sensing but without the focus on building a multiple depth-camera system. Since the second generation Kinect v2 sensor is based on ToF depth sensing technology, the results have only limited validity for the new sensor which is analyzed below. For Kinect v2, Otte et al. evaluate the accuracy and reliability of tracking data with respect to clinical measurements of motor functions [220]. The presented evaluation results have directly influenced design considerations for the following skeletal fusion algorithms.

1	Head
2	Neck
3	Spine_Shoulder
4	Spine_Mid
5	Spine_Base
6	Hip_Right
7	Hip_Left
8	Shoulder_Left
9	Shoulder_Right
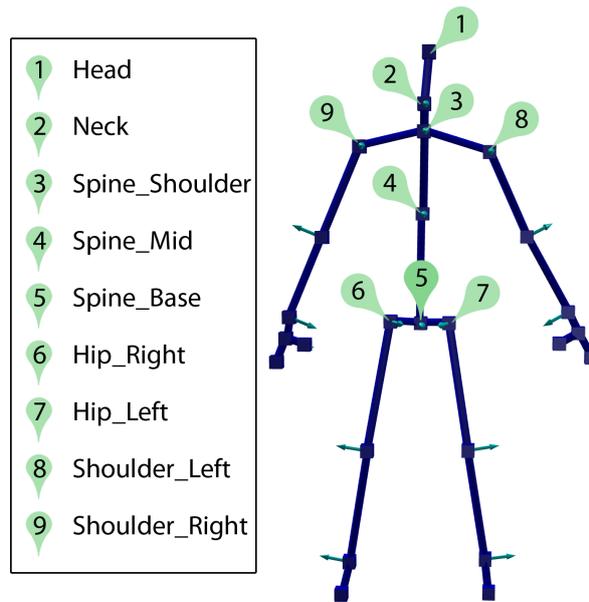
Figure 7.3: Relevant joints nomenclature of SDK body tracking

7.5.1 *Experimental setup*

To gain insights into the run time behavior and sensing performance of the Kinect body tracking system, a setup for reproducible human postures and trajectories is created for conducting blackbox-like tests. In order to achieve reproducible trajectories, a high precision robot "UR10" made by "Universal Robots"[1] is mounted horizontally on a table at a height of 1.0m (see Figure 7.4). This robot has six DoF and an operating radius of 1.3m. Each axis can be rotated by 720°. All tests are carried out at slow speeds with a maximum velocity of 0.1m/s to avoid tracing effects. The repeatability of each trajectory is specified to ± 0.1mm. The so-called tool center point (TCP) of the robot is defined in the center of the mounting plate directly linked to the last rotational axis.

Mounted to the robot arm, a 1.75m tall mannequin is used for all experiments. It is dressed in a regular t-shirt and jeans. As black clothes cause problems with the depth image (and the body tracking results) due to low IR reflection, bright colors are chosen. In all experiments, the mannequin has a symmetric posture with open hanging arms.

In order to retrieve precise and low jitter input data of the mannequin's movements, two experiments with reproducible trajectories are carried out: The 360° and 85° experiment. The latter is carried out in order to bring the rotation axis out of the skeletal center-axis, which is not possible in the 360° experimental setup due to axis limitations and occlusions by the robot itself.

---

1 http://www.universal-robots.com
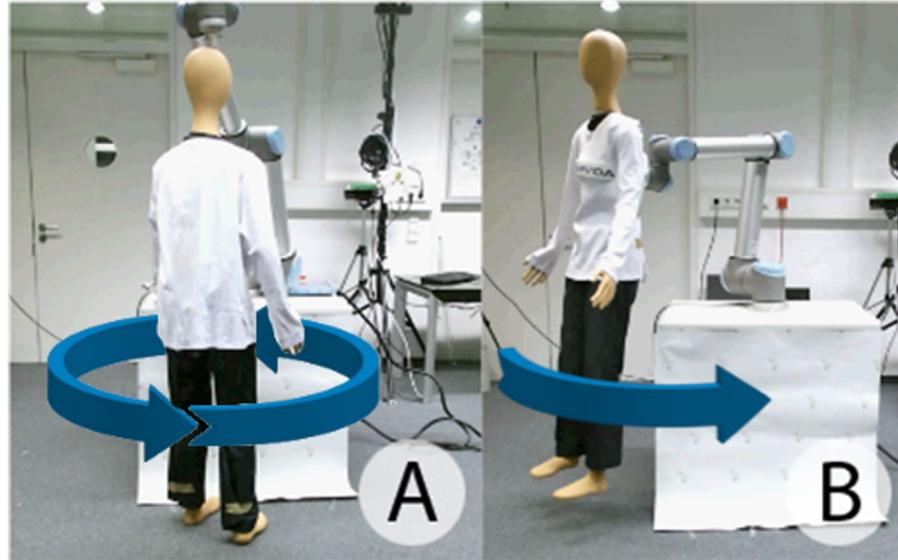
Figure 7.4: Experimental setups: A) 360° rotational test: Mannequin hanging on robot, B) Mannequin's spine mounted on robot
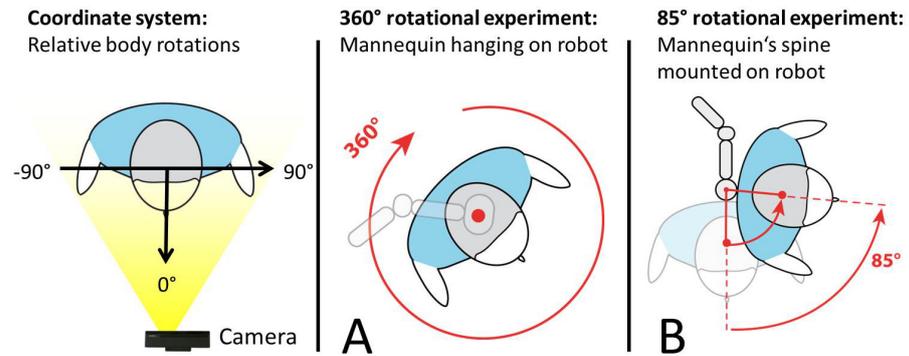
Figure 7.5: Drawing of coordinate system and two experimental setups: A) 360° rotation, B) 85° rotation

For the 360° experimental setup, the mannequin is mounted on the TCP at the top of its head with a 10cm long separator in between (see Figure 7.4 A and Figure 7.5 A). The robot spins the mannequin around its vertical axis at a constant angular velocity. With this experimental setup, skeletal orientation and positional stability can be assessed with regard to the user's rotation.

In the 85° experiment, the mannequin's spine is mounted on the robot's TCP at 1.3m height (see Figure 7.4 B and Figure 7.5 B). The mannequin's feet are 1mm above the floor. The sensor is leveled and stands at a distance of approx. 2m at a height of 1m. The robot rotates the mannequin around the TCP from 0° (orthogonal) to 85° (close to side view) as depicted in Figure 7.5 B. The 85° experimental setup is used to quantify the positional precision of the center-axis joints (see Figure 7.3) and therefore the applicability of joints for multi-camera registration purposes.
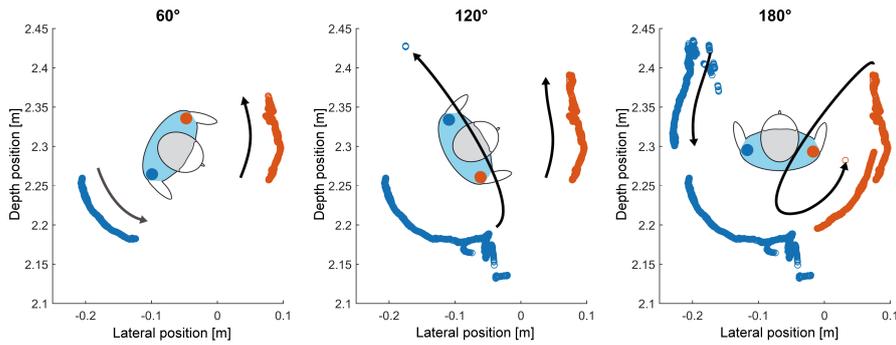
Figure 7.6: 360° rotational experiment: Front/back ambiguity during rotation for the non-center axis joints Shoulder_Left and Shoulder_Right. Color change indicates the front/back ambiguity

7.5.2 *Evaluation 360° experiment*

In general, body tracking was originally designed for gaming use cases in which the user is facing the camera. Therefore, a quickly decreasing tracking performance when slowly rotating to a lateral or even rear view is expected.

Figure 7.6 shows the course of shoulder joints in a top view over rotation from 0° and 180°. At 0° the mannequin is facing the camera as depicted in Figure 7.5. In contrast to the two expected continuous semicircles (the lower semicircle representing the right shoulder whereas the upper semicircle is related to the left shoulder), the SDK results show a discontinuous behavior. It can be observed that the position of the right shoulder jumps in the mid section in Figure 7.6 from point [−0.05m,2.2m] to [−0.15m,2.4m]. This ambiguity is attributable to the SDK assuming that the user's body is always oriented towards the camera. Consequently, an ambiguity appears when the user turns his back towards the camera. The analysis shows that this effect can be observed for all non-center-axis joints such as arms and legs. Even the non-centered torso joints, e.g. Hip_Left, Hip_Right, Shoulder_Left and Shoulder_Right, suffer from the same problem. However, center-axis joints like Head, Neck, Spine_Base, Spine_Mid and Spine_Shoulder are invariant to the front/rear ambiguity and could therefore be used for registration purposes.

Having excluded non-center-axis joints, the estimated user orientation is subsequently analyzed with regard to its validity. Figure 7.7 depicts the expected SDK behavior (turquoise) given by the rotating robot axis. Moreover, this plot shows the estimated vertical orientations for the center joints Neck, Spine_Base, Spine_Mid and Spine_Shoulder. Arm and leg joints are not included in Figure 7.7 since they only provide valid data up to 30° due to occlusion.

The SDK reliably estimates the orientation for the center joints within a range of 0° to ± 40°. Beyond this point, the user's vertical orienta-
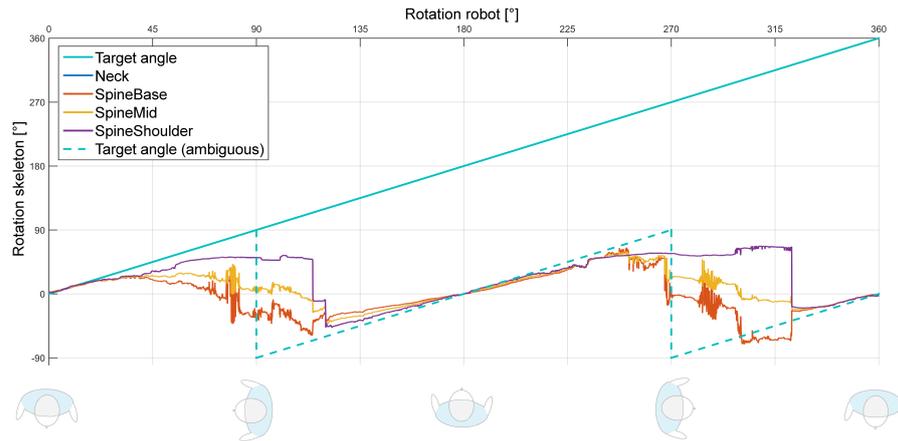
Figure 7.7: 360° rotational experiment: Expected SDK behavior vs. estimated shoulder and hip orientation

tion is increasingly underestimated until the mentioned ambiguity takes place at approx. 130°. Between 40° and 130° Spine_Shoulder and Neck orientation perform better than Spine_Mid and Spine_Base. From 130° to 260° an angular behavior, similar to the range of 0° to ± 45°, can be observed, whereas between 260° and 320° the user's orientation switches again. In Figure 7.6 the dotted line additionally depicts the expected ambiguous behavior with two discontinuities at 90° and 270°. However, comparing this expected behavior to the described results, it can be seen that these distinctive points are located 30° to 40° beyond their predicted location due to the SDK trying to maintain the current orientation of the person. Consequently, the analysis reveals that the vertical orientation of each joint can be considered as ambiguous and error-prone outside the range of 0° to ± 40°.

Figure 7.8 illustrates the mean Euclidean distances of spatial inter-frame jitter during the 360° experiment. During the entire 360° rotation, one can see jitter differences depending on the respective joints and the user's orientation towards the camera. The head joint has lower jitter compared to the other joints (mean jitter <0.072mm). Consequently, it can be assumed that the Kinect SDK filters head joint position in order to reduce spacial jitter enabling viewpoint control applications. The joints Spine_Shoulder and Neck have modest jitter below 35° on both sides (mean jitter <0.2mm), whereas in rear facing orientation (180°) the jitter is increasing (mean jitter <0.5mm). In contrast to that Spine_Mid and Spine_Base joints reveal an increasing jitter on the side view at around 90° and 270° user orientation (mean jitter <4.1mm). Therefore Spine_Mid and Spine_Base joints cannot be used for flexible multi-sensor skeletal registration purposes.

As a result, Head, Neck and Spine_Shoulder could be used for registration purposes since they are independent of the positional ambiguity and offer the lowest jitter during user rotation. Information
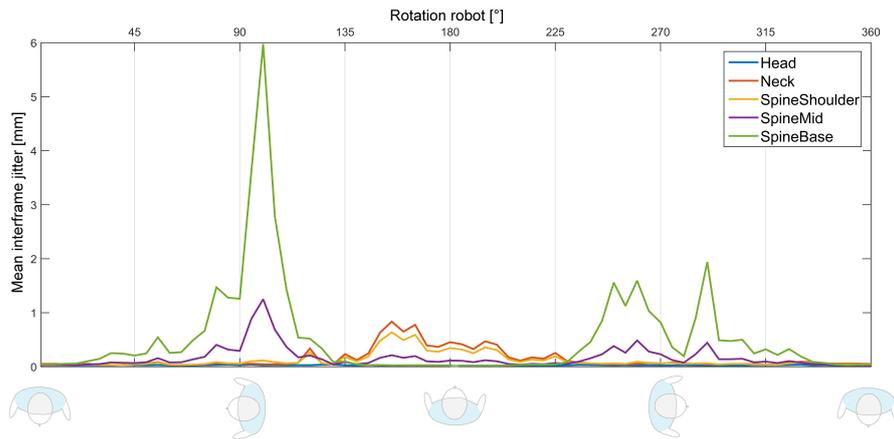
Figure 7.8: 360° experiment: Mean inter-frame jitter for all center-axis joints

about orientation cannot be used since each joint displays rotational ambiguity.

7.5.3 Evaluation 85° experiment

The 85° experimental setup compares the trajectories of the center-axis skeletal joints (Head, Neck, Spine_Shoulder, Spine_Mid and Spine_Base) to the ground truth robot trajectory as depicted in Figure 7.5B. These insights are used to understand their absolute positional accuracy when varying the human's orientation towards the sensor.

In this experimental setup, the mannequin's back is mounted on the TCP at a distance of 27mm, thus rigidly coupling both trajectories. Consequently, the mannequin trajectory is known in advance and can be directly compared to the result of the skeletal joints. Figure 7.9 shows the resulting paths of the baseline and the five joints from a top view.

The Spine_Base joint does not follow a circular trajectory. Above 60° the trajectory is getting noisy and unstable. Spine_Mid shows similar behavior. For both of these points, the distance is overestimated. Spine_Shoulder has a good circular performance up to 80° while the distances at low angles are still overestimated. Head and Neck joints match the baseline for lower angles, whereas at more than 45° the head joint tends to move forward in the mannequin's perspective and away from the sensor (almost 90°). These effects could be reproduced throughout all repetitions.

Overall, these experiments lead to the conclusion that for registration purposes the **Neck joint** of the Kinect SDK skeletal tracking provides the most stable and suitable positional data even at high rotation angles relative to the sensor. Additionally the head joint offers a low jitter 3D joint position.
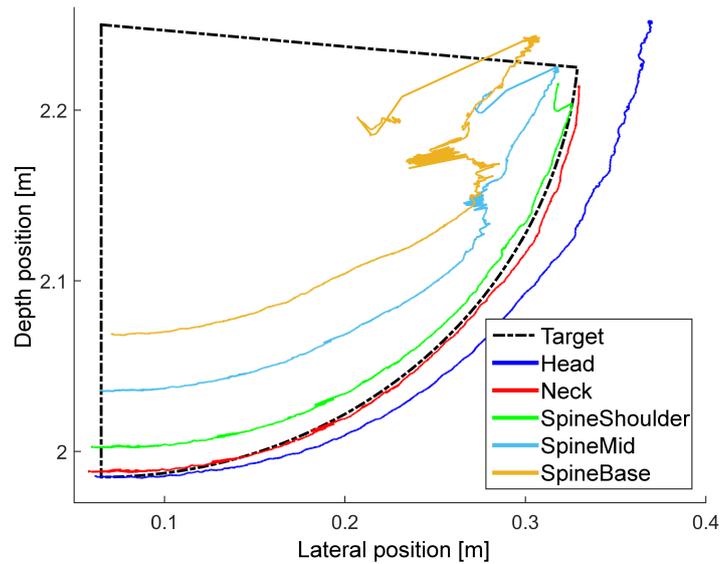
Figure 7.9: Evaluation of 85° experiment: top view of center-axis joints in relation to ground truth real trajectory
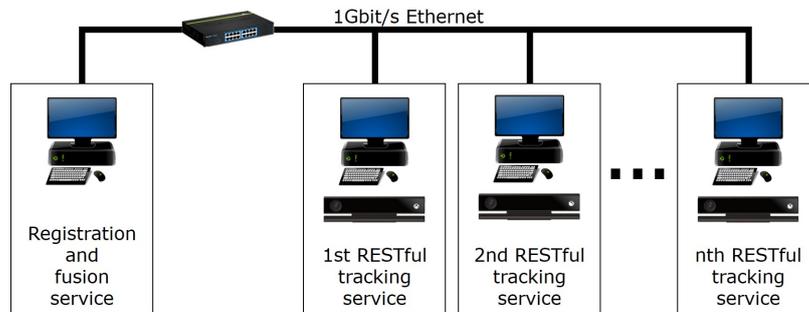
Figure 7.10: Hardware setup for the tracking system with service-oriented, distributed sensor services

7.6 IMPLEMENTATION OF MULTI-DEPTH SENSOR SYSTEM

The proposed multi depth-camera system consists of several Kinect v2 sensors. In comparison to the first generation Kinect, the Kinect v2 sensor uses Time of Flight as depth sensing technology, which induces improved depth accuracy and better interference resistance. The proposed system consists of several tracking computers accommodating the tracking services and one central fusion computer (see Figure 7.10). Each tracking computer is connected to one Kinect v2 sensor via USB 3.0 and additionally to a fast local area network.

There are two main software components: Both the "service-oriented tracking service" and the "fusion software" are described in the following:
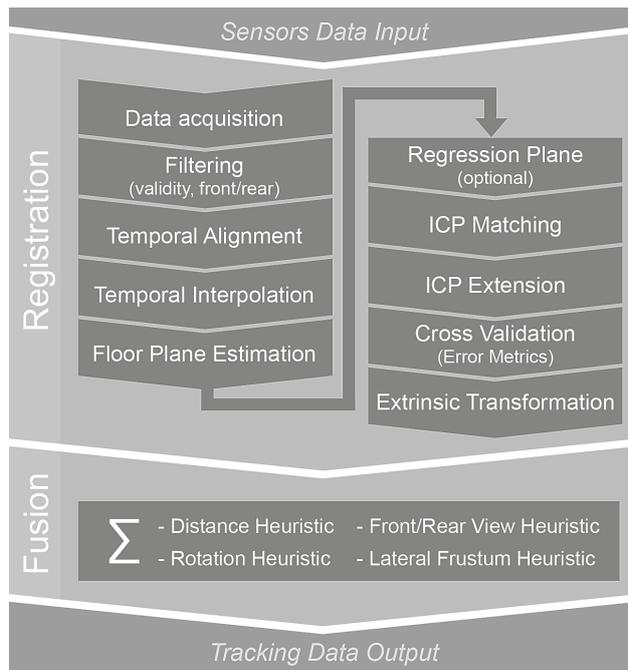
Figure 7.11: Block diagram of fusion service

7.6.1 *Service-oriented tracking service*

Implementing a service-oriented RESTful tracking service instead of conventional streaming architecture has several advantages: Third party integrators can easily reuse the services for implementing clients. Additionally, the use of standardized and publicly available tracking vocabulary and Resource Description Framework (RDF) helps achieve interoperability between tracking devices which is also the goal of the ARVIDA project. In this context, the tracking services presented use a RESTful polling-based approach with linked data which conforms to the ARVIDA standard. Keppmann et al. [200] show that RESTful Linked Data resources can be applied for virtual reality environments.

In the tracking service, information is gathered by the event-based Kinect SDK. The web service offers all skeletal information, the status of each skeleton, the floor plane estimation and color and depth camera views as RESTful resources. RDF datagrams are serialized using the Turtle format. Each datagram contains time stamps for subsequent synchronization.

7.6.2 *Fusion service*

The fusion and multi-sensor service runs as a central component and handles registration, fusion and data input/output in the tracking environment. Figure 7.11 depicts the architecture of the registration and fusion service.

The fusion service polls data of each attached tracking service. Each sensor requires a dedicated tracking service. As depicted in Figure 7.11, the fusion service first calculates the extrinsic registration and uses this registration for fusing data in real-time.

Pre-assumptions for the extrinsic registration process:

1. Only one user is inside the tracking frustum during the registration process

2. The sensors are arranged on the outside of a tracking frustum and have overlapping tracking frustums

3. The user moves within overlapping tracking frustums, so that data points can be captured over time.

During **extrinsic registration** Neck joint data is captured over time from each sensor and is used as an input point cloud for calculating the transformation matrix. Umeyama least-squares algorithm [166] iteratively minimizes the difference between two point clouds gathered by the sensors and outputs an extrinsic transformation between each overlapping pair of cameras.

Having a valid registration with low error-metrics, the **heuristic-based fusion component** is able to combine skeletal data from all registered cameras and provides them as an output to possible domain-specific application components.

7.6.2.1 ICP extension

Gathering only the neck joint information has a drawback that must be compensated: Since the user's movement takes place on the flat floor and the height of the user's neck joint does not vary a lot, the gathered point cloud data lies almost on a single plane. To compensate for this lack of variance, additional information is gathered. The floor plane estimation compensates for the missing information by using an approximation of the distance to the floor and the pitch angle of the sensor. Fusing this information with the ICP's transformation matrix (see [207] for ICP variants) offers an improved transformation for extrinsic registration between one sensor relative to the master sensor. In addition to that, a regression plane has shown to further enhance the ICP results if enough feature points have been gathered during the user's movement.

7.6.2.2 Front/rear detection

In order to achieve maximum flexibility for the hardware sensor setup, the fusion service recognizes whether the user is facing the Kinect or if he is turning his back on the corresponding sensor. A robust indicator of whether the user is turning his back towards the camera is an evaluation of the angle between the shoulder joints: By evaluating
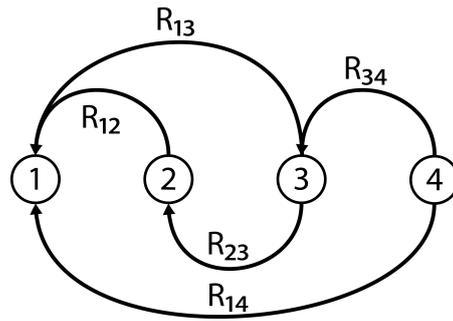
Figure 7.12: Concatenated transformation chains for a four sensor arrangement

the discrete skeletal states of the collar joints, one can determine the user's orientation to the camera in each frame. These assumptions lead to optimized RAW data correcting even laterally reversed data robustly in the fusion service.

7.6.2.3 *Scalability*

To achieve a fully scalable motion capture system with a common coordinate frame, extrinsic transformation matrices between all overlapping tracking frustums are calculated (see Figure 7.12). Missing transformation matrices indicate that no overlapping tracking frustum is used in the calibration routine (for instance R_{24} in Figure 7.12). These matrices are concatenated so that all transformation chains for each sensor are calculated to the master sensor. For N sensors sharing an overlapping tracking area, there are $(N-1)!$ transformation matrices. Having more than two sensors sharing the same tracking area the system is over-determined and a cross-validation of transformation chains is carried out with regard to the absolute transformation accuracy. Therefore an error metric is introduced which consists of the summed up and normalized Euclidean distances of the reprojection error. Based on this error value, the best interlinked transformation chain between master and each slave sensor can be determined.

7.6.2.4 *Time alignment*

Time synchronization is crucial for the interpolation of asynchronously captured body tracking frames generated by multiple depth sensors. Since the user walks slowly during the registration process and Kinect v2 only captures skeletal data at 30Hz, a worst case offset of several centimeters (3.33 cm @ 1 m/s speed) is induced just by event-based, non-synchronized image acquisition. To generate synchronized timestamps within the whole sensor network, NTP protocol is utilized. Based on these precise timestamps, skeletal body tracking frames are virtually synchronized within the fusion software through offline time interpolation. The depth sensor's skeleton acquisition time is

assumed to be constant for all sensors. Since the user's body has a certain inertia and the refresh rate varies at around 30Hz, the inter-frame trajectory between two skeleton datagrams can be assumed as linear movement.

7.6.2.5 *Fusion process and quality heuristics*

Having registered all sensors via extrinsic transformation chains into a common coordinate frame, the tracked body frames - generated from different views - are placed in the same coordinate frame and must be fused. For large-frustum and rotation invariant motion capture for posture analysis, the following set of skeletal fusion heuristics is proposed. Each skeleton within each sensor is given a certain weight. The higher the weight the higher the influence of a specific sensor on the tracked subject's fused skeleton. A set of weight penalties is presented in the following for real time skeletal fusion. For each sensor and skeleton, these weight penalties are calculated as follows:

First, a **distance measure** penalizes unfavorable distances between the user and a sensor. This quality measure weights the sensor's skeleton over the distance to the neck joint respectively. Evaluations have shown that the most reliable tracking results can be achieved when the user is tracked at a distance of two and three meters between himself and the sensor. Other distances are penalized.

$$
w(d) = \begin{cases} 0 & \text{for } d \leqslant 1m \\ 1 - (d - 2.0m) & \text{for } 1m < d \leqslant 2m \\ 1 & \text{for } 2m < d \leqslant 3m \\ 1 - (d - 3.0m) & \text{for } 3m < d \leqslant 4m \\ 0 & \text{for } 4m < d \end{cases} \tag{7.2}
$$

Second, the **rotation quality measure** penalizes when a user is standing in a less frontal position to a sensor. This is helpful for rotation invariant human activity analysis. Full frontally captured skeletons get high weights. Rear views are set to zero weight. The user should stand as orthogonally to the sensor as possible and 30° has been determined as the maximum vertical user orientation for reliably tracking limbs:

$$
w(\varphi) = \begin{cases} 1 - \frac{|\varphi|}{30°} & \text{for } |\varphi| < 45°, \\ 0 & \text{for } |\varphi| \geqslant 45° \end{cases} \tag{7.3}
$$

Kinect V2 has a horizontal field of view of 70° in the ToF sensor. In order to achieve a smooth transition between the multiple tracking frustums at the horizontal edges to the tracking frustum, a **lateral frustum quality measure** limits the tracking frustum to a horizontal field of view of 50° so that the limbs are still likely to be within the

tracking area of the sensor (70°). Weights are set to zero if the user's center axis joints exceed 50° on the horizontal axis of the local camera coordinate frame:

$$w(\alpha) = \begin{cases} 1 - \frac{|\alpha|}{25°} \text{ for } |\alpha| < 25° \\ 0 \text{ for } |\alpha| \geqslant 45° \end{cases} \tag{7.4}$$

The weights are fused to a total penalty weight for the sensor's skeleton with weights K_{Dist} (e.g. 1) and K_{Rot} (e.g. 10). Different weight profiles can be used to optimize the fused result for special circumstances, e.g. laboratory settings without any objects occluding the user or more complex settings such as automotive PV workshops where physical mock-ups placed in the middle of the workshop environment may weaken the tracking quality of some sensors:

$$w_{Tot}(d, \alpha, \varphi) = w(\alpha)\left(K_{Dist}w(d) + K_{Rot}w(\varphi)\right) \tag{7.5}$$

These quality heuristics are used to subsequently fuse the skeletons geometrically in a weighted manner. Skeletons are fused in a weighted manner for all joints:

$$\overrightarrow{p}^i_{Fusion} = \frac{\sum_{j=0}^{k} w^j_{Total}(d, \alpha, \varphi)\overrightarrow{p}^{i,j}}{\sum_{j=0}^{k} w^j_{Total}(d, \alpha, \varphi)} \tag{7.6}$$

Practical implementation has shown that for six sensor setups, typically only the best three weighted skeletons can be used for weighted overall skeletal fusion. Secondly an implementation of a linear Kalman Filter was used [221]. The gathered weights from the heuristics are integrated in the measurement covariance matrix of the filter, increasing or decreasing the variance of the measurement. Since the tracking data of the Kinect sensors are always afflicted with jitter, the Kalman Filter can improve the overall tracking quality by its probabilistic modeling, and in case of partial or full loss of tracking data it can predict the skeleton state.

7.6.2.6 *User identification*

Having fused skeletal data between many sensors, skeletons are handed over between registered sensors. If users move in the extended tracking frustum, the same ID is used every time and theoutput skeleton does not jump to other tracked users. Even though Kinect SDK already provides skeletal IDs, a persistent ID of fused skeletons is generated by using dynamic lookup table with minimal Euclidean distance for the neck joint. As long as the user stays in the multi-sensor tracking system, the persistent ID does not change for smooth transmission of skeletal data via standardized protocols.

7.6.2.7 *Standard tracking protocols*

Three different standardized protocols are implemented in the multi-kinect system for interoperability. The fusion service layer can be used fully transparently, so that the receiving software component cannot distinguish whether the single tracking web service or fused skeletons are transmitted. Even though frequencies of data acquisition in the fusion service may vary, the transmission of tracking data is carried out at an adjustable but fixed frame rate. For both variants the following protocols are implemented:

Virtual-Reality Peripheral Network [143] is an open source and broadly used protocol. It is implemented so that data can be transmitted to every compatible receiver.

dTrack is a proprietary protocol of ART company [183]. Nevertheless, because the unidirectional protocol uses unencrypted and uncompressed messages, it has come be widely used in the tracking community. That is why this protocol received a pseudo-standard for transmission of tracking data. Many simulation environments natively support AR-Tracking components. Therefore, implementing this protocol opens the doors to use with standard assessment environments.

ARVIDA RDF protocol and vocabulary was developed in the publicly funded "ARVIDA" research project (see Keppmann et al. [200]) This novel standard for polling spatial data using linked data RDFs is supported by the multi-kinect system.

7.6.2.8 *3D Point Cloud Fusion*

Besides skeletal tracking data, the tracking service also allows polling of the latest reconstruction of the 3D scene point clouds. Due to large volumes of 3D point cloud data, the tracking service uses protobuf protocol to compress these 3D point clouds in real-time. Even when on the fly compression algorithms are used, sub-sampling the point clouds 1:5 is required in order to transmit data from six sensors in a 1 Gbit/s local area network.

Fusion service has determined an extrinsic transformation matrix between the depth cameras of each sensor. Using this common coordinate frame between the master sensor and the slave sensors can be utilized to generate a merged, live real time point cloud consisting of depth information of multiple sensors. Figure 7.13 illustrates the fusion of two point clouds in real time. Red and black dots indicate information from each camera.

Background substraction of static contents is implemented easily, since silhouette points of the tracked human are given by Kinect SDK. For instance, Shapiro et al. uses such rapid avatar captures for interative simulations [222] Li et al. for Kinect 3D self-portraits [223] and Zhao reconstructs human body shapes from commodity depth

Figure 7.13: Point cloud fusion with two Kinect v2 cameras. Human depth information becomes more dense and concealed structures can be captured using multi-sensor arrangements

cameras [224]. This kind of self-representation may create a sense of embodiment in the VE by giving the subject the sense of self-location, sense of agency and sense of body ownership (compare Kilteni et al. [225]).

7.7 EVALUATION OF REGISTRATION ACCURACY

To determine the accuracy and precision of extrinsic transformations and thus the spatial registration error using the fusion service, a series of experiments is carried out.

7.7.1 *Experimental setup*

Since an absolute accuracy evaluation is needed, a high precision marker-based tracking system was chosen as a ground truth measurement tool. This OptiTrack system consists of 16 OptiTrack Flex 13′ cameras and uses retro reflective markers trees to measure a 6 DoF coordinate system in the tracking frustum. The OptiTrack system reports a residual mean error of 0.624mm at its initial calibration for the entire tracking volume.

Both Kinect v2 sensors are equipped with a rigid body trackable (tree of retroreflective balls) so that the position and orientation of each sensor can be located precisely in the OptiTrack system (see Figure 7.14). The pivot point translation of the rigid body markers are set to the Kinect's depth camera focal point to match the origins of Kinect body tracking and the OptiTrack rigid body markers.
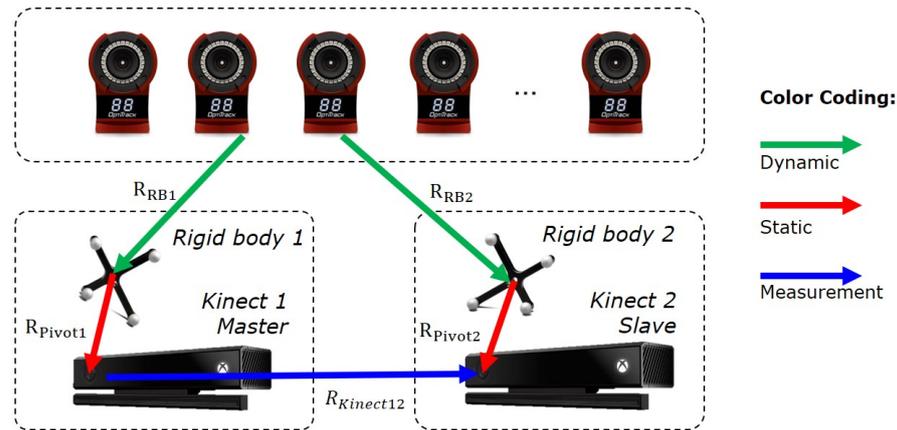
Figure 7.14: Experimental setup for evaluation of Multi-Kinect tracking accuracy

7.7.2 *Design of experiment*

All registration scenarios are conducted using two Kinect v2 sensors (see Figure 7.14). The above described fusion service gathers the tracking data of both tracking services, each running on a standard PC overfulfilling the minimum requirements mentioned by Kinect SDK (Core i7, 8GB RAM, GTX 970 GPU).

Five scenarios are carried out. These scenarios differ in the sensors' relative positions and their angles around the vertical axis: 0°, 45°, 90°, 135°and 180°. For each scenario, 20 runs are performed resulting in a total of 100 measurements.

Each run follows the same procedure. The operator tells the user to move inside the overlapping tracking frustum for 10 seconds. During this time, skeletal tracking data is gathered for extrinsic markerless registration. Registration automatically ends when there are enough data points for registration. No outliers are removed for the following evaluation.

7.7.3 *Results*

Figure 7.15 illustrates the registration performance of the fusion service. Circles depict the calculated ideal OptiTrack positions.

For these scenarios the Euclidean distance in the floor plane is always less than 15mm to the ground-truth position. The vertical axis reveals maximum deviations of 1.5° for the sensor's pitch axis. The body tracking estimator within the SDK reveals uncertainties especially in the vertical axis. The uncertainty of the joints can vary up to 20mm, depending on the angle between the user and the sensor.

The reproducibility of this approach is summarized in Table 7.2. The mean deviation from the center point ranges from 9.6mm to
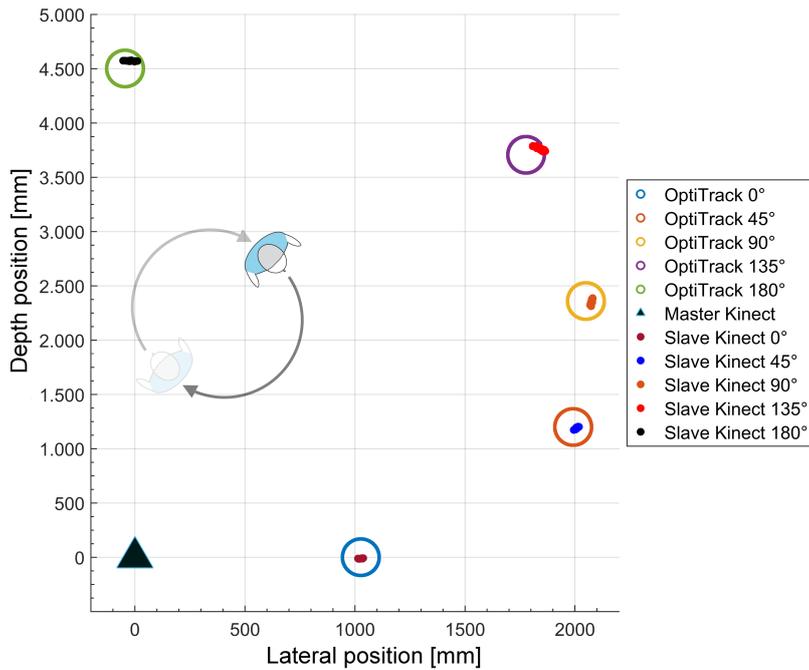
Figure 7.15: Top view of the registration results: Master sensor at [0,0], 5 scenarios with 20 registrations each, circles indicate the ground truth of the OptiTrack measurements

Table 7.2: Reproducibility and deviation of registration results for the 5 scenarios (N=100)

Scenario	Mean Error	Standard Deviation (SD)
0°	9.6mm	4.5mm
45°	16.0mm	7.1mm
90°	17.8mm	13.6mm
135°	26.6mm	10.0mm
180°	12.3mm	10.0mm

26.6mm with a standard deviation ranging from 4.5mm to 13.6mm. The maximum deviation was 42.2 mm in the 135° scenario.

7.8 PRACTICAL MOTION CAPTURE IN PV WORKSHOPS

During PV workshops in the VMS typically six Kinect sensors are utilized. They all face the center of the assembly cell and are evenly distributed on the edges of the tracking area. This area covers at least 6m x 6m since movements within real workstations in automotive end-assembly lines match these dimensions. For less challenging tasks, two Kinect sensors with a parallel arrangement and a distance of 1.5 m proved to work reliably.
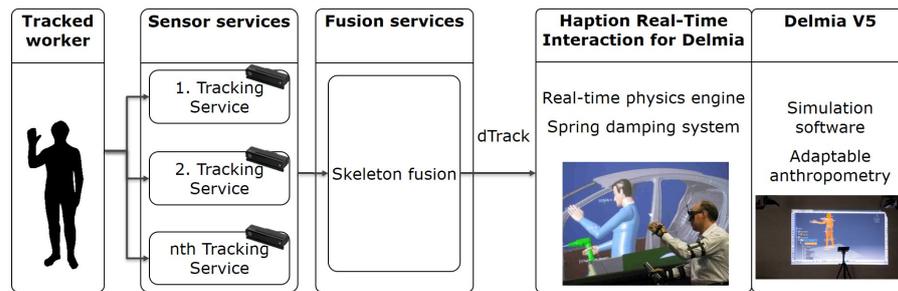
Figure 7.16: Block diagram of data pipeline for digital human simulation in Delmia V5

Having registered all cameras to a common world coordinate frame, the system architecture presented in combination with the fusion heuristics enable constant tracking of the worker regardless of his position and his orientation within the concatenated tracking frustum.

Besides the presented "veo" tracking pipeline (see Section 6.3.5), another working simulation pipeline for ergonomic assessments in the VMS includes the usage of "Delmia V5-6 R23" in combination with "Haption RTID" plugin, as shown in Figure 7.16. The "fusion service" exposes all fused skeletons as 6 DoF tracking data via the A.R.Tracking dTrack protocol. Haption "Real-Time Interaction for Delmia" receives skeleton datagrams in the integrated real-time physics engine and maps them onto a fully flexible virtual human in Delmia V5. In this case, only 20 out of 25 skeletal joints are used to manipulate the DHM interactively. RTID uses a spring-damper retargeting algorithm with dynamic weighting of skeletal joint influences. This modification is required as marker-based tracking systems deliver reliable orientation data in contrast to Kinect skeletal tracking. Calibration between virtual and physical human is carried out by executing a T-Pose in physical space, offseting between tracking data and the virtual humans is minimized.

Using this pipeline, the DHM matches the size of the physically tracked user. Even though theoretically the DHM should be the same size as the real user, practice has shown that the tracking results are more stable when the DHM is 5 cm smaller than the physical user.

Figure 7.17 depicts a Delmia V5 VE having a virtual car body in the assembly status of the respective workstation. Dynamic parts are simulated and attached to the right hand joint. The anthropometry of the virtual human is adjusted to the real worker's size and weight. Limitations of the pre-planned process and unfavorable ergonomic situations can be identified using this virtual assessment pipeline. Additionally, the results gathered could be verified by subsequent traditional hardware workshops.
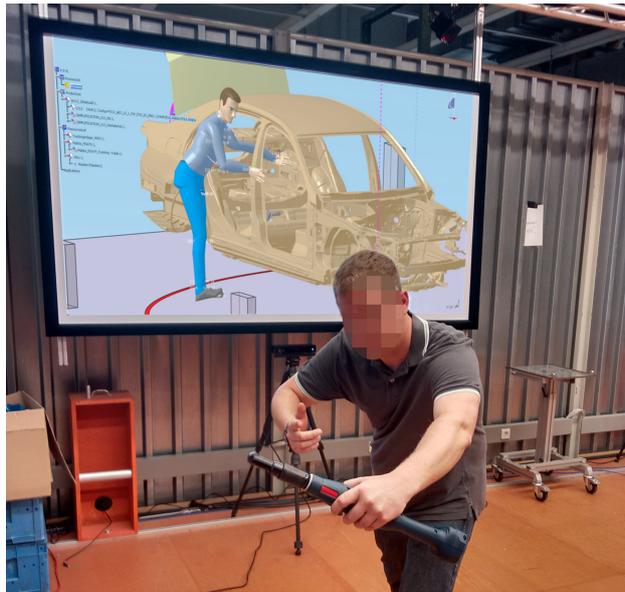
Figure 7.17: Delmia V5 DHM directly manipulated by marker-less motion capture approach. All assessments can be carried out in real time

7.9 WORKING POSTURES EVALUATION IN ERGONOMIC ASSESSMENTS

In this section, the presented multi-depth sensor motion capture system is evaluated with respect to its applicability for standardized working postures and its ergonomic assessments. The applicability for real PV workshops is seen as given when all relevant working postures can be reliably carried out with the presented system. Even though multiple publications propose Kinect v2 as suitable for use in ergonomic assessments, none of them have evaluated it with respect to specific working postures (see Haggag et al. [226], Geiselhart et al. [174], Bortolini et al. [227], etc.).

For both single- and multi-sensor arrangements, closed-source Kinect SDK by Microsoft is applied for skeletal tracking, based on the principles presented by Shotton et al. [185]. Even though fusion results optimize the overall tracking results, they rely on the single skeletal tracking results of each sensor. Therefore the following results are valid for both single Kinect and multi Kinect approaches.

7.9.1 *Related work on standardized ergonomic assessments*

To avoid musculoskeletal complaints and disorders of workers, manufacturing companies carry out ergonomic risk assessments. Poor ergonomic design of workstations can be reliably detected using DHMs and even be optimized [167], whereas deriving repetitive forces in the virtual domain is still hard to assess. Typical ergonomic assessment

methods for tackling workstation assessments are screening tools for physical workload.

In the automotive industry, plenty of ergonomic screening methods are applied, such as the rapid upper limbs assessment (RULA) [228], rapid entire body assessment (REBA) [229], NIOSH [230], OCRA [231] and more regional ones, such as NPW, Design Check [232] and AAWS. Besides those screening methods, EAWS by Fraunhofer IAD [67] is primarily used by European automakers since the aforementioned ergonomic assessment methods are either predecessors or compatible with EAWS. International standards for minimum ergonomic requirements are presented in ISO 11226 for postures and ISO 11228 for actions [67]. These screening tools are often integrated in assembly assessment simulation systems [233], [170]. At present, synthesized motions of DHMs often lack accuracy and parameterization capabilities which leads to vague assessment results and time-consuming authoring [123]. Therefore pen & paper-based assessment methods are still state of the art.

EAWS is widely used by European car manufacturers and automotive suppliers [227], [234]. It penalizes unfavorable physical workloads with "load points" and deduces an overall risk assessment. " The EAWS consists of four sections for the evaluation of working postures and movements with low additional physical efforts (< 30-40 N or 3-4 kg respectively), action forces of the whole body or hand finger system, manual materials handling and repetitive loads of the upper limbs." [67]. Working postures are assessed as "static working postures and high frequent movements are estimated" [67]. "Symmetric working postures for standing, sitting, kneeling & crouching and lying & climbing are rated" as well as "asymmetric effects like rotation, lateral bending, and far reach" [67].

7.9.2 *Study goal, setup and evaluation method*

An ergonomics expert using a virtual environment with an animated DHM must be able to reliably assess the overall process and come to the same conclusions as by assessing the physical domain. The following evaluation aims at answering the question whether Kinect v2 as a standalone system as well as the multi-sensor system presented by Otto et al. [172] are able to deliver assessable results for EAWS working posture assessments. Similar to the following applicability analysis, Haggag et al. evaluated Kinect v1 for rapid upper limb assessment (RULA) using an automated assessment approach in 2013 [226].

EAWS working postures were evaluated below with regard to whether they can be carried out by using the presented markerless motion capture system. The intended goal was achieved when the ergonomic
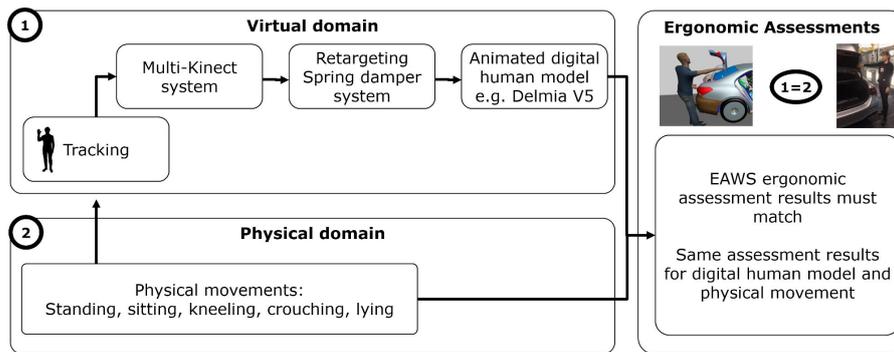
Figure 7.18: Block diagram of the EAWS applicability evaluation pipeline

expert reached the same assessment results by visually inspecting all working postures of the animated DHM in the simulation scene.

Figure 7.18 depicts the system block diagram and pipeline. While having been tracked, a worker performed all EAWS relevant working postures in different symmetric and asymmetric postures. The "Multi-Kinect" tracking system consisted of six sensors and gathered real-time motion capture data. Latency was negligible for the applied motions. All six sensors were evenly distributed and have been facing towards the middle of the tracking frustum. There were no optical occlusions in this laboratory setup.

The generated data stream of skeletal tracking information has been transmitted via the standardized dTrack protocol and then retargeted onto a DHM by using the IPSI spring damper system made by Haption. The DHM moved within an simulation scene only consisting of a floor plane and the DHM avatar representation.

7.9.3   *Procedure and participants*

For evaluation, 3 participants carried out all EAWS working postures (i.e. standing, sitting, kneeling & crouching and lying) as well as two non-standardized dynamic postures as the "tracked workers" (see Figure 7.18). Their motion has been captured and recorded using the multi-depth sensor tracking system and an RGB camera for the physical domain. Subsequently, an ergonomics expert - working for an automotive OEM - visually assessed all recorded datasets and tried to fill out the EAWSheet. Both the virtual and physical domain were displayed side-by-side so that the expert could compare both domains. All participants took part on a voluntary basis. The results indicate whether the expert would come to the same conclusion by assessing only the virtual domain.

7.9.4 *Results of working posture analysis*

The results for **standing** working posture levels are presented in Table 7.3.

Table 7.3: Evaluation results for EAWS working postures "standing"

Body posture	EAWS	Resulting Posture	Lower body	Upper body	Overall	Remarks
Standing Upright (0° - 20°)			✓	✓	✓	
Bent forward (20° - 60°)			✓	✓	✓	
Strongly bent forward (> 60°)			⚠	✓	⚠	Strong optical occlusion of lower body limbs due to missing line of sight
Arms at / above shoulder level			✓	✓	✓	

In general, the working postures "standing upright" and "little bending forward" are applicable for EAWS assessment without any further limitations. Those working postures can be assessed properly by the ergonomics expert. The working posture "strongly bent forward" causes optical occlusions of the legs due to the missing line of sight to the Kinect v2 sensor. Fusion heuristics of multi-sensor systems do not improve the "strongly bent forward" results of a single sensor. However, the overall body posture can be used even though the legs are getting jittery. Moreover, the "arms above shoulder" posture as well as combinations with symmetric effects allow a feasible EAWS assessment for all standing postures. In particular, rotation, lateral bending, and far reach are possible. No visible limitations apply for the space within reach.

Analyzing the gathered results for **sitting** with related body postures, it can be denoted that EAWS "sitting" working postures can be assessed reliably as long as one of the sensors is placed frontal (+/- 30°) of the tracked user. All EAWS "sitting" postures with symmetric and asymmetric combinations are reliably detected and can be used for ergonomic assessments (see Table 7.4)

EAWS working postures **"Kneeling & Crouching"** capture upper body movements reliably whereas lower body parts are error-prone. With the "kneeling" posture, the lower legs are completely occluded for the sensors which results in the lower legs are penetrating the floor plane. In this case, the motion capture data delivers outstretched legs. Despite this tracking inaccuracy, the knees, upper legs and upper body are at the correct height, so that an overall interpretation is still possible. Same holds for "crouching" postures. If the leg can

Table 7.4: Evaluation results for EAWS working postures "sitting"

Body posture	EAWS	Resulting Posture	Lower body	Upper body	Overall
Seated Upright (0-20°)			✓	✓	✓
Seated strongly bent (>20°)			✓	✓	✓
Seated Overhead work			✓	✓	✓

be seen by a frontal sensor, crouching is working properly. Symmetric and asymmetric upper body combinations are also feasible (see Table 7.5)

Table 7.5: Evaluation results for EAWS working postures "Kneeling & crouching"

Body Posture	EAWS	Resulting Posture	Lower body	Upper body	Overall	Remarks
Kneeling & Crouching Upright (0 - 20°)			⚠	✓	✓	"Crouching" works "Kneeling" does not work properly. The occluded lower legs are tracked as outstretched posture. Nevertheless, the overall body height matches the real scenario, so that the virtual knee touches the virtual floor as can be seen in the pictures on the left.
Kneeling & Crouching Strongly Bent (0 - 20°)			⚠	✓	✓	
Kneeling & Crouching Overhead Work			⚠	✓	✓	

In contrast to the aforementioned working posture types, EAWS **"lying"** postures do not work properly with the presented markerless motion capture system. Each sensor in the array was placed at a height of 1.3m with a horizontal view. Due to the steep viewing angle, lying completely flat on the floor generates only jittery skeletal tracking data which cannot be applied for ergonomic assessments. If the upper body bends up > 20° towards a sensor, tracking data can be used again for EAWS (see Table 7.6).

Assessing the multi-sensor system's performance for **dynamic postures**, two non-standardized dynamic posture evaluations are carried out. In a 360° rotation experiment, the tracked user constantly turns around a pole so that skeletal hand-over between sensors can be evaluated. Tracking data is handed over properly between all six sensors,

Table 7.6: Evaluation results for EAWS working postures "lying"

Body Posture	EAWS	Resulting Posture	Lower body	Upper body	Overall	Remarks
Lying Flat			✖	✖	✖	Data not usable for ergonomic assessments. Too jittery.
Lying >20° upper body (sit-up)			✔	✔	✔	Applicable MoCap results are generated for sit-ups >20°.

even the hand sticks on the virtual pole and the movement is properly mapped on the DHM. Second, the entire space within reach is evaluated. No additional limitations compared to physical restrictions have been detected using skeletal tracking (see Table 7.7).

Table 7.7: Evaluation results for working postures "dynamic postures." 360° rotation only applies for multi-depth camera setups

Dynamic movements	Resulting Posture	Lower body	Upper body	Overall	Remarks
360° Rotation		✔	✔	✔	Movement quality is depending on sensor density. For six sensor setup working properly.
Space within reach		✔	✔	✔	No additional limitations induced by tracking technology could be revealed - compared to anthropometric space of reach possibilities.

7.9.5 Discussion

All in all, the Kinect v2 skeletal tracker extended with the multi-depth camera tracking algorithms and the retargeting system proved to be suitable for use in ergonomic assessments in accordance with EAWS. Overall, 9 out of 11 full body postures can be used for ergonomic assessments. Optical occlusions cause jittery and unusable motion capture data while strongly bending forward, lying flat on the floor and kneeling on the floor. These optical occlusions become more intense when having large-scale PMUs inside the VMS as they potentially occlude tracked users.

7.10 FLEXIBLE, VISUALLY LOW-OCCLUDING PMUS

As determined in the ergonomics study described above, the overall results are heavily influenced by sensor arrangement and amount of sensors. As ToF sensors are optical tracking devices they require a direct line of sight for robust skeletal tracking. Despite the developed fusion heuristics, at least one sensor needs a good, frontal view (distance > 1m, low vertical rotation) of the tracked subject. A proper arrangement of depth-sensors minimizes the optical occlusions for tracked users and optimizes the views for frontal user captures.

PMUs are used by the tracked users to generate a sensation of **physical barriers** for kneeling, sitting, bending over. For instance, the user has to bend into the trunk to mount a control unit or has to sit in the empty interior of the PMU in order to mount car ceiling parts. As depicted in Figure 7.1, in PV workshops PMU are regularly integrated in the assessments, for instance a full car chassis. However, the use of PMUs in combination with skeletal tracking may result in many occlusions in the VMS.

In order to maintain a good markerless tracking performance, in the following **abstract, flexible, optically low intrusive PMUs** are proposed: These PMUs are constructed so as to minimize optical occlusions. They can be either built from wooden-based, cardboard-based, carbon fiber-based or aluminum-based materials. Four types of dummies are proposed. Each of them is constructed with a diameter of 2 cm to have as little optical occlusion as possible and made out of aluminum profiles to minimize the influence on the skeletal tracking result. Nevertheless, these profiles are strong enough to provide haptic support.

The requirements of these four flexible PMUs have been deduced from ergonomic posture assessments (Ergonomic Assessment Worksheet (EAWS)), practical motion capture sessions and constraint affordances. The flexible mock-ups are presented in Table 7.8. For ergonomic assessments, typical movements are sitting, bending over, lying, leaning and getting lateral support and holding on to sth. (compare with EAWS [67])

All presented flexible PMUs can be tracked with markers so that the physical mock-up is registered in relation to the DHM in the virtual scene. This facilitates rearrangement of the PMUs and of course orientation of the user in the virtual scene. For example, different work heights can be validated simply by changing the height of the tracked horizontal bar of the PMU. The virtually represented object itself, such as the trunk door, changes its height accordingly.

Table 7.8: Visual light-weight physical mock-ups for occlusion reduction

Purpose	Example	Use case description
Sitting		Applied for seated work processes with adjustable sitting height. Can be combined with other PMUs. Example: Interior overhead assembly
Lean against sth./ hold on sth.		Applied when vertical support is needed to perform a manual assembly task. Users can also support themselves by holding on to the vertical aluminum profile. Example: Mounting a control unit beneath the steering wheel while getting support from the A-pillar.
Bend / bow over		Applied for bending over at a variable height or kneeling. Example: Mounting control units in a trunk.
Haptic constraints		Applied for realization of a physical constraint or physical orientation, such as a height constraint, which should not be exceeded. Example: Mounting cover panel of an opened trunk door at a certain height.
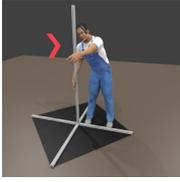
7.11 CONCLUSION

A system for large-scale, markerless motion capture using multiple depth-sensors has been presented and evaluated. This system has been developed in order to acquire rotation invariant tracking results, large scale tracking frustums and optimized motion capture data for the VMS's simulation environments.

An in-depth analysis of skeletal tracking performance of Microsoft Kinect v2 sensor revealed insights into how this tracking data can be used optimally for a multi depth camera setup. For evaluation purposes, a "UR 10 Universal Robot" was used in combination with a mannequin to generate reproducible trajectories and constant velocity. Best results were achieved by processing the neck joint information. Interestingly, the head showed less inter-frame jitter compared to the remaining joints of the tracked skeleton. Additionally, orientation limitations and front/rear ambiguities were discussed in detail.
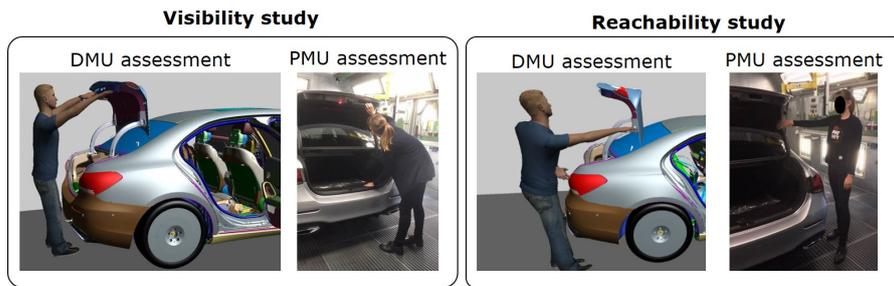
Figure 7.19: Practical usage of marker-less motion capture in PV workshops in comparison with PMU assessments

For the extrinsic registration of several depth sensors and fusion of sensor data, several registration-relevant techniques were presented and evaluated such as time-synchronous interpolation, front/rear detection and error measures. In addition, a comprehensive set of quality heuristics were derived for the skeletal fusion process which showed an improvement in skeletal tracking. These heuristics profit from the in-depth skeletal tracking performance analysis. Technically the number of possible tracking nodes are only limited by computing power and network throughput. Sensor setups of up to ten tracking services were successfully tested. Practically, for VMS workshops with 360° rotations of the tracked user, a sensor array of six cameras proved to be working reliably. The fusion service itself can be used transparently and acts externally as if it were a single sensor tracking service. Standardized tracking protocols (VRPN, dTrack, trackD, ARVIDA) are implemented to achieve interoperability with virtual assessment environments and game engines.

The ergonomics of hundreds of workstations have been assessed using the above presented system in PV workshops at Mercedes-Benz Cars. Figure 7.19 depicts a comparison between the PMU and DMU assessment for two real use cases of a PV workshop. Both use cases can be carried out in the physical and virtual domain. Only slight differences apply: In the visibility study, DMU assessment seems to be slightly more feasible, whereas in PMU assessments, large bending is required to see the clip's location. In contrast, a conducted physical reachability study has found that the intended working location can be reached more easily than indicated by the DMU assessment due to choosing a different stance relatively to the car chassis.

Ergonomic experts point out multiple positive effects compared to marker-based systems: Users do not have to put on a special suit with retro-reflective markers. This is time-consuming and cumbersome for the tracked persons. Movements may be influenced by the marker suits and seem not as natural as in regular working clothes. In addition, users can swap immediately without any preparation time, so that multiple users can test the process without any prior work.

Furthermore, experts appreciate the side benefits of this tracking approach such as visibility checks through interactive viewpoint control.

Nevertheless, on the down side the markerless system induces more latency and jitter to the tracking data than the marker-based tracking system. Ergonomic experts point out that the motion capture data quality is still sufficient for most work tasks to identify and solve the issues. Latency of several frames is considered to be irrelevant. It becomes apparent that registration as well as fusion accuracy and precision are sufficient for human posture analysis, profound ergonomic simulations and for large-scale view point control applications in virtual environments.

7.11.1 *Limitations*

Overall, by using the system some limitations have been revealed, which could not be resolved even by fusion of multiple sensors:

Dark to **black clothes** of tracked subjects have problems being tracked reliably since reflection parameters for the ToF sensors deliver jittery data.

High precision tracking data (<1cm) cannot be delivered by Kinect SDK. Even when fusing these data, precision does not reach sub millimeter precision. This impedes the use for buildability product assessment with narrow clearances. A marker-based optical tracking system can be fused for end effector tracking, such as virtual hand representations or tooling such as screw drivers.

Finger tracking is not possible by using standard SDK tools. For PV workshops, eight standard hand poses can be either statically set or a fusion with an external leap motion finger tracking sensor is used in order to manipulate fingers as well (see Qian et al. for hand tracking using depth sensors [195])

Since Kinect SDK does not calculate the **orientation of head joint** it is approximated using the Shoulder_Left and Shoulder_Right joints. For visibility assessments, reconstructing a virtual view cone is possible but sometimes not as accurate as required.

Usage of optically heavy PMUs is generally feasible. To avoid jitter, multiple ways are proposed, namely (1) adjusting the position and orientation of the sensors in PV workshops, (2) adjusting the amount of sensors in the arrangement or (3) using optically low intrusive PMUs. The latter is presented as a method for achieving haptic feedback in the physical domain, while maintaining good tracking results in the virtual domain.

7.11.2 *Fulfilled requirements*

In accordance with the VMS key properties (see Section 5.2), Table 7.9 shows a realization analysis of whether the developed multi-depth camera tracking system meets the VMS's tracking requirements:

The presented system is capable of **practically supporting many PV verification criteria**. For example, the motion capture system can be used as a tracking device for simulation of dynamic part assembly and disassembly paths when attaching a tracker to the physical object. Reachability and overhead assembly simulation works well (see Figure 7.19). Visibility assessments of mounting points work but they are limited by bending the neck in a horizontal manner. Body forces, hand/finger forces cannot be simulated using this motion capture system at this time, even though the literature has proposed force estimation using MoCap data. Motion capture data are not used for time measurement of physical processes due to legal aspects. All operational resources, accessibility for tooling, intra-logistics and designing cell layouts can be used in combination with the proposed motion capture system.

A requirement analysis showed that for **ergonomics-related assessments**, such as EAWS body posture analysis, the presented system can be applied properly. Analogously, Haggag et al. conclude that Kinect's "motion capture capability makes it suitable for detecting the risk of injury and recording posture in the workplace" [226]. No force analyses are feasible and orientation tracking is limited as well. Even though no joint forces and torque can be measured currently, Planard et al. propose a system to use skeletal Kinect data to estimate these measures using an inverse dynamics method [235]. With this method, they find that forces and torques can be determined reliably for shoulder joints even in cluttered environments.

Overall, the multi-depth sensor system has proved its applicability and therefore reduces preparation times compared to marker-based systems, reduces costs and increases ease-of-use due to fast tracked subject change. It is shown that performance and applicability of the system is suitable for use in manufacturing industry and can be seen as a complementary system to conventional high-end marker-based systems in this domain.

Table 7.9: Realization analysis of requirements of the presented motion capture system

Requirement	Status	*Description*
Room-scale tracking frustum	*Works*	*Infinite scalability, limited by network transmission and calculation performance. Up to ten sensors tested in arrangement*
Multiple tracked users	*Works*	*For each sensor, up to six people can be tracked. Maximum 2 people required for PV workshops. Fused skeletal data get a unique ID for data transmission.*
Instantaneous tracking	*Works*	*Shotton et al. algorithms [184] in Kinect SDK enable instantaneous full body tracking*
Low-Cost Tracking Components	*Given*	*Kinect v2 sensors are commercial off-the-shelf products. Solely computing hardware for fusion service requires good computing performance*
Interactive frame rate and latency	*Works*	*Adjustable, fixed frame rate up to 120Hz for data transmission from fusion service. Optimized in fusion service leveraging the effect of multiple unsynchonized 30Hz sensors and time-interpolation.*
Level of DHM articulation	*Partly fulfilled*	*25 joint full body posture tracking is implemented. Level of articulation in fingers and hands are to be optimized. Facial expression tracking was not required and is not implemented*
Accuracy and precision	*Partly fulfilled*	*Precision of tracking cannot be regarded as high precision, rather an first approximation (>1cm). Not all PV workshop assessments can be carried out (e.g. product assessments) but most verification tasks can be carried out*
Usage of standardized protocols	*Fulfilled*	*VRPN, dTrack, ARVIDA protocols are implemented for both single and fused skeletal data.*

WALL-SIZED, HIGH RESOLUTION VISUALIZATION HARDWARE SYSTEMS

Large High-Resolution Displays (LHRD)s provide an enabling technology to achieve immersive, isometrically registered virtual environments. Wall-sized visualization hardware is becoming more and more common in research, entertainment, public signage - and of course industry. The reasons for this are that technical specifications are constantly increasing while prices are simultaneously decreasing.

This trend especially applies for the two most common types of wall-sized visualization hardware: LED walls and projector systems. Both LHRD technologies are able to display large amounts of information such as alphanumeric data, 3D CAD files and immersive experiences. LHRDs can also be found in increasing numbers of public signage installations as shared interaction spaces for content which is presented in an interesting manner, new interaction forms and collaboration (see Peltonen et al. [236]). Even more, immersive display technologies are expected to allow more accurate size judgments, improved collaboration performance in workshops and reduced task completion time [237]. However, according to Andrews et al. [238], the advantages of additional pixels — generated by high pixel densities and large installations — can only be realized "through understanding of the interaction between visualization design, perception, interaction techniques, and the display technology" [238].

The realization of such an output device follows the key properties of the VMS framework (see Section 5.2) and consequently aims to improve PV workshops. These requirements have been deduced in multiple, iterative workshops with potential customers:

- **True to scale output:** An entire manufacturing workplace in automotive general assembly requires a true-to-scale visualization so that the whole product and work place can be seen at once. According to VMS requirements, a floor visualization system is defined to the minimum dimensions of 6m x 3m since that is the size of a typical work place.

- **Co-location:** Workshop participants can collaborate in the same VE at the same time without being remotely connected.

- **Spatial integration of PMU and DMU:** As a consequence of true-to-scale output, physical dimensions of tracked objects, parts and humans can be adjusted to match the virtual dimensions. This allows for augmentation of physical parts with virtual contents. No teleportation of single virtual components is intended

to be realized (in contrast to co-located teleportation approaches such as those presented by Langbehn et al. [107])

- **Multi User Support:** The output devices must meet workshop requirements for collaborative use. Collaboration performance is a key performance indicator for properly implemented VMS concepts.

- **Asymmetric output for VR/AR:** In contrast to symmetric collaborative VR/AR concepts, where all workshop participants perceive individual views on the VE using the same visualization device, an asymmetric visualization approach is proposed (compare with Gugenheimer et al. [109]). Large workshop groups share a common view of the virtual scene using power walls, whereas selected users choose an individual view of the scene and interact with it simultaneously.

An iterative implementation and evaluation of two visualization hardware systems is presented below. First, a prototype of a floor projection system is described using scalable floor projection in order to prove technical feasibility and demonstrate application use cases of true-to-scale floor visualization systems. Followed by this, a multi-person-display ecosystem, LED-based implementation is presented. Both the prototype implementation and productive LED-based system share the same requirements in order to fulfill the expected use cases and both are realized in accordance with the VMS's key properties and take advantage of true-to-scale visualization (see Section 5.2).

Using LHRDs, multiple research questions arise: Amongst other benefits, pervasive displays are said to allow more accurate size judgments, improved collaboration performance in workshops and less task completion time [237]. In this chapter, task completion time is compared to state-of-the-art workshops and size judgement accuracy and precision is evaluated using a LED floor showing 2D content.

> **Related publications**
>
> Multiple parts have already been presented in peer-reviewed conference posters, book chapters, conferences papers and journal papers. In 2014, the conference poster "Using Scalable, Interactive Floor Projection for Production Planning Scenario" by Otto et al. (see [110]) was presented at the ACM conference "Interactive Tabletops and Surfaces" (ITS). Use cases and further algorithms were presented and described by the author together with Agethen et al. [239], [240]. The overall setup was described in the "Motion Capture" publication in the Springer book entitled "Web-basierte Anwendungen Virtueller Technologien" [132]. Analogously, the study on size perception using large scale floor visualization systems was presented at the ACM Pervasive Displays 2019 as a peer-reviewed conference paper [241]. In 2020, the paper "Using Large-Scale Augmented Floor Surfaces for Industrial Applications and Evaluation on Perceived Sizes" was published in the Springer journal "Personal and ubiquitous computing" [242]. Passages of the following texts have been extended and revised.

8.1 STATE OF THE ART IN LHRD VISUALIZATION SYSTEMS

Large high-resolution displays displays are used to display large amounts of data either for alphanumeric graphics (2D) or 3D data. In 1991, Mark Weiser published an article on "The Computer for the 21st century" [243] and proposed three different clusters of visualization devices. He clustered these devices in three groups, namely "tab", "pad" and "board"-sized devices. These "boards" were defined as yard-sized (>91 cm) displays and it was proposed that they should be used "in the home, video screens and bulletin boards; in the office, bulletin boards, white boards or flip charts." In 2009, Terrenghi et al. extended Weiser's size taxonomy of displays to "Inch", "Foot", Yard", "Perch" and "Chain", since all of them differ in their form of social interaction [244] within the multi-person-display ecosystems. Nevertheless, in 2015 Lischke et al. [245] concluded that digital "boards" are still rarely used. There was, however, a good chance that wall-sized display-"boards" would become commonplace within the next decade like smartphones and tables did in the last decade" [245].

Production validation scenarios in the automotive industry require interactive collaborative spaces and large display devices. Such combinations of multiple visualization devices are called **Multi-Display Environment (MDE)s**. Garcia-Sanjuan et al. present a general taxonomy of MDEs, classifying their topology with respect to "homogeneity of surfaces", "spatial form", "regularity of shape", "size", "mobility" and "scalability" [246] and also present many use cases in this liter-

ature review. According to Lischke et al. [245], software is the key enabler for a simple setup of multi-display environments and easy usability of these systems, which is still a hindrance factor for broad use. As parallelization of workflows becomes more and more important, LHRDs allow for visualizing complex data or switching between tasks without hiding required information at the same time. Rogers and Lindley analyzed collaboration in vertical and horizontal large displays [247] and found that the physical arrangement of publicly shared displays affects the social roles of collaboration members, such as switching roles more frequently, greater awareness of each other and exploration of more ideas. These findings by Rogers and Lindley, Lischke et al. and Garcia-Sanjuan et al. have directly influenced the automotive production planning application scenarios.

Additionally, industrial use cases aim for **high efficiency** when using LHRDs. In 2003, Czerwinski et al. analyzed performance using large scale displays in comparison with regular sized desktop screens [237]. They discovered that users' task completion times and productivity can be significantly increased for specific tasks by using larger visualization techniques. Analogously, in 2009 Bi and Balakrishnan supported these findings by using even larger LHRDs [248]. In their week-long survey they supervised users working with an LHRD and found that LHRDs enhance the user's awareness of peripheral applications, facilitate multi-window and rich information tasks and provide an immersive experience. "The results indicate that users unanimously prefer using a large display" [248]. Interestingly, Bi and Balakrishnan observed that "users tend to utilize the center part as the focal region and the remaining space as the peripheral region. The results also reveal that users on a large display perform more window moving and resizing but less minimizing and maximizing operations as compared to a single- or dual-monitor" [248].

8.1.1 *Content representation with LHRDs*

Having physically installed an LHRD system, users want to make the best use of its capabilities. Therefore, adapted **content representation and interaction** for this type of visualization devices is crucial. Andrews et al. present design considerations, outline challenges and future opportunities for designing visualizations on LHRDs [238]. They analyze physical display technologies, visual encoding, visualization designs and user interaction. Andrews et al. describe the benefit of additional pixels with the following, explicitly non-exhaustive list:

- More data entities

- Ability to show greater data dimensionality

- More data details

- Multiscale data

- More data complexity or heterogeneity

- Space for processes

- Space for sensemaking

- Enable collaboration with private and shared spaces

Complex data and large amounts of data are visualized on LHRDs. Lischke et al. conclude that using such systems enables humans to scan large areas quickly for objects and visual cues [249]. All of these aspects are necessary in PV workshops as they bring together PPR information. Many data entities are shown in parallel and represent multi-scale data and high complexity. For windowed applications, Lischke et al. explored the design space of LHRDs in 2017 by proposing four different graphical interfaces to be displayed in such arrangements [250]. They focus on windowed arrangements and admit that it continues to be a challenging task. They propose four new alignment techniques, namely "curved zooming, window grouping, window spinning and side pane navigation" and summarize their work by exploring the design space for focus switching but keeping spatial relations for related windows contents.

Andrews et al. examine how LHRDs support sense making and how the increased space affects the cognitively demanding task of **sense making** [251]. They explore the idea and show that this type of spatial environment supports "sense making by becoming part of the distributed cognitive process" and find "clear evidence of analysts using the space both as a form of rapid access external memory and as an added semantic layer providing both external memory and a semantic layer." This flexible semantic layer adds meaning to the displayed information such as ordering, proximity and alignment for clusters. This leads to a reduced need "for elaborate internal models by replacing memorization and computation with perception."

Observing the content representation and interaction with the graphical interfaces on LHRDs, Lischke et al. conclude that completely novel concepts are required to present content on LHRDs. Classic ways of arranging and hiding windows are neither suitable nor required anymore as visualization can be parallelized thanks to the by far higher resolutions.

8.1.2 *Interaction with LHRDs*

Lischke et al. also argue that "the success of wall-sized display installations highly depends on the **interaction technique** used in the particular setup" [245] and propose to get "a clear understanding of advantages and disadvantages of interaction techniques" used with

LHRDs. All these findings are required in manufacturing, as validation scenarios typically rely on commercially available software which is not natively built for LHRDs. Industrial applications want to leverage the benefits of showing multi-scale data entities and massive amounts of data in parallel, such as PPR information at the same time.

For LHRDs, multiple **input devices** can be utilized: Classical interfaces (e.g. mouse and keyboard), natural user interfaces such as direct or indirect touch (such as pointing or clicking by Vogel et al. [252] and Malik et al. [253]), 3D interaction devices (e.g. Microsoft Kinect DK) and multi-device strategies (body-attached interaction devices - smartphones [254], glasses, etc.).

8.1.3 *LHRD augmented floor surfaces*

In contrast to common vertical LHRD wall setups, **LHRD floor systems and their applications** have not been widely researched. LHRD floors are also called "augmented floor surfaces" and "floor visualization systems" in the literature. It is the only form factor where an infinite scalability for visualization and direct touch can be achieved due to its horizontal alignment.

Two **groundbreaking studies** must be mentioned in the context of augmented floor surfaces. In 1993, Cruz-Neira et al. presented the "Cave Automated Virtual Environment (CAVE) that also included a floor display for immersive environments. Cruz-Neira et al. used their "floor wall" projection system for the first presentation of a CAVE setup for virtual reality (VR) applications [255]. Another important study was conducted by Pinhanez in 2001. Their novel projection system utilizes a rotating mirror to augment all areas of a room including the floor surfaces [256]. These studies have inspired many further research activities.

In the literature, augmented floor surfaces are set up in various ways. The following, non-exhausive list clusters these **form factors**:

- **Back projection systems**
  e.g. MultiToe by Augsten et al. [257]

- **Front projection systems**
  e.g. laser projection-based system by Müller et al. [258] and "The Everywhere Displays Projector" by Pinhanez [256]

- **LED-based systems**
  e.g. by Dalton et al. [259]

- **Low-poly lumination systems**
  e.g. hexagon arrangement by Delbrück et al. [260]

- **Irregular lumination systems**
  Camba et al. [261]

- **Peripheral halos**
  Vermeulen et al. [262]

8.1.4 *Application scenarios for augmented floor surfaces*

Research has proposed only a few **application scenarios** for augmented floor surfaces, mostly in the domain of entertainment and gaming such as ShareVR [109], MultiToe [257], IGameFloor[263], Space-Hopper [264] or Kickables [265]. For outdoor advertisement in public spaces Camba et al. propose a [261] tiled floor visualization system with irregular lumination realized with optical fiber rods. This makes the low resolution of 6x6 LEDs look more interesting. Additionally they give a general overview of tactile floor setups. In the domain of health and sports use cases, Heller et al. [266] present a smartfloor for motivating people to exercise more by using a floor projection with interactive floor cells. Interaction with the games on the floor are carried out with 50cm x 50cm force weight cells. They present 3 different games "tightrope", "smartdance" and "pong", each intended to encourage people to work out.

Petersen et al. [267] present **design considerations** for floor interaction in architectural environments. They present three interactive floor concepts and use them to derive design issues for interactive floors. They divide the design space into "plaza interaction" and "street interaction." For plaza interaction, no dominant direction is given for the content, due to the multidirectional access vectors. Street interaction on the other hand is characterized by unidirectional access. Hence, more efficient interaction can be assumed.

Law et al. present a **multimodal floor** for immersive environments [268] in 2009. They combine auditory, tactile and visual feedback from users' steps in order to create the impression of "walking over natural ground surfaces, such as snow and ice." The authors argue that by just presenting visual and auditory feedback and leaving out tactile feedback creates a perceptual conflict, which lacks the desired immersion.

Vermeulen et al. provide insights into **dynamic peripheral floor visualizations** with isometrically registered tracking systems [262]. They explore the design space and discuss design considerations for peripheral floor visualizations for conveying users' information on the tracking fidelity of a system, to show borders and interaction zones and to give cues to invite users to perform spatial movements. These kinds of design considerations are applied for the MDE in the VMS, especially the concepts for connecting and mediating interactions with primary interaction devices.

Schmidt et al. argue that the price of having a large-scale floor visualization with direct touch induces bad **ergonomics** for users [269]. They evaluate ergonomics while interacting with a floor visualization

system and derive a novel system for an interactive adaptation of the content to the operator's pose in which the user interacts with the floor. This pose-aware system enables a smooth transition between views and is a countermeasure for "prolonged standing, especially in combination with looking down, quickly causes fatigue and repetitive strain" [269]. The interaction design even lets multiple adapt their personal view to their optimal individual body posture.

Interaction focused research is presented by Schmidt et al. as they propose a set of foot-based interaction tangibles, called "Kickables" [265]. They are intended to be used for "very large interaction surfaces." A set of tangibles for UI controls is proposed and evaluated with the perceptual affordances they take, such as knobs, switches, sliders, radio buttons. They propose them for use in walk up installations.

The literature review reveals that by now there are no research publications on industrial use cases utilizing large augmented floor visualization devices. The author's own previous publication on true scale visualizations for automotive production planning [110] using a scalable floor projection system has been the only contribution in the literature.

8.2 INDUSTRIAL APPLICATION SCENARIOS USING AUGMENTED FLOOR SURFACES

As there is a huge potential of using augmented floor surfaces in the manufacturing domain for collaborative workshops, six application scenarios for automotive production validation are provided below. They are set into the design space context of the above mentioned publications.

8.2.1 *Interactive walk path optimization*

Assembly workplace layouts must be optimized with respect to product, process, worker, ergonomic aspects and the reduction of non-value adding tasks such as walk paths. As proposed by Otto et al. [110] in 2014, an augmented floor visualization system can be used for walk path validations and optimizations. When planning and validating assembly workstations in the VMS, cell layouts must be modified by production engineers, so that simulated walk paths are reduced as they are not value adding to the product. Assessment of drift situations in assembly flow lines is crucial as they degrade the overall efficiency as well. Therefore, virtual simulation tools are applied and workstation layouts are generated virtually for variant-rich simulations.

Large scale augmented floor surfaces display the bird's eye view of these virtual workstation layouts and the corresponding simulation

results as depicted in Figure 8.1. Participants interactively optimize and validate these generated results.
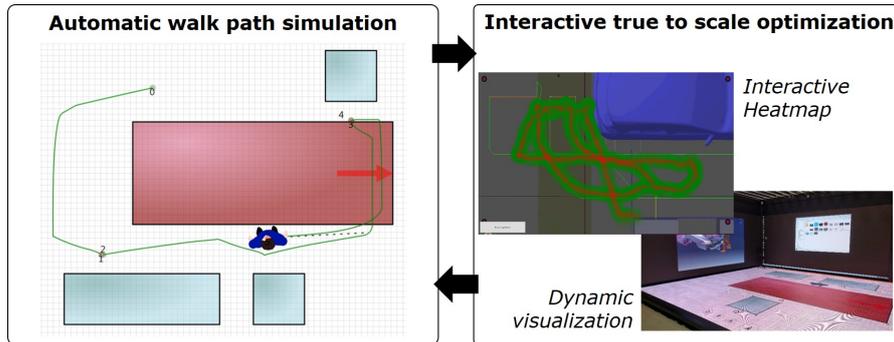
Figure 8.1: Walk path optimizations on the augmented floor surface. A combination of automatically simulated walk paths and interactively recorded walk paths are shown true to scale

In-depth insights into these optimization tasks are provided in the co-authored paper Agethen et al. [68]. Walk paths can be optimized by visualizing (see Figure 8.1 right) the automatically simulated walk paths (see Figure 8.1 left). These simulated walk paths stem from a dedicated motion simulation framework [68] and are displayed on the augmented floor surface true to scale. Subsequently, a user validates the walk path simulation's outcome by means of re-enacting the walking tasks. Walk paths actually carried out can be displayed as heat maps showing parameters of the captured motions, e.g. visualizing speeds, process flows, times, etc.
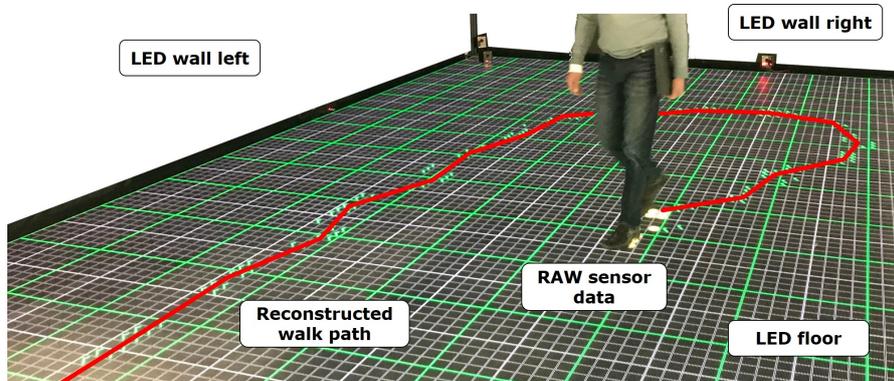
Figure 8.2: Optical sensors of the augmented floor surface are able to detect objects e.g. foot steps. Walk path trajectories can be reconstructed using this sensor information

The presented approach can be regarded as a closed feedback loop between dynamic simulation and the real user's movement. As the synthesized and the user's motions are directly compared, invalid simulation outcomes can be detected at an earlier stage. This increases the maturity of planning data. Additionally, the automatic simulation framework can generate process variants or predict the impact of dif-

ferent parameters (e.g. the height of a person or the weight of a part) according to the acting user's motion.

Consequently, assembly tasks performed on the augmented floor surface (see Figure 8.2) can be enriched further using synthesized data. It is thus possible to cover a wide range of process variants in the overall assembly process simulation while ensuring a high degree of realism by means of comparing the simulation model with the captured walk paths. Setting this application scenario in context with the design space ideas of Vermeulen et al. [262], the user's captured positions are visualized as a "halo" on the heatmap. Walk paths are trails with "historic information" about their actions.

8.2.2 *Layout assessments*

Similar to the use case of walk path optimization, virtual cell layouts are applied to optimize the overall arrangement of resources (i.e. racks, carriers, AGVs) and product parts at a cell, not limiting the scope to walk paths. This also implies checking the availability of all resources, overall fluent processes and process robustness (see Figure 8.3).

In this way, the augmented floor surface helps to display true-to-scale virtual cell layouts visualized from a bird's eye view. Process flows and variants of the process are augmented on the floor so that all participants can geometrically assess the product, processes and tasks at a workstation. As the workshop participants share the same CVE, they are able to discuss cell layouts in-depth.

Virtual workstation layouts consist of simplified representations so that racks and carriers are reduced to 2D or 3D boxes with text labels. By using an orthographic bird's eye view of the augmented floor surface, even complex virtual 3D environments are reduced to 2D projections.
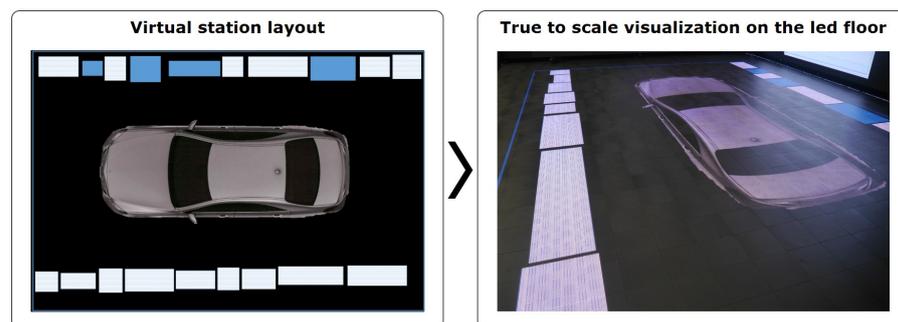
Figure 8.3: Bird's eye view of a virtual assembly workstation layout (left). The content is displayed on the augmented floor surface and presented true to scale to multiple people (right). White and blue boxes represent outlines of carriers and racks

8.2.3   *Virtual true to scale stencil*

Based on the above-mentioned virtual cell layouts in production validation, occasionally these virtual layouts must be built as real-life physical mock-ups, e.g. as a cardboard PMU. Serving this use case, the LED floor can be used as a virtual stencil.
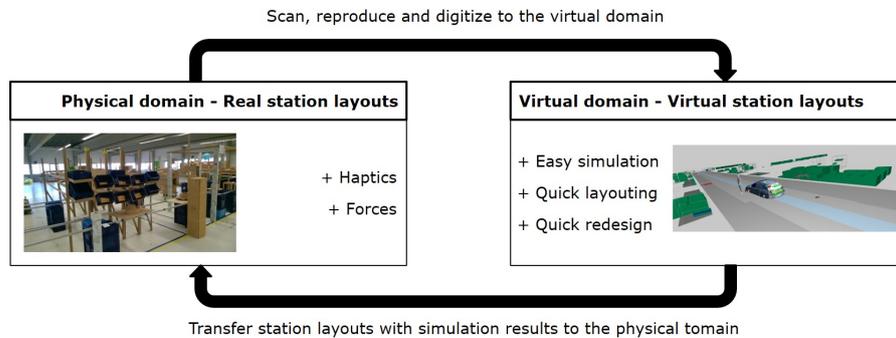
Figure 8.4: Block diagram for the use case cell layout planning. Both the physical and virtual domain bidirectionally influence each other. (left) physical cardboard workshops make use of haptic materials and real-life forces. (right) In the virtual domain, cell layouts can be quickly altered and simulated

True-to-scale representations of virtual contents simplify the process of physically replicating the pre-planned workstation and validating layout variants. Figure 8.4 shows the interaction cycle for both digitization of a workstation and hardware realization in the physical domain. No rulers and no protractors need to be used to arrange physical items, as the augmented floor surface functions as a virtual stencil.

Having physical items present, such as carriers, racks or tools, they can be scanned, remodelled and tracked again. This closes the loop from the physical to the virtual domain. This represents the idea of a digital twin and interaction cycle using the augmented floor surface in the VMS.

Having also tracked physical items, the augmented floor surface is capable of visualizing its virtual meta-information, such as contents of a rack, dimensions of a carrier or simulation results for walk paths. These findings are in accordance with the concepts presented by Müller et al. in BaseLase [258].

8.2.4   *Size perception and engagement*

As production validation engineers make use of these application scenarios on a daily basis, they must be able to estimate sizes properly. In contrast to relatively scaled visualizations, the augmented floor surface helps identify problems with clearances and other geometric

details of virtual cell layouts when showing it true to scale. Rogers et al. found that the arrangement of output devices has a huge influence on the process of idea generation and discussions [247]. Exactly this effect is leveraged by using an augmented floor. In addition, the size perception of true-to-scale contents is expected to be more precise and more accurate.

8.2.5 *Self-navigation in virtual space*

Similar to the concepts presented by Gugenheimer et al. [109] the augmented floor surface helps to opt-in and opt-out in the virtual scene. By registering all display devices isometrically to the virtual scene, production validation engineers can easily see the virtual borders of their region of interest. For example, when assembling the rear back light of a car, the interactively tracked user must get a reference where he is located in the virtual domain. Seeing the car from the bird's eye view of the augmented floor surface, one can easily walk to the region of interest and perform the assembly task there. Even when using a VR head-mounted display, the user can walk to the respective region of interest and put on the VR goggles when he is already perfectly aligned with the virtual domain. Vermeulen et al. propose "Halos" for self-navigation [262] and also visualize borders of the tracking environment on the floor. This feature helps simplify interaction in the virtual domain.

8.2.6 *Virtual travels with interactive maps*

Planning new work contents and tasks for an existing work place requires changes in the respective work cell (brown-field adjustments). It is thus necessary to have a clear understanding of the current baseline situation. Just like "google street view" photospheres for outdoor situations, current indoor workstations are scanned as 360° photospheres and 3D point clouds at defined intervals. In production validation, this data is used to take a look at remotely located factories all around the world.

As depicted in Figure 8.5, the workshop participants simultaneously see the 360° images of the distant workstations and the map. The two L-shaped walls show two perfectly stitched images. The virtual cameras always keep a 90° offset around the vertical axis. Additionally, the augmented floor visualizes a circular map of the factory perfectly aligned with the heading of the viewport of the two walls. This provides the user with a perfect overview of where he is currently located and in which direction he is looking. They dive into real world scans and understand the circumstances directly. They can walk through the distant factory, look around and measure directly in 3D space.

Figure 8.5: Usage of the augmented floor surface for virtual travels. (Left) The schematic shows the information architecture on both walls and the floor. (Right) The walls allow an immersive deep-dive in 360° photospheres, each virtual camera having a 90° offset around the vertical axis. The floor map orientation changes accordingly to the wall viewport

8.3 IMPLEMENTATION AND EVALUATION OF A SCALABLE FLOOR PROJECTION SYSTEM

A scalable floor projection system is presented. Its technical specifications and implementation details enable the showcasing of the above-mentioned real use cases and an evaluation of the effectiveness of a true-to-scale floor projection system in the context of the VMS.

The augmented floor projection system is developed for PV workshop environments with several constraints:

- **Size:** The projection area must be large enough to visualize work places true to scale.

- **Portability:** The system must be portable so that it can be moved within a day.

- **Height:** The overall height of the system must be limited to 2.5m due to factory restrictions.

- **Light density and contrast**: The projection system must be brilliant enough, so that all workshop participants have good legibility with modest ambient light.

- **Low-cost implementation**

- **Pixel pitch:** Typical virtual cell layouts do not inherit high information density. So using a pixel pitch of less than 5mm is presumed to be an acceptable trade-off between realization costs and visualization quality.

Aber nur, weil es in dieser Anwendung im Wesentlichen um Pläne (Zeichnungen) geht!

Taking these technical requirements into consideration, a scalable floor projection system is developed and evaluated.
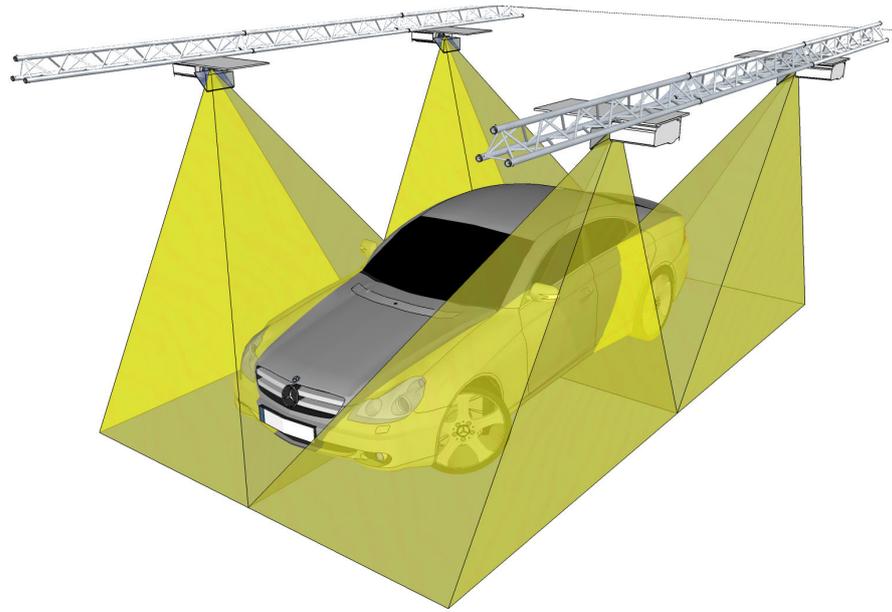
Figure 8.6: Hardware arrangement of floor projection with light paths and projection cones

8.3.1 *Hardware implementation*

The system is composed of four overhead short-throw DLP projectors, four front facing mirrors, video wall controllers and several pieces of event trussing and profiles. Figure 8.6 depicts the 2x2 arrangement with a car standing in the projection cones.

The implementations inherits four projectors, which allow for a maximum projection area of approx. 18 sq.m. in a medium ambient light.

8.3.2 *Image distribution in the system architecture*

The floor projection system consists of multiple projectors. A multi display signal controller with four outputs is used, which can be driven as a dual-link DVI display and can represent an arbitrary crop region of the original input signal. Input images can be cropped, scaled, mirrored and bezel corrected. Input signals are accepted up to 4k x 4k resolution via Dual Link DVI signal. One single virtual display is presented to the computer. The native resolution of this virtual screen is set to 4k (3840 x 2160) at 60Hz image refresh rate. As depicted in Figure 8.7, multiple concatenated display wall controllers are used in order to achieve a large size stitched image effect.

Multiple controller instances enable scalable flexible floor projection systems. Datapath dL8 enables 1:8 image cropping, in the second concatenation step, datapath x4 enables 1:4 flexible image cropping and output. These standalone controllers can crop arbitrary regions,
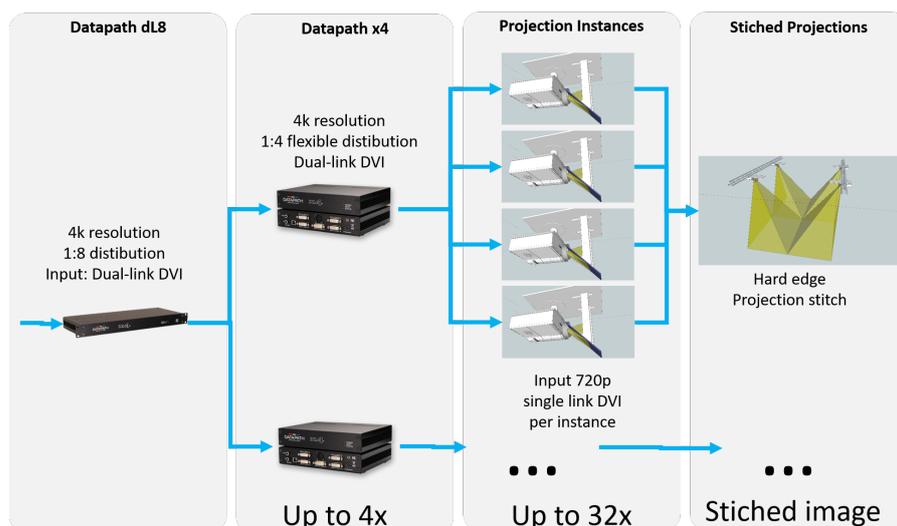
Figure 8.7: System architecture of floor projection system

scale, rotate and replicate in the input channel and (up-)scale it to native DVI output resolutions at a desired frame rate (here 60p). The system can be scaled for up to 32 projectors without altering the hardware setup (Figure 8.7).

8.3.3 *Implementation of projection units*

The projectors used in this setup are off-the-shelf, semi-professional BenQ projectors. Each projection unit integrates the projector with front surface mirrors in order to be able to mount the projector horizontally and redirect the beam to the floor. The projectors are not mountable in each orientation but only in upside-down ceiling mount and upright stand because the hot air must be transported out of the device. If mounted vertically, the projector would be damaged quickly. Each projector has a short throw lens with a variable throw ratio from 0.72 to 0.87. The luminous flux of each projector is specified to 2500 lumens integrated to redirect the picture to the floor.

Since the optically effective coating of the mirror is located directly on top of the surface of a glass, good projection quality is maintained with this mirror. Regular back surface coated mirrors would generate ghost projections. The minimum size of each mirror depends on the throw ratio, optical offset, distance of the projector to the mirror, the elevation angle of the mirror and the projector.

Two hardware setups have been realized in different setups. One floor projection instance is designed to be used in a 2x2 (see Figure 8.6 and Figure 8.8 left) configuration and the second hardware setup is a 4x1 configuration, where the long sides of each image are stitched. Each single picture has a typical size of 3 m x 1.68m (variable with height of rigging and zoom factor) and the resulting augmented floor
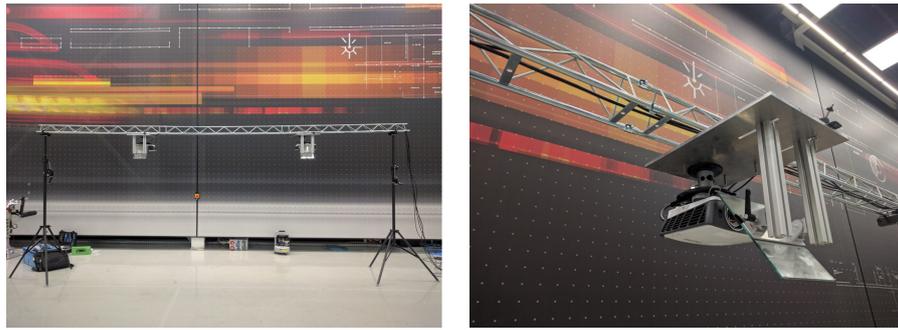
Figure 8.8: Implementation of the floor projection system. (left) Two projectors on trussing (right) close view of projectors with front-facing mirror

areas have sizes of 6m x 3.36m and 6.75m x 3.00m respectively. This is sufficient for a true-to-scale visualization of a station on one working side. The stiffness and tolerances of the ball joint mountings should be improved. A well-calibrated system setup has minor visible edges, which differ less than 3mm due to hard edging technology. Yet, no soft-edging functionality is implemented, so that the adjacent edges are smoothed out by a software generated deformation which could blend two or more projections smoothly. Typical software component for enabling soft-edging is Wings AV.

Figure 8.8 (right) depicts a single side of the setup and shows the possibility for fine adjustments of the projectors. The angle of the projector, angle and height of front surface coated mirror and distance can be adjusted flexibly depending on the required hardware setup.

No special software is required to utilize the floor projection system as a regular display unit. All images are presented just as they would be on a regular desktop screen. Nevertheless, true-to-scale visualization of digital simulation content is driven by correct parametrization of the simulation software components used with the floor projection system. Using true-to-scale software, the output must be parametrized to the correct display height, width and aspect ratio.

8.3.4 *User study*

To obtain insights into the implementation quality and limitations of the presented floor projection system, a brief user study has been carried out. Users were observed using the system in PV workshops and qualitative feedback was collected at the end of the supervision time frame. A non-standardized questionnaire was handed out, focused on effectiveness in PV workshops rather than its enjoyable character. Trade-offs in system design were discussed with the users.

Twelve production planning engineers took part in this expert survey. All attended one specific PV workshop. The following results do not claim statistical significance due to the low number of participants
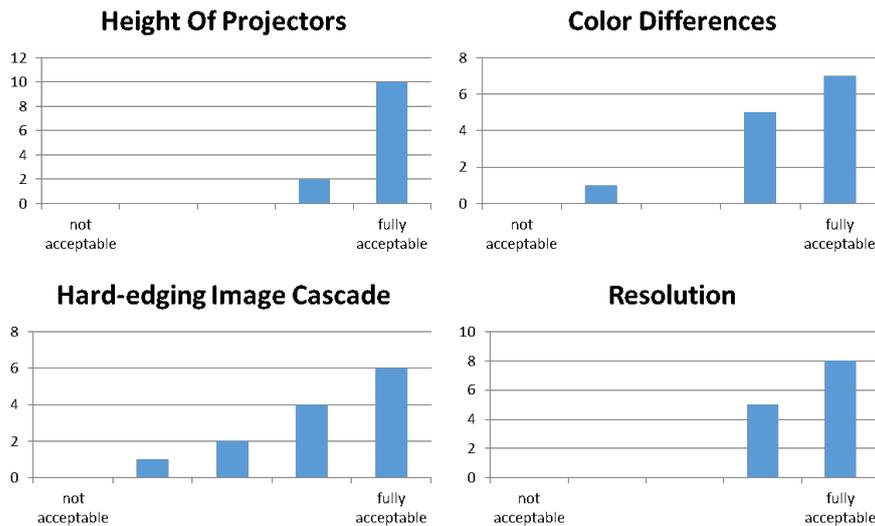
Figure 8.9: Survey results on: height of projectors (upper-left), color differences (upper-right), hard-edging image cascade (lower-left) and resolution (lower-right)

(N=12), but show a tendency regarding how workshop participants work with the system. All participants took part on a voluntary basis, did not receive any extra reward and reported normal to corrected vision.

The participants worked with the system for multiple days. At the end of the last workshop day, the author handed out a questionnaire to each interview partner after they had worked with the floor projection system for an entire day in PV workshops. They used the system for the above mentioned walk path optimization tools with true-to-scale floor visualization. A non-formal, verbal discussion took place after filling out the questionnaires.

Results show that all twelve production engineers agreed that the height of the projectors at 2.5m was absolutely sufficient (M=4.84, STD=0.37) to be unobtrusive for daily work (see Figure 8.9 upper-left). Most engineers rated the color differences as acceptable (M=4.38, STD=0.84), because they were not crucial for reaching their assessment goals, and items within virtual assessment scenes were often modeled using random colors (see Figure 8.9 upper-right).

The same feedback was provided on the resolution of the augmented floor surface image (M=4.62, STD=0.49) (see Figure 8.9 lower-right), even though the system had a pixel density of 4.26 px/cm. The hard-edging properties of the system were rated more diversely (M=4.16, STD=0.95) but tended to be positive. Discussions with experts pointed out that a small gap between the projection areas could be accepted more easily than a small overlap with double the brightness. More elaborated projectors offer soft-edging ability natively.
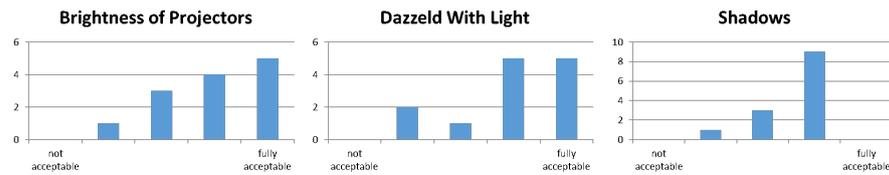
Figure 8.10: Survey results on: brightness (left), dazzled with light (middle) and shadows (right)

Figure 8.11: Survey results on: true to scale visualization (left), efficiency improvements (middle) and verification of walking paths (right)

The results of the feedback regarding the system's brightness were surprising (see Figure 8.10 left). The experts rated the brightness as sufficient overall (M=4.00, STD=0.96), with a projectors light intensity of 2500 ANSI lumen. Since the brightness of projection systems can be increased, one limiting factor is the available budget. Interestingly, some people felt dazzled by the light (see Figure 8.10 middle) when walking through the projection cone. The majority of people tend to be able to cope with this property (M=4.00, STD=1.04). One way to improve this property could be using projectors with higher throw-ratio lenses and mounting them higher in the ceiling, in order to have a smaller projection cone and a higher angle of incidence. The most diverse results and most discussions of the experts were related to shadowing effects. By tendency, people found self-occluding shadows "neutral to acceptable" (M=3.61, STD=0.62) (see Figure 8.10 right). For the stakeholders of such a projection system, it seemed to be the biggest drawback. As depicted above, there are several solutions to eliminate shadowing effects using LED floors, LCD floors or rear projection systems.

In a nutshell, the experts were satisfied with the hardware implementation with some potential for improvements.

An unambiguous result of the survey was that true-to-scale visualization supports the experts in spatial tasks (M=4.69, STD=0.46). All engineers tended to support the statement (see Figure 8.11 left) that the true-to-scale augmented floor surface is more suitable for estimating lengths, spatial relations and speeds, compared to common desktop displays or a powerwall.

When asked about generic efficiency improvements (see Figure 8.11 middle) by usage of true to scale visualization in contrast to common desktop systems, the results were diverse (M=3.62, STD=1.00). In the

more precise case of walk path analysis, production engineers saw larger (M=4.31, STD=0.82), efficiency benefits (see Figure 8.11 right).

The experts responded to the overall question whether they would like to use the system in general. 12 out of 12 answered "yes" to the question, meaning that they would like to use the system operative. This fact proves the system's big potential.

8.3.5 *Discussion and conclusion*

An unambiguous result of the user study is that true-to-scale visualization helps the experts in spatial tasks like material zone planning. All engineers supported the statement that they would like to work with the system. Users had the subjective impression that it is more suitable for estimating lengths, spatial relations and speeds than using common desktop displays or a power wall.

Asking users about the system's overall efficiency revealed the question to be too generic as efficiency is not only determined by the hardware, but by all components involved in the system, such as hardware, software, users and environment. More specific experiments with objective task completion time evaluations will be carried out in order for the latter to better distinguish between the performance of the visualization software and the augmented floor surface hardware.

This **prototype setup** of a scalable floor projection system in combination with true-to-scale visualization software for assembly line cell layouts proved work reliably. The proposed hardware setup is movable, easy to use and scalable. The introduced hardware setup offers a cell-sized visualization area realized with low-cost hardware. The system is installed, introduced and piloted during on-site PV workshops for a mid-size premium car.

The projection cascade can easily be enlarged to show multiple cell layouts simultaneously in-true-to-scale visualization. Since the prototype proved to work reliably, the VMS concepts apply the augmented, interactive floor.

8.4 IMPLEMENTATION AND EVALUATION OF A LED VISUALIZATION SYSTEM

As the prototype installation proved to work reliably and the PV workshop participants were satisfied with the implemented floor projection capabilities, the VMS framework applies the concept and extends the presented visualization capabilities of the aforementioned prototype with a different technological approach:

In the past few years, more and more **LED walls** have started to replace projection systems, since their pixel pitch is drastically reduced, contrast is higher, the viewing angle is optimized and, most importantly, the price per square meter has decreased. In this implementa-

tion approach, no projection units are used due to their limited contrast and susceptibility to ambient light changes. Therefore, the VMS framework's output realization is based on state-of-the-art LED technology. In comparison with the floor projection setups, no stitching, soft or hard-edging capabilities are required, which results in less maintenance work. Since LED tiles for floors and walls are rectangular and visualization units can be seamlessly concatenated in two dimensions, only the physical arrangement of modules has to match properly on a sub-millimeter basis.

8.4.1 *Multi LED display environment arrangement*

The VMS display apparatus represents a MDE. The apparatus consists of three LHRDs. Two identical LED walls, each sized 6.0m x 2.7m, are arranged in a 90° L-shape with a closely attached large scale LED floor (see Figure 8.12). The specifications of the LED walls and the LED floor are shown in Table 8.1. Each powerwall has a pixel pitch of 1.25 mm and allows a 5k resolution.The LED floor contains proximity sensors for further applications such as walk path reconstruction without any external 3D tracking systems.

Figure 8.12: The VMS apparatus consists of two 16 sqm, L-arranged LED walls next to a 54 sqm large-scale LED floor

Even though the setup resembles a CAVE with its three adjacent LED displays, the setup intentionally leaves a gap of 120mm between each wall and the floor. On purpose, the setup partly forgoes the immersive effect of a typical CAVE (see Cruz-Neira et al. [255]) in order to achieve maximum mechanical flexibility. At the borders of each display, standardized aluminum profile structures make it possible to mount additional devices, such as sensors, haptic devices, cameras and other tracking equipment.

Table 8.1: Specifications of VMS output apparatus

Property	LED walls [each]	LED floor
Active Area	16 sqm	54 sqm
Pixel pitch	1.25 mm	5 mm
Resolution per wall	4800 x 2160	1728 x 1152
Size	6 m x 2.7 m	9 m x 6 m
Max. power consumption	20x 440W	216x 240W
Typical power consumption	20x 100W	216x 100W
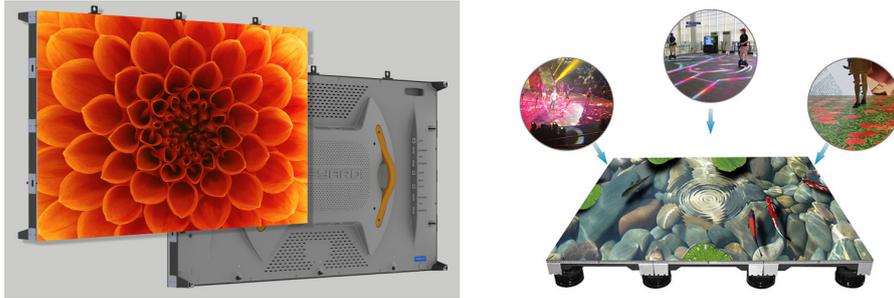
Figure 8.13: Product image of LED wall display cabinet Leyard TWS 1.2 (left) and product image of Uniview LED I Series 5.2mm (right)

On each of the two LED walls, seamless images are generated by using 20 concatenated cabinets of LED modules summing up to a 6.0m x 2.7m setup. Each LED wall consists of 5x4 54" Leyard TWA cabinets [270] (see Figure 8.13 left) with a pixel pitch of 1.25mm, resulting in an overall native resolution of 4800x2160 pixels (see Table 8.1). The maximum brightness is defined as 800cd/sqm with a horizontal viewing angle of 160°. For indoor usage, brightness is set to approximately 20% intensity without daylight influence. Colors change when looking from steeper angles due to a limited horizontal field of view of the single LEDs (160° horizontal viewing angle).

The LED floor consists of 216 "Uniview LED I Series" cabinets resulting in a 9.0m x 6.0m setup. Each cabinet has a dimension of 500 mm x 500 mm and therefore has an arrangement of 18 by 12 cabinets for the whole floor (see Table 8.1). Dalton et al. found that "pixel density, over the range of tests, is less important than visual artifacts introduced by carpet tile edges" [259]. Carpet tile edges can hardly be seen for the concatenated floor tiles in this setup. These seamless tiles also offer interactive direct touch capability with optical proximity sensors for floor step detection as depicted in Figure 8.2. There are 16 sensors in each floor module with a latency of 10ms. For practical design considerations, the LED floor tiles are water resistant IP65, have an anti-scratch coating, can be replaced without additional adjustments and can carry point loads of up to 2000kg/sqm. For automotive production validation, the latter specification is required for
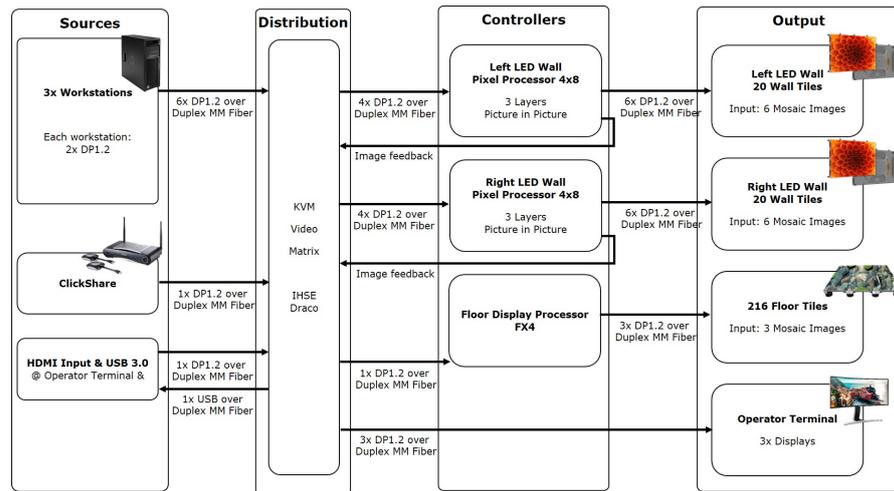
Figure 8.14: System architecture of the VMS image distribution and controllers to the devices

mixed reality assessments, putting a physical car body as a mock-up in the VMS and registering it to the virtual scene.

8.4.2 *System architecture*

The MDE is designed to display up to five pictures in parallel for massive data visualization. The system's architecture inherits four stages of image processing (see Figure 8.14):

Sources: PV workshop participants can either use locally available hardware at the VMS or bring their own devices to visualize content. Three CAD workstations are available at the VMS which output three images, each consisting of 2 mosaic pictures due to DP1.2 bandwidth limitation and the high resolution of the LHRDs. DP1.2 limits the resolution to 4k at 60Hz. That's why each monitor output is split up to 2 4k images and stitched together in image processors. These mosaic tiled images are frame-locked and sync-locked for 3d visualization. For wireless sharing of notebook, tablet and mobile phone content via standardized protocols, e.g. AirPlay, GoogleCast, MiraCast, a video receiver "Barco Clickshare CSE-800" is applied. This is attached to the video matrix using two 4k DP.1.2 connections. The last input source is a regular, wired HDMI cable specified for HDMI 1.4 standard. Figure 8.14 shows a block diagram of the entire image distribution process.

Distribution: All images of the sources mentioned above are acquired from input capture cards and passed to the modular signal distribution system "IHSE draco" for video and Keyboard, Video and Mouse (KVM) switching. All images can be distributed to one or multiple arbitrary output cards or devices. Fiber connections enable distant remote output devices, no restrictions apply in terms of distance,

compared to regular display cables. For changing the distribution configuration, a Crestron interface is applied.

Controllers: All display controllers capture the image signals received by the video matrix and create a composite image distributed to the output devices. For both the left and right LHRD, powerwall image processors made by "Brainsalt" are applied for resizing the image, splitting the image onto the 20 wall tiles via 6 DP1.2 connections and to merge multiple (mosaic-) images from different sources as Picture in Picture (PiP) or Side-by-Side Views. The resulting image composition is visualized on the powerwall and is fed back into the video matrix, so that each device can also retrieve the video composition of the pixel processors. In contrast to the highly flexible pixel processors, the "Datapath FX4" floor display processor just resizes the image and splits it up in three signals to be output on the floor device.

Output: Both floor and LED wall components are DisplayPort daisy-chained systems, so that the image is passed through module by module. For the walls, the image is split up to FullHD signals, whereas the floor is fed by 3 low-res mosaic images. 5k operator terminal displays are directly fed by the video distribution system without any display controllers.

8.4.3 *Graphical user interface for image distribution*

The technical realization is complex in terms of switching behavior as can be seen in Figure 8.14. That is why a tablet-based Crestron UI system is used for realization of an intuitive user interface for switching both the video distribution system and the pixel processors for the LED walls. Reducing complexity for easy usage is the main aspect of the tablet-based UI. The overall image distribution process is shown in Figure 8.15.
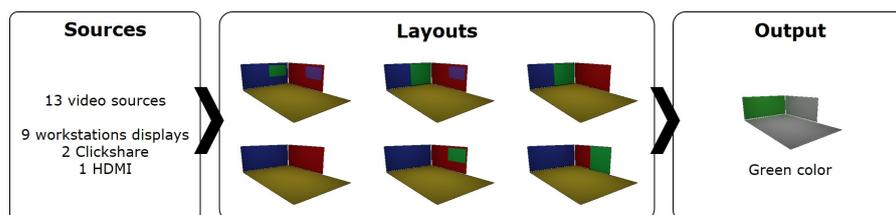
Figure 8.15: Block diagram of Crestron UI image distribution

Output: Single monitor outputs can either be displayed on the full physical screen or be split via predefined layouts. Output areas are virtual displays and - for simplicity's sake - each output is assigned to a certain color. Sources are not directly assigned to physical walls, but on virtual displays first, respective colors in the UI. These virtual displays are arranged using layouts. Each assignment from sources to output areas can be seen as a 2D matrix routing table. For simplicity, a two-click procedure for pairing sources and outputs has been de-

veloped: Choosing the desired input source in a list and afterwards assigning it to a virtual display or the other way around.

Layouts: Arrangements of virtual displays are called layouts. Layouts arrange content on the floor and walls, for example in a side-by-side or picture-in-picture manner. Therefore each LHRD can be either used as a regular display or with multiple source contents.

Assistant: The assistant offers the users six presets which automatically distribute a whole setting to the complete system. With one click, all sources and outputs are set to the respective layout, the console display outputs and the KVM HID devices. For example, all six mosaic images generated by the CAD workstation are distributed on the walls, on the floor and on the 3 preview monitors of the operator terminal. Additionally, the HID devices of the operator terminal are automatically assigned to the correct workstation. This assistant reduces the single display assignment work of switching between different layouts drastically. Without the assistant, the whole procedure of routing all sources to the proper output units takes 10 clicks, whereas the assistant reduces this work to one click.

Having implemented such a LHRD MDE, one research question is how people perform in size estimations using this technology in the context of PV workshops.

As this LHRD setup is utilized by production engineers on a daily basis for the above mentioned use cases, the question arises whether they really improve in their personal size estimations and whether they become faster using these tools provided by the VMS. For validation purposes, people have to estimate sizes as precisely and accurately as possible in order to judge the validity of planning data, even when utilizing complex 3D models in combination with an orthographic, non-tracked virtual camera, visualization contents, such as racks and carriers. Czerwinski et al. also indicate that LHRDs have an influence on productivity and satisfaction [237]. Size perception is thus presented, using 2D content representations, just like in the aforementioned use case "layout assessments."

8.5 TRUE TO SCALE SIZE PERCEPTION STUDY

This study focuses on how people perceive sizes of virtual contents using immersive display technologies with true-to-scale data visualization. Therefore, size judgment performance is compared between three scenarios, two of them showing to-scale data representations on an LED floor and one showing relative-sized visualizations on a tablet computer.

Amongst other benefits, pervasive displays are said to allow more accurate size judgments, improved collaboration performance in workshops and reduced task completion times. But research does not yet provide a definite answer on whether the usage of LHRD displays

leads to better size judgements and whether users can become faster when using it. "Better size judgments" imply higher spatial accuracy and higher spatial precision in the user's estimations. Therefore, this study follows a call for research [271] and investigates size judgment accuracy, precision and task completion time by using the VMS's large led floor showing absolute scaled contents compared to tablet computers with relative scales.

> **Related publications**
>
> This study was initially created for this doctoral thesis and was already been published as a peer-reviewed conference paper [241] by Otto et al. at the "ACM Pervasive Displays" 2019 and as a Springer journal paper in "Personal and ubiquitous computing" [242]. This section is based on the papers mentioned, while texts have been extended and revised for this doctoral thesis.

However, basic research on using LHRD technology has not focused on human perception. Bezerianos et al. present a call for research [271] on the perception of data on wall-sized displays as little research has been carried out in this domain so far: "We do not yet know how the perceptual affordances of a wall, such as the wide viewing angles they cover, affect how data is perceived and comprehended" and call "for more studies on the perception of data on wall-sized displays." Using different types of display devices directly influences spatial perception, visual space and the control of spatial behavior, especially when using display arrangements such as an LED floor. We follow this call for research and add an additional element: Data visualization on large-scale augmented floor surfaces showing contents in absolute scale.

In the **purely physical domain**, the literature has focused on size and distance judgments for a long time. In 1963, Epstein [272] presented their key findings, that distance and size judgments were not systematically related and deviations of size judgments varied with distance. Later, Epstein and Broota [273] presented another evaluation of the judgment of sizes and distances and the corresponding reaction times. They found a positive correlation between viewing distance of objects and the reaction time. In Wagner's publication "The metric of visual space" [274], he provides insights into judging distances, angles and areas as conducted in this study. Cleveland and McGill present groundbreaking work in the visual decoding of information, namely graphical perception. They propose a set of elementary perceptual tasks, detailing how people extract quantitative information [275]. More recently, Talbot et al. pick up these publications and analyze the reasons for the differences in perception of charts [276].

For **virtual environments**, extensive research has been carried out on perceived spaces in VR, such as distances, sizes, speeds and spaces. Loomis et al. show that egocentric distance judgments in physical environments nearly match 100% of the actual distance [277], whereas in virtual environments they are frequently underestimated. Renner et al. present a literature review and conclude "mean estimation of egocentric distances in virtual environments of about 74%" [278]. Renner et al. also cluster possible influence factors for this under perception of sizes in four different clusters: measurement methods, technical factors, compositional factors and human factors. In contrast, current state-of-the-art head mounted displays seem to ameliorate these effects [279]. Kelly et al. show that when using modern HMD devices this effect is reduced but has not been completely resolved. In comparison with the literature, no relative size judgment has been carried out in VR by providing users with relative scales.

To the best of this dissertation's author's knowledge, this is the first time such a size judgment experiment is being executed using a large-scale LED floor setup in comparison to a small-sized baseline measurement.

8.5.1   *Study goal and predictions*

One of the striking benefits of a large-scale display is the possibility to visualize true-to-scale data, contents and virtual scenes. In the context of the presented use case within the automotive industry, 3D contents with individual view points have been intentionally excluded, whereas 2D representations (see Figure 8.12) have been chosen for this study as the aforementioned use cases are limited to data visualization of 2D data.

This evaluation gives insights into whether participants can assess sizes of 2D contents more accurately and precisely when they are shown true to scale compared to relative-scaled representations. The baseline scenario represents relative-sized visualizations on a tablet computer, showing exactly the same visual cues as in the true to scale scenarios. In this study, size judgment refers to the edge length estimations. To date there is no published research documenting the extent to which true-to-scale floor content supports participants in estimating sizes using augmented floor surfaces. To address these issues thoroughly, this study employs verbal distance judgments and objective measurements. Four different aspects are evaluated in this study:

- **Accuracy:** Is there a systematic over- or underestimation (accuracy) of size judgments? (Mean absolute percentage error, see Armstrong and Collopoy [280])

- **Precision:** In which scenario do participants come to the most precise size judgments. (SD of mean absolute percentage error).

- **Task completion time:** Is there a difference in task completion time for the three different scenarios? (Objective time measurements)

- **Qualitative feedback:** Do the user's subjective judgments on precision and task completion time match the objective measurements? (Non-standardized questionnaire)

8.5.2 *Participants*

For this study 22 voluntary participants were randomly selected, such as production engineers, research engineers, PhD candidates and students from different production planning departments in the manufacturing industry. 15 males and 7 females took part, ranging in age from 21 to 57 years. (M=31.57, SD=11.52). All participants reported normal to corrected vision and chose the metric system as their preferred unit.

8.5.3 *Setup, stimuli and design*

For this evaluation, the VMS's augmented LED floor surface is applied as described in Section 8.4.1. Three different modes of perception are evaluated. For all three scenarios, the same visualization software, visual cues and interaction (besides user's movement) are used, only the output modality is changed (see Figure 8.17):

- **Tablet scenario (T):** Relative-sized visualizations as a baseline

- **Floor scenario (F):** True-to-scale visualization restricting user's viewpoint on the side of the LED floor

- **Floor and Interaction scenario (FI):** True-to-scale visualization allows the users moving on the whole LED floor

The rendering and evaluation software is a custom application which displays virtual squares in a randomized order (six different sequences for 3 scenarios). The randomized scenario work flow and evaluation data logging (square size, square rotation, pixel per meter, scenario completion time) is handled by the individual software as well. In all three scenarios, the participants are shown 2D white squares on a black background. These squares have randomized sizes from 50 cm to 200 cm with random positions and orientations (+/-15°) on the screen (see Figure 8.16). Additionally, a virtual ruler represents the absolute length of one meter and remains at the same position (center bottom) throughout all scenarios. The scenarios (F) and (FI) utilize the
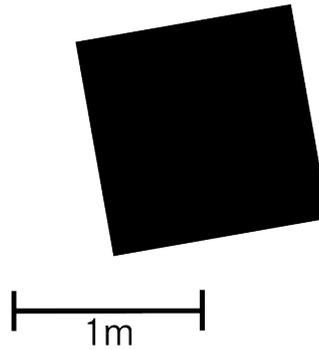
Figure 8.16: In all three scenarios a square is shown. It is randomly scaled, positioned and rotated. Additionally a 1 meter ruler is given as an additional visual cue

Tablet (T)
relative scale

Floor (F)
True to scale

Floor Interaction (FI)
True to scale
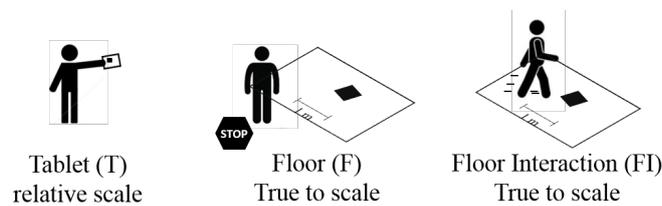
Figure 8.17: Three evaluation scenarios: Tablet(T), Floor(F) and Floor Interaction(FI). Left: The user carries out the size estimations using a tablet computer. Center: The user utilizes the augmented floor surface, standing on the outside. Right: The user moves on the floor while performing the size estimations

aforementioned $9\,\text{m} \times 6\,\text{m}$ LED floor apparatus with $10.81\,\text{m}$ screen diagonal, whereas scenario (T) visualizes the content on a 12,3" tablet screen. All scenarios are set to the same display aspect ratio as the LED floor (3:2). The LED floor pixel pitch is $5\,\text{mm}$.

8.5.4 *Procedure*

After signing the informed consent, the participans are given verbal instructions on the goal and evaluation procedure. Each participant executes all three scenarios (T), (F) and (FI) (within-subject design) in a randomized order to abolish learning effects. There is no interaction with the virtual contents, so that the focus is limited to the differences in spatial perception. In each scenario 20 randomized (size, rotation, position) squares are visualized. After presenting each square, the participants verbally express their size estimate to the experimenter in the unit centimeters. The experimenter writes down the response for each estimation in parallel.

The three different scenarios are depicted in Figure 8.17 and described as follows:

- **Tablet (T):** The software visualizes the squares on the tablet computer as relatively sized content. The users have to judge the absolute edge length in relation to the visualized ruler.

- **Floor (F):** The software visualizes the squares on the LED floor to scale. The participants are directly facing the LED floor from a static location (compare [279]), standing on the outside border, centered on the long edge of the LED floor (3m to the center) and may not access it.

- **Floor&Interaction (FI):** Same setup as in scenario (F), but in contrast, the users have the opportunity to move freely on the augmented floor during the study, so that the subjects may position themselves directly above the respective square.

The experiment was conducted a total of 22 times with different participants. Each evaluation took approximately 20 minutes including the subsequent completion of the questionnaire. A total of 1320 datasets were collected (22 participants, 3 scenarios, 20 trials), each one containing the actual & reported length [cm], spatial deviation/error [cm], task completion time [ms], pseudonym, scenario, square rotation and position.

Finally, participants were asked to fill out a questionnaire after completing all three scenarios to gather their subjective feedback. They were asked about their personal scenario preferences for direct comparison. In addition, each subject was to select their preferred method and specify the reason for their decision.

8.5.5 *Results*

The following results are clustered in the three sections: Accuracy, precision and task completion time. Spatial deviation is the difference between the actual edge length (ground truth) of the squares and the estimation of each participant for the respective square edge length. Negative values represent an underestimation of size and positive values represent an overestimation.

Figure 8.18 shows a scatter plot of all three scenarios depicting the true length [cm] over the difference between true and estimated length. All three scenarios show that in mean, there is only little overall over- or underestimation of the user's size judgments with (T) having a mean of $0.951\,\text{cm}$ (SD=30.204), (F) $-0.634\,\text{cm}$ (SD=22.499) and (FI) $-5.694\,\text{cm}$ (SD=17.850). However, regarding the relatively large standard deviations compared to the small means, the interpretability of the aforementioned spatial deviation is disputable due to over- and underestimation. Furthermore, spatial deviation tends to increase with a growing edge length of the squares, especially when considering (T) and (F). In order to normalize these effects, in the following the mean absolute percentage error (MAPE) and mean stan-

richtig!
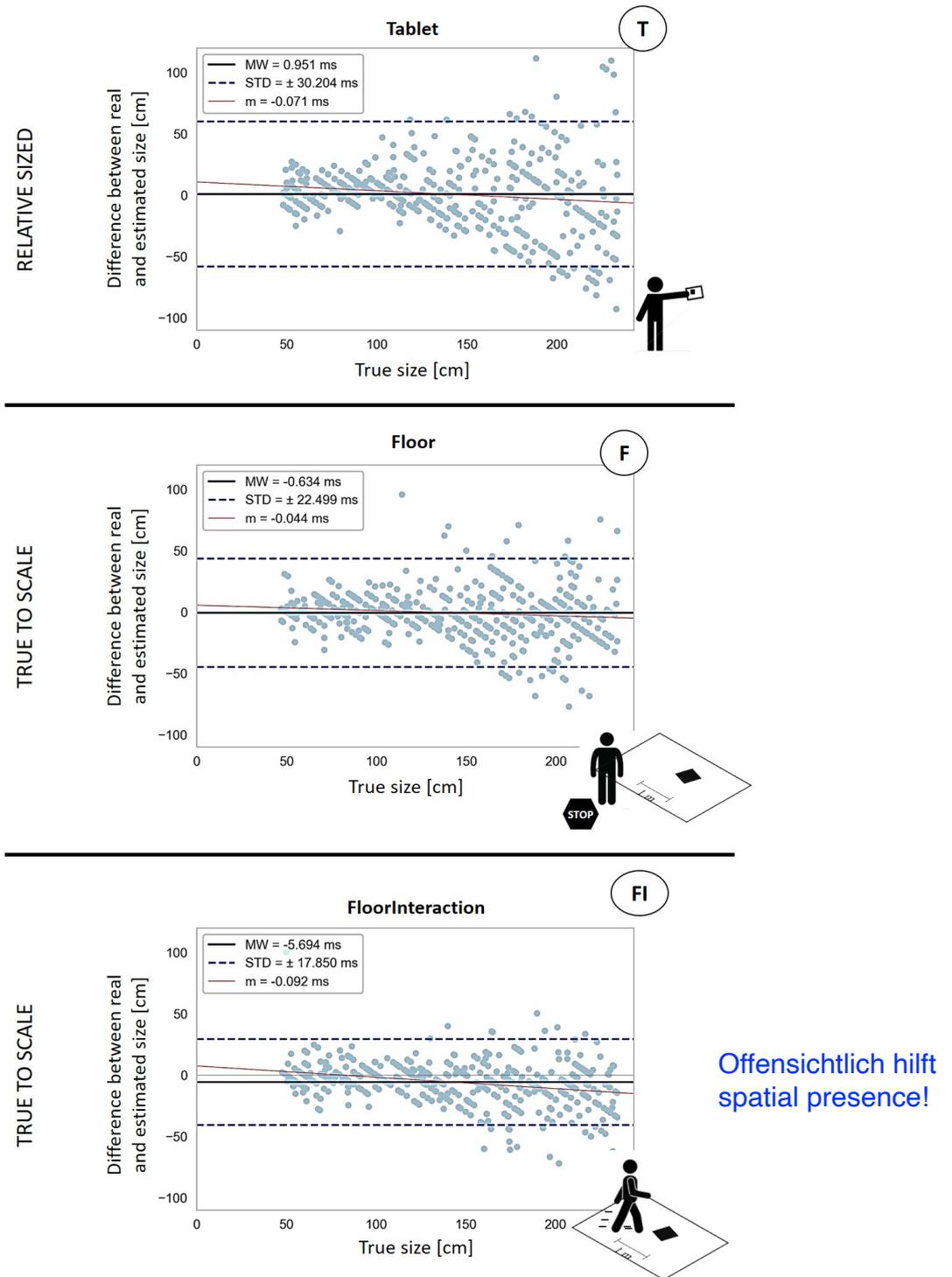
Figure 8.18: Scatter plots of all three scenarios show the spatial deviations. Each plot follows Bland-Altman plot [281] style, additionally showing the mean values, standard deviations and a linear regressions over the actual visualized cube sizes. One can see that variance of spatial deviations over the true sizes tends to increase in all scenarios
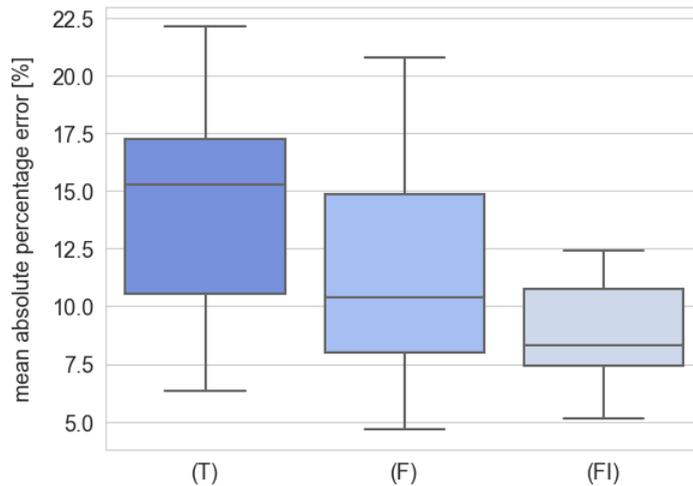
Figure 8.19: Box plot for MAPE of scenarios (T), (F) and (FI)

dard deviation (SD) of MAPE for trials within subject are used to evaluate accuracy and precision in all three scenarios.

8.5.5.1 *Accuracy*

MAPE is a measure of prediction accuracy. (T) shows a mean absolute percentage error of 14.783% (SD = 5.612%), (F) 11.369% (SD = 4.599%) and (FI) 9.814% (SD = 3.957%). Figure 8.19 depicts the box plots of the MAPE of all three scenarios. A statistical comparison is performed considering (T), (F) and (FI). Levene's test shows that variance homogeneity is given for this data ($F_{(2, 63)}$ = 0.942, p = 0.395), therefore the standard one-way ANOVA can be used. One-way ANOVA reports statistically significant differences between the three scenarios ($F_{(2, 63)}$ = 6.242, p = 0.003). The post-hoc pairwise t-test with Holm correction reveals that there is no significant difference between (FI) and (F) (p=0.284), but for both other scenarios (T) and (F) (p=0.041) and (T) and (FI) (p=0.003).

Overall, the MAPE of both true-to-scale visualization scenarios (F) and (FI) can be regarded as significantly different from the relative scaled (T) scenario. As both mean MAPE values are lower, the scenarios (F) and (FI) have a higher accuracy compared to (T).

8.5.5.2 *Precision*

The mean SD of MAPE for trials within subject demonstrates the precision of size judgments represented by the "variance of absolute percentage errors." (T) shows a mean SD of 10.006% (SD = 3.394%), (F) of 9.759% (SD = 6.051%) and (FI) of 8.921% (SD = 7.898%). Figure 8.20 depicts the SD of MAPE for trials within subject box plots of all three scenarios. Levene's test is utilized for testing equality of the variances in distributions. With $F_{(2, 63)}$ = 0.329, p = 0.721 shows
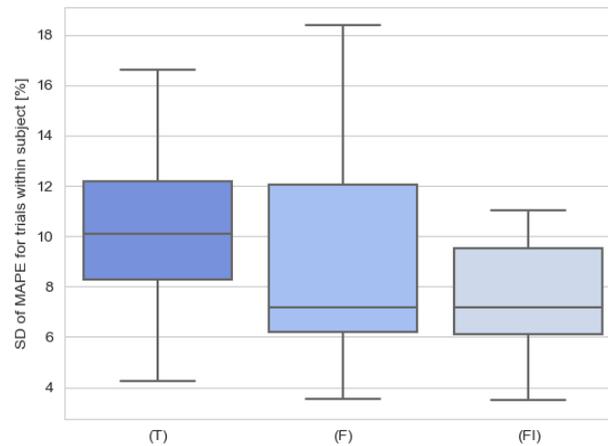
Figure 8.20: Box plot for SD of unsigned percentage errors in scenarios (T), (F) and (FI)

that variance homogeneity is given for the SD. Therefore standard-one way ANOVA with post-hoc pairwise t-test with Holm correction can be used in this case which reports $F(2, 63) = 0.184$, $p = 0.832$. Since one-way ANOVA shows no significance, post-hoc test results are not reported here.

No significant difference in precision can be found using true-to-scale visualization scenarios (F) and (FI) compared with (T). However, considering the descriptive statistics of mean SD of MAPE for trials within subject, a minor tendency of lower precision of (T) compared to (F) and (FI) is depicted (see Figure 8.20).

8.5.5.3   *Task completion time*

The participants received no instructions on task execution time or on the priority between precision and speed. Nevertheless, task completion time was tracked throughout the experiment. Time measurements were gathered for every single size estimation in all scenarios, stating when a square is displayed and finishing when verbally passing the size judgment to the study manager.

Participants showed a training curve throughout the 20 runs of each scenario (see Table 8.2). All in all, run 2 to 20, the median of scenario (T) was 5.063 ms, whereas the scenarios (FI) (9.959 ms) and (F) (8.429 ms) were slower. For all three scenarios the very first runs showed higher median values caused by non-existing training. A detailed graph on all median task completion times throughout all runs is given in Figure 8.21.

Table 8.2: Comparison of median task completion times

Scenario	Median of first run [s]	Median of run 2 - 20 [s]
Tablet (T)	10.84	5.06
Floor (F)	9.90	8.43
Floor&Interaction (FI)	16.23	9.96
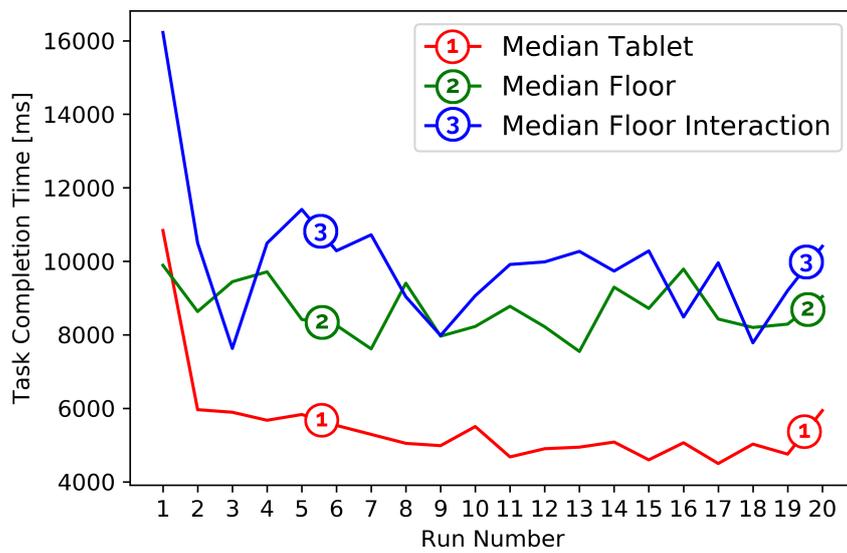
Figure 8.21: Median task completion time of all participants (N=22) for all 3 scenarios throughout the 20 runs
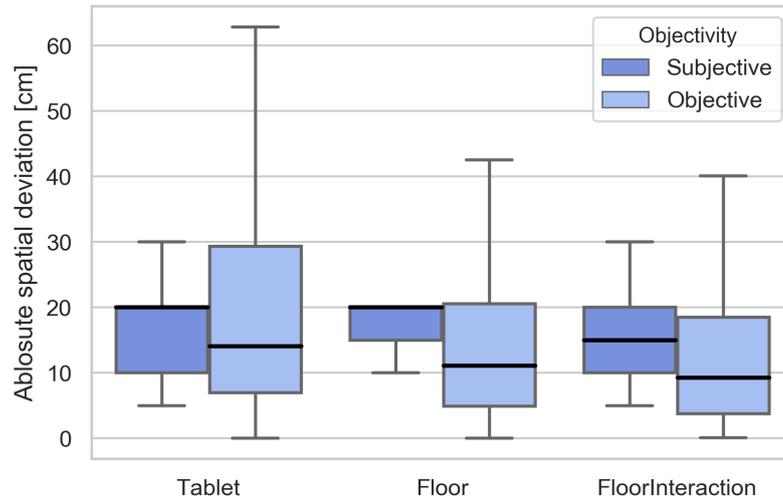
Figure 8.22: Comparison between perceived (subjective) absolute spatial deviation and objective absolute spatial deviation. N=22 for each scenario

8.5.6 *Questionnaire results*

After the experiment, all 22 participants filled out a questionnaire on their subjective perception. The non-standardized questionnaire compares the objective metrics to the participant's subjective perception.

8.5.6.1 *Task completion time*

"For this method, I was able to judge the sizes more quickly." The participants had to decide on each possible pairwise combination of all three scenarios: "(T) or (FI)", "(T) or (F)", "(F) or (FI)". Overall, the subjectively fastest scenario was (T). Comparing scenarios (F) and (FI) the results are equal (50% vs. 50%). Comparing both floor scenarios (F) and (FI) to the (T) scenario a subjective time benefit of (T) was reported 72.73% in favor of (T) compared to (FI) and 63% in favor of (T) compared to (F). The subjective questionnaire feedback matches the objectively measured times. 86.67% of the participants were really quicker when they were in favor of the (T) scenario in terms of task completion time. In contrast to that, only 7.14% of the participants in favor of (F) or (FI) scenarios were really quicker.

8.5.6.2 *Precision*

"Using this scenario, I'm able to assess the sizes more precisely". As for task completion time all pairwise combinations of scenarios were tested: (FI) was estimated to be the most precise scenario (46.97%) followed by (T) (31.82%) and (F) (21.21%). Interestingly, participants clearly preferred (FI) over (F) (86.36%), whereas when comparing (FI) to (T) and (F) to (T) there was no clear preference (50.00% and 54.55%

in favor of both floor scenarios). Comparing those subjective results with objective error metrics, there was a false impression for the subject's error estimation capability using (T) scenario. Only 28.57% objectively performed more precisely using (T) even though they were estimating this scenario as the most precise one. In contrast, 78.26% of the participants who were in favor of either (F) or (FI) scenarios also objectively performed better using these scenarios. Additionally participants reported on their absolute subjective size judgment error. In general, participants objectively performed better with a lower absolute median error than they subjectively expected it to be (positive values only) (see Figure 8.22). For (T) scenario the perceived median absolute error is 20.00 cm, whereas objective median error was 14.08 cm. The same was true for (F) (perceived 20.00 cm, objective 11.08 cm) and (FI) (perceived 15.00 cm, objective 9.25 cm)

8.5.6.3 *Personal preference*

"I personally prefer the following scenario": The highest ranked scenario was (FI) with 59.09%, followed by (T) (31.82%) and (F) (9.09%). Even though (T) was ranked second as a preferred scenario, participants who preferred this scenario never performed best (0/7) in terms of precision and most of them even performed worst (5/7).

Additionally, the questionnaire gathered free answer possibilities: Participants reported that when using (FI) they felt "more confident estimating sizes" (3x), "used natural walking" (1x) to estimate the absolute lengths and to change their "viewing perspective" (2x) so that the squares were "right in front of them" (1x). They reported having a better "spatial sense" (1x) and realism degree (2x). Additionally such a true-to-scale visualization was helpful. Participants who preferred the (T) scenario subjectively mentioned a better "overview" (3x) and better "comparison with ruler" (2x) due to the smaller display size and "higher resolution" (1x).

8.5.7 *Discussion*

The results of this study indicate that both absolute (true-to-scale) and relative-scale visualizations have advantages:

For absolute-scale visualizations, there is a significant change in size judgment accuracy between tablet and both floor scenarios (F) and (FI). Using the LED floor with true-to-scale visualization has a positive influence on the precision of size perception. These experimental results are in accordance with earlier findings by the authors (see Otto et al. [110]). There, cascadable room-scale projection systems are also used to realize industrial applications. In addition to LED-based and projection-based systems, more and more industrial application scenarios are realized using VR/AR interaction techniques. Using these HMDs lacks two main benefits compared to augmented floor

surfaces: First, HMDs are single user devices, whereas augmented floor surfaces can be utilized by groups of up to 30 people. Second, perceived spaces and sizes are frequently underestimated using VR HMDs (following Kelly et al. [279]), even though this effect becomes smaller with state-of-the-art headsets. In contrast, LED-floors do not show effects of over or under estimation in accordance with the results of this publication. Therefore, using true-to-scale visualization enables participants to judge sizes more accurately.

For relative-scale visualizations, task completion times tend to be lower. Overall, using the scenarios (F) and (FI) is slower than using (T). Even though lower task completion times could be a hindrance factor for other use cases, in automotive production validation task completion time is less important than high accuracy.

Woran liegt es?

Another interesting effect in human size judgments are rounding habits: All participants reported size judgments in a rounded form. Typical reports of size estimation granularity were $5\,cm$ (5/22), $10\,cm$ (16/22) and $25\,cm$ (1/22) steps. None of the participants gave sub-centimeter precision results. Therefore, rounding effects are still smaller than the perceived size judgment capability (compare Figure 8.22).

8.5.8 *Conclusion of size perception study*

This study provides in-depth insights into human size judgments and task completion times, as all of the applications mentioned above rely on accurate and precise user's size judgments, different modes of perception and faster task completion times.

When comparing true-to-scale (absolute) and relative scale visualizations, size judgment accuracy is better with absolute visualization scenarios (F) and (FI), whereas task completion time increases with those scenarios compared to the baseline scenario (T). In comparison with VR spatial estimations, where sizes are frequently underestimated, for true-to-scale floor visualizations no generalizable deviations could be revealed. Various use cases depend on reliable spatial estimations of humans such as collaborative production validation workshops.

8.6 VIRTUAL STENCIL STUDY USING AUGMENTED FLOOR SURFACES

As presented in Section 8.2.3, one potential use case of augmented floor surfaces in production planning are virtual stencils. Virtual cell layouts are models of real world workstations. In production preparation, virtual plannings must be physically built up in assembly or logistics at some point, even though virtual layouts already exist. For realization, cardboard engineering is a popular method to replicate cells as PMUs in order to simulate assembly or logistics processes.

To date, in PV workshops tablet computers are commonly used to display virtual layouts when physically realizing physical cardboard models. Therefore, this method is the baseline scenario. However, this method of comparing small-sized visualizations with physical arrangements is cumbersome, time-consuming and error-prone. In contrast, by using an augmented floor surface and visualizing virtual cell layouts in absolute scale, the overall setup can be utilized as a virtual stencil. Augmented floor surfaces are expected to simplify the physical realization process (see Figure 8.4). Using this latter method, PV workshops can potentially be completed in less time and with fewer errors.

8.6.1 *Study goal and predictions*

The following study aims to investigate four aspects of using an augmented LED floor as a virtual stencil:

1. **Task completion time:** Can PMU cell layouts be set up faster?

2. **Quality:** Can PMU cell layouts be executed at a higher quality with fewer errors?

3. **System Usability Score:** Are people satisfied working with the floor stencil system?

4. **Design space evaluation:** For which additional use cases can the augmented floor system also be effectively applied?

8.6.2 *Participants*

All subjects were production planning employees at Daimler AG and had prior knowledge of the PV process. They were selected randomly after taking part in a PV workshop. Each of the 16 subjects volunteered to take part. All were either production engineers, product developers, doctoral candidates or interns, ranging in age from 23 to 51 years (M = 29.7, SD = 6.5). There was no extra reward for taking part in this study. All subjects reported normal to corrected vision.

8.6.3 *Setup, stimuli and design*

As the study investigated the realization process from virtual cell layouts to physical mock-ups, two different scenarios were compared: Tablet-based scenario and LED-floor-based scenario. The setup comprised six physical objects and ten virtual cell layouts as pre-requisites. Figure 8.23 depicts the two different scenarios.
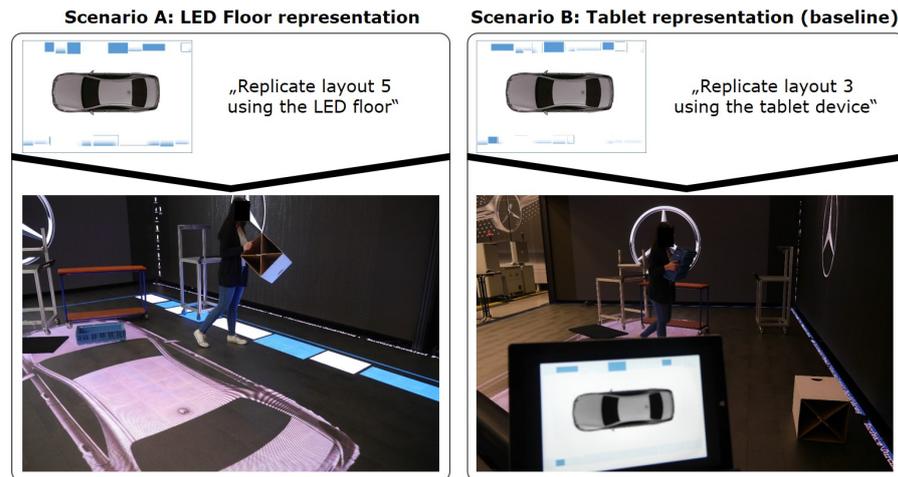
Figure 8.23: Example of both scenarios during each run. Blue squares indicate the physical object's destinations

- **"Scenario A"** utilized the augmented floor surface (technical specifications in Section 8.4.1). In this scenario, cell layouts were visualized true to scale on the 9m x 6m large LED floor.

- **"Scenario B"** utilized a Microsoft Surface tablet with 12.3" and showed image representations of the virtual cell layouts.

Both scenarios utilized the XnView image visualization tool to display virtual layouts. Overall, ten cell layouts were prepared using a walk path simulation tool and were converted to images for better reproducibility. Cell layout images were saved in the corresponding aspect ratio 3:2 as this corresponds to the native resolution of the LED floor (1728px x 1152px).

Each cell consisted of 15 to 20 objects, such as carrier representations or rack representations. These representations were visualized as white or blue rectangles. In each of the ten layouts, exactly four objects were "active boxes" which were colored blue instead of white and had to be physically built up during each session. The study participants were provided with six different objects as shown in Figure 8.24 with the following base area dimensions:

- Small material wagon (50cm x 50cm)

- Big material wagon (70cm x 60cm)

- One plate (104cm x 56cm)

- Box (59cm x 39cm)

- Cardboard cube (46.5cm x 34.5 cm)

- Table on wheels (118.4cm x 35.5cm)

Each cell layout used four out of six physical objects in a randomized manner. This is a representative mix of items in a physical cell.

Figure 8.24: Set of six physical objects each having different sizes. Four out of six objects are used during each run to replicate the virtual cell layouts

8.6.4 *Procedure*

In both scenarios the subjects had to physically build the virtual cell layout as quickly as possible with as few errors as possible.

The experimenter welcomed the subjects to the VMS. He provided an introduction to the goals and procedure of this experiment and clarified that the task completion time and error rate are measured during the experiment. Afterwards, subjective feedback was gathered using a standardized (SUS) and a non-standardized questionnaire.

The overall procedure took approximately 25 minutes including the completion of the questionnaire. Since a between-subject experimental setup has been chosen, each subject was randomly assigned either to group 1 (N=8) or group 2 (N=8). The experiment was carried out with 16 participants (N=16) in total. Five runs in each of the two scenarios A and B were carried out, each with all 16 subjects (see Table 8.3). Sequence effects were counterbalanced by alternating both scenarios. This led to a total of 160 datasets.

Table 8.3: Sequence of runs in the 2 groups for each participant. Counterbalanced scenarios A and B for group 1 and 2

Layout	L01	L02	L03	L04	L05	L06	L07	L08	L09	L10
Group 1 (N=8)	A	B	A	B	A	B	A	B	A	B
Group 2 (N=8)	B	A	B	A	B	A	B	A	B	A

For each run, the participant has been shown one out of ten different cell layouts with alternating starting scenarios. Before each run, all six physical objects were placed in the middle of the LED floor.

Scenario A: LED Floor

| - | Interpret layout | Choose proper object (4x) | Bring object to position (4x) | (Refine layout) | Finish |

Scenario B: Tablet (baseline)

| Walk to tablet device | Interpret layout | Choose proper object (4x) | Bring object to position (4x) | Refine layout | Finish |
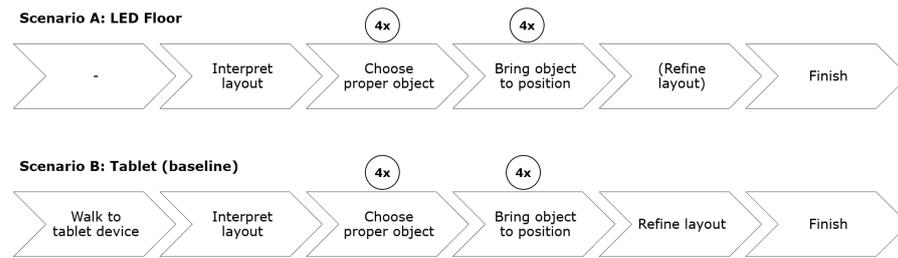
Figure 8.25: Typical sequence of tasks during each run for scenario A and B

The moment the cell layout was displayed to the subject, time measurement began.

For both scenarios, the participant analyzed the virtual cell layout, then had to pick the correct four out of six physical objects and had to bring each single object to the designated visualized spot. These six objects are depicted in Figure 8.24.

Each cell layout used a different subset of objects and were partly difficult to distinguish visually by its dimensions. The objects had to match the designated destination, represented by four blue boxes, as closely as possible. Errors were defined to be either the wrongly picked objects or geometrically misplaced destinations, which differed more than 50% outside of the virtual representation. Each run and thus time measurement ended when the subject verbally stated that he or she has finished the task.

Each subject completed 10 runs, 5 times scenario A and 5 times scenario B, in an alternating order. The schematic procedure of each scenario is depicted in Figure 8.25.

Scenario A: The participant has been shown the virtual cell layout with the respective objects in true-to-scale size on the LED floor. The subject must determine which physical objects have to be used, and the subject must place them correctly on the surface.

Scenario B: The LED floor only visualized the vehicle and the border of the entire cell as a rough landmark. The complete cell layout including all objects was displayed on the tablet in relative size. Thus, the size and geometric shape must be estimated using the tablet computer. After the end of each run, the experimenter revealed the solution on the LED floor to count the errors of wrongly picked or misplaced objects.

After finishing all ten runs, the subject completed a questionnaire. In this questionnaire, the VMS has been evaluated with respect to the usage as a virtual stencil using the "System Usability Scale" [28]. The System Usability Scale (SUS) [28] is a standardized questionnaire for measuring the usability of a product or system.

Participants were asked additional questions regarding the usability of the LED floor and the rating of additional application scenarios, suitable for the augmented floor surface. For these additional usability questions a five point Likert scale has been used.
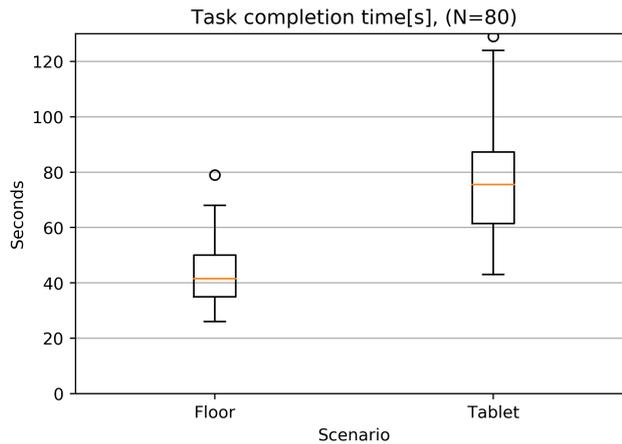
Figure 8.26: Comparison of the two different scenarios in a box plot

8.6.5 *Results*

Task completion time and number of errors for each scenario are shown in Table 8.4. This task completion time comprises a full physical replication of the virtual cell layout. If the subjects select a wrong object or if the placement of the object differs by more than 50% from the designated position, one error is counted per misplaced object. Table 8.4 depicts the overall mean numerical results for scenario A (M=42.95, STD=10.93) and scenario B (M=78.21, STD=23.09). Hence, scenario A using the LED floor revealed no errors (0) whereas scenario B counted 75 errors.

According to Table 8.4, Figure 8.26 depicts a box plot of the respective data differentiated for both scenarios A and B.

Table 8.4: Results on task completion time and absolute number of errors for each scenario

	Taks completion time		Errors
	M [s]	**STD[s]**	**# of errors**
Scenario A: LED Floor	42.95	10.93	0
Scenario B: Tablet	78.21	23.09	75

Hence, complexity is considered to be constant throughout all cell layouts. Mean complexity similarity is revealed by between-subject evaluation groups. Group 1 (N=8) finished scenario A runs (Layouts 1, 3, 5, 7 and 9) in 43.4 seconds, compared to group 2 (N=8) who finished the complementary layouts (Layouts 2, 4, 6, 8 and 10) in 42.5 seconds. Same holds for tablet scenarios. Group 1 (N=8) finished scenario B runs (Layouts 2,4,6,8 and 10) in 76.1 seconds, compared to group 2 (N=8), who finished the complementary layouts (Layouts 1,3,5,7 and 9) in 80.4 seconds (compare Table 8.3 for group sequence
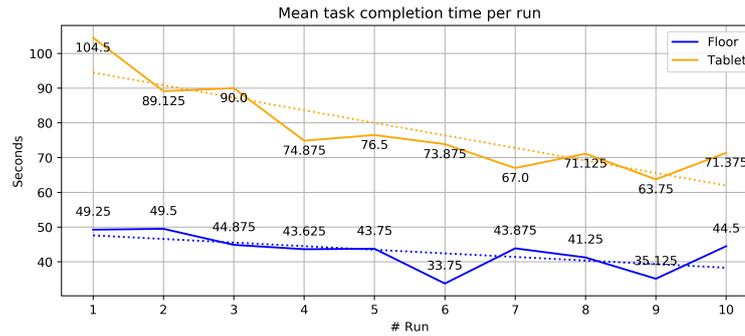
Figure 8.27: Comparison of mean task completion time values [s] over the different runs

assignment). Therefore, the difficulty of each layout roughly can be assumed to be similar since the amount of active objects is kept the same for all layouts.

Learning effects throughout all ten runs can be defined. By randomizing the starting condition (layouts), each participant carries out ten runs in a different sequence of layouts. Figure 8.27 shows the resulting task completion time sorted over the number of each run.

For instance, run #1 shows all mean results of all participants, regardless of which layout they had to complete during that run. All ten runs are carried out eight times (N=8) by different participants. In scenario B (Tablet), the first run took in mean 104.5 seconds, the ninth run only took 63.75 seconds. For scenario A (Floor), the learning rate was smaller as the first run took an average of 49.25 seconds, whereas the ninth run only took 35.13 seconds.

An independent-samples t-test was conducted to compare the two scenarios with each other. Even though there are only eight samples per scenario and run, all independent t-test reported significance levels below $t(8)=-2.923$ $p=0.011$ (second run). Therefore, the visualization methods of the two different scenarios do not have the same expected mean values.

Results of the "System Usability Scale" questionnaire [28] report an overall satisfaction with the usability of the LED floor system. The system scored an average of 86.56 points (SD=7.17) in SUS. 86.21% of the subjects "fully agree" that the system was "easy to use" and by tendency all of them agreed.

In addition to the standardized SUS questionnaire, a subjective questionnaire is assessed containing 4 questions. It evaluates the personal opinion of the subjects regarding the usage of the LED floor compared to conventional desktop-sized visualization methods. All subjects agreed on the statement that they were subjectively "quicker" (N=16), would "prefer to work with this method" (N=16), and agreed that "this method helps to better understand the real assembly processes in workshops" (N=16) than current state-of-the-art methods.

Only 1 out of 12 subjects found the "LED floor method is less suitable in workshop situations", compared to small-sized visualization devices (N=12).
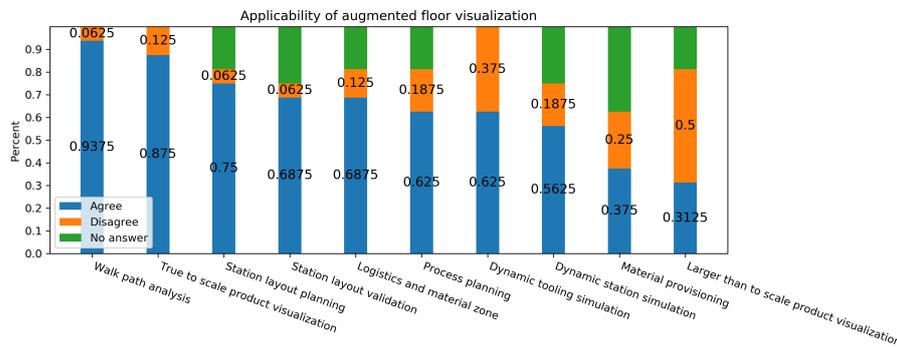
Figure 8.28: Evaluation of potential application areas of the LED floor visualization and interaction

In addition, the subjects ranked potential use cases of the LED floor with regard to pre-suggested applications. Subjects were asked to judge which potential application scenario would be beneficial for their work. The results of this ranking are shown in Figure 8.28.

8.6.6 *Discussion*

As already depicted in Figure 8.26, results show that there is an effect on task completion time using either scenario A or B. Replicating virtual cell layouts using the LED floor outperformed scenario B by the factor of almost two in terms of task completion time (42.95 seconds in scenario A compared to 78.21 seconds in scenario B). Due to the counterbalanced experiment design, no difference between the groups performances could be revealed. The linear regression in Figure 8.27 shows a learning curve in task completion time over the ten runs of each subject, especially in scenario B (tablet).

Errors are induced by choosing the wrong objects or by choosing the wrong place. Results show fewer errors are made in scenario A compared to B. No mistakes (0) were made using the LED floor method, whereas 75 mistakes were made using the tablet device. Verbal feedback by the subjects revealed that choosing the proper objects is error-prone using the compressed visualization scale of the tablet visualization. For example, subjects pick the box instead of the material wagon because the absolute size of the blue rectangle is misinterpreted. Additionally, even when participants picked the correct object, they oftentimes placed it in positions with a less than 50% overlap of the designated place. Due to the direct spatial coupling of the augmented floor surface, there was no room for misinterpretation or misplacement of objects, and subjects were able to directly check their

results and correct them, even if the objects were initially incorrectly assigned. So overall, scenario A is less error-prone.

In addition to the quantitative results, subjects completed a questionnaire for the **qualitative data** collection. First, the method using an LED floor is evaluated with respect to its usability by using the SUS questionnaire. Subjects rated all ten SUS statements, predominantly regarding the LED floor as positive. The evaluation of the SUS test showed a total score of 86.56 points. According to the inventor of the SUS score, Brooke [164], SUS scores above 85 points indicate an excellently perceived usability [164].

Insights into **further industrial application areas** for the interactive augmented floor surfaces are provided in Figure 8.28 and applications were ranked by the subjects. Most strikingly, "walk path analysis" was rated best with a 93.75% agreement score. Also, "true to scale product visualization" (87.5%) and "cell layout planning" (75%) received high interest for potential use. Average results for potential application areas were given for "logistics and material zone" (68.75%), "cell layout validation" (68.75%), process planning (62.50%) and "dynamic tooling simulation" (62.5%). "Material provisioning" (31.25%) and "larger than life product visualization" (37.5%) were judged with the lowest acceptance rates.

Due to the small number of participants, these figures do not claim quantitative reliability but an overall prioritized list of potential application areas for future realization for the VMS. Even though the questionnaire calls for additional ideas of application areas, none of the participants added to the pre-proposed list of application areas. Some application scenarios were not demoed to the participants, which might influence some low agreement scores for "larger than life product visualization" or "route planning." Additionally, participants came from assembly planning, not logistics planning departments. Therefore, "logistics route planning" is not a intrinsic use case for this population, even though it potentially could be a striking use case for other subjects.

8.7   OVERALL CONCLUSION AND OUTLOOK

Multiple research question have been addressed. First, an in-depth review of current LHRD research and augmented floor surfaces was given and set into context with industrial planning and validation processes. Six novel application scenarios in the domain of virtual assembly validation were presented:

- Interactive optimization of walk paths

- Layout assessments

- Virtual to scale stencil

- Size perception and engagement

- Self-navigation in virtual space

- Virtual travels with interactive maps

Derived out of this, **two industrial setups with multi-display environments** were developed and implemented. As a proof of concept, the first implementation in the context of the VMS was a scalable floor projection system which proved to work reliably in PV workshops, evaluated by using questionnaires. This revealed benefits of true to scale visualization but also found some drawbacks in implementation, such as lacking brightness or self-occlusion by shadowing effects. Overall, even this scalable, mobile implementation found huge acceptance during PV workshops. With the prior knowledge of the proof of concept implementation, the LED-based VMS implementation including an augmented floor surface, was built and presented for usage in PV workshops. This implementation was evaluated quantitatively and qualitatively with two experiments:

A **basic research study on size perception** of users showed that subjects were able to estimate sizes of virtual objects more accurately by using true-to-scale visualization compared to tablet-sized relative scale visualizations. Nevertheless, task completion time in this abstract experiment increased compared to tablet-based scenarios.

In a second experiment, a **real-life application study** of realizing physical manufacturing cell layouts was evaluated, comparing tablet-based baseline interaction scenarios with an augmented floor surface interaction scenario. Results showed that the task completion time of replicating virtual cell layouts declined while the quality of results increased. Subjects appreciated the novel true-to-scale interaction cycle of physical and virtual cell layouts, so they always had a strong mental coupling by true-to-scale visualization of objects.

Overall, both experiments indicate that true-to-scale visualization helps estimate sizes more accurately and with fewer errors. LHRD multi-display environments help achieve the VMS goals in PV workshops, since large amounts of data have to be visualized in parallel. According to Andrews et al. [251], this also has an impact on the quality of work results. In accordance with the VMS's key properties, these results on LHRDs, size judgment capabilities, task completion time and quality multiple aspects have been achieved:

- **Multi-user support** is given, since the multi display environment can be used with up to 30 people

- **True to scale** is enabled by the presented system including an isometrically registered interaction cycle.

- **LHRDs** are an enabling technology for asymmetric output as not all PV workshop participants are intended to have personal views using HMDs.

- **Isometric registration** of true-to-scale visualization allows geometrically close coupling of PMUs and DMUs.

- **Entry barriers** are reduced since no mental transfer must be made when true-to-scale visualization helps show massive amounts of data in parallel as well as geometric information true to scale.

In future research, using the presented LHRD apparatus, further interaction research will be carried out since the LED-based augmented floor surface contains optical approach sensors for direct touch capabilities. Footstep recognition, as depicted in Figure 8.2, can thus be realized (see also Agethen et al. [68] and compare Schmidt et al. [265]). Such reconstruction of human walk paths can be used in some application areas to partly replace complex 3D motion capture systems.

9

PROJECTIVE SPATIAL AUGMENTED REALITY

Augmented reality for manufacturing environments has been an important focus of research activities over the last decades. For instance it has been applied to concurrent engineering [41], training [282], factory planning [283] and maintenance [284]. Industrial Augmented Reality (IAR) is defined as the application of AR for the support of industrial processes (see Fite-Georgel [285]). In accordance with Fite-Georgel, typical IAR application domains are product design, manufacturing (assembly guidance, training), commissioning, inspection and maintenance and decommissioning. However, even though Bimber and Raskar state in 2005 that "industrial use of augmented reality is [...] on the rise" [286], in 2018 Uva et al. summarize that "Augmented Reality solutions are still struggling to reach the factories. This issue is mainly due to display solutions that still do not fulfill strict industrial constraints on ergonomics, color coding, training of operators, and the reliability of proposed solutions" [287]. At the moment, most of the solutions proposed for augmented reality use hand-held or head-worn devices, both having disadvantages in industrial environments and constraints in real life application. That is why this chapter focuses on a different non-body-attached variant of augmented reality, namely projective (SAR). "Projection Augmented Reality features one or more optical devices (projectors) that project a beam of light onto a specially designed work surface and in some cases directly at the parts a user is working on. This provides immediate guidance for tasks and reduces the need to interrupt workflows to obtain information elsewhere" [288].

PV is about assessing the details of the upcoming PPR. The intention is to integrate PMUs, such as individual parts or even the entire car, in the virtual assessment workflow. These physical objects are intended to be superimposed with spatially aligned digital information in PV workshops. Therefore, the objective is to integrate DMU and PMUs in the validation process by using spatial augmented reality. This is in line with the VMS framework's key property to achieve a closer "coupling between PMU & DMU" (see Section 5.2).

> **Related publications and research project**
>
> Multiple parts of this chapter's research have been supported by internal and external partners within the BMBF-funded ARVIDA research project. The company EXTEND3D has supported this work with their integrated projective SAR systems called Werklicht Pro and Werklicht Video, their proprietary SAR software, as well as Fraunhofer IGD Darmstadt with their markerless CAD object tracking software and Daimler Protics with their integration in the veo validation software environment. These studies were presented in mutual publications in the Springer book "Web-basierte Anwendungen Virtueller Techniken" [132]. However, the studies described in the following chapter were designed and carried out for this doctoral thesis.

This chapter is structured as follows: First, the requirements for IAR in PV workshops and ways to integrate PMUs are described. Subsequently, a comprehesnive literature review is presented on AR and SAR with a focus on all-purpose driven manufacturing environments. The implementation of a working demonstrator is described in the context of employing SAR technology within the VMS framework and combining it with the assembly simulation software discussed in Section 6.3. A brief comparison between OST-HMD with projective SAR is given. Subsequently, insights, two studies are presented: The first study gives insights into the design space of SAR in comparison with usage with a head-worn augmented reality device for manual assembly tasks. In the second study, SAR is evaluated with respect to its ability to support collaboration performance in workshops with distributer knowledge in terms of task completion time for typical localization tasks.

9.1   REQUIREMENTS FOR PMU-BASED PV WORKSHOPS

The overall goal of PV workshops is to validate and optimize the entire production system by simulating the build-up of the product. This can be conducted in either the physical or the virtual domain. Blending both domains, AR represents an alternative user interface, which allows the addition of 3D-registered digital content to the physical environment in real time [286]. By blending both domains using IAR, production engineers aim to optimize the overall workload during PV workshops to reduce potential error sources, to increase the resulting quality, to save time and possibly enhance motivation (see Neumann and Majoros [289]).

In the literature, Neumann and Majoros present an overview of cognitive, performance and system issues for AR applications in manufacuring [289] and cluster AR activities as a list of typical activi-

ties. These clusters of manufacturing tasks also apply for PV workshops. Therefore, Table 9.1 extends the original content of Neumann and Majoros with exemplary application scenarios in PV workshops. Two types of activities are distinguished: Informational activities and workpiece activities. "The distinctions implied in the table are often blurred, but in general, document-related activities tend to be cognitive, workpiece activities tend to be kinesthetic and psychomotor, and both involve visual and auditory factors" [289].

Table 9.1: Classes of AR supportable activities in manufacturing tasks following Neumann and Majoros [289]

Clusters of AR activities	Exemplary use cases of AR within PV workshops
Informational tasks	
Direct attention to storage medium	Assessment of logistics zone (Pick by Beamer)
Read, comprehend, interpret, calculate	Predetermine motion times, Process quality, Projection of manufacturing sequence
Understand speech	Ambiguities in commands of PV workshop colleagues
Form hypotheses	Reachability assumptions
Transpose information from documents to workpiece	Derive actions from assembly task descriptions
Workpiece tasks	
Direct attention to workpiece	localization of points (see also Marchi et al. [290])
Inspect, discriminate, compare, select, align	PMU/DMU quality align Bolts
Orient to sound, interpret sound	N/A
Adjust or actuate devices, detect movement	Variant handling
Manipulate devices	Assembly paths

According to these clusters proposed by Neumann and Majoros as shown in Table 9.1, they are used to evaluate the design space of PV workshop activities with respect to the usage of SAR in the following evaluations. The results of the subsequent design space evaluation will give insights into applicable types of technological approaches for specific validation tasks: projective SAR or AR.
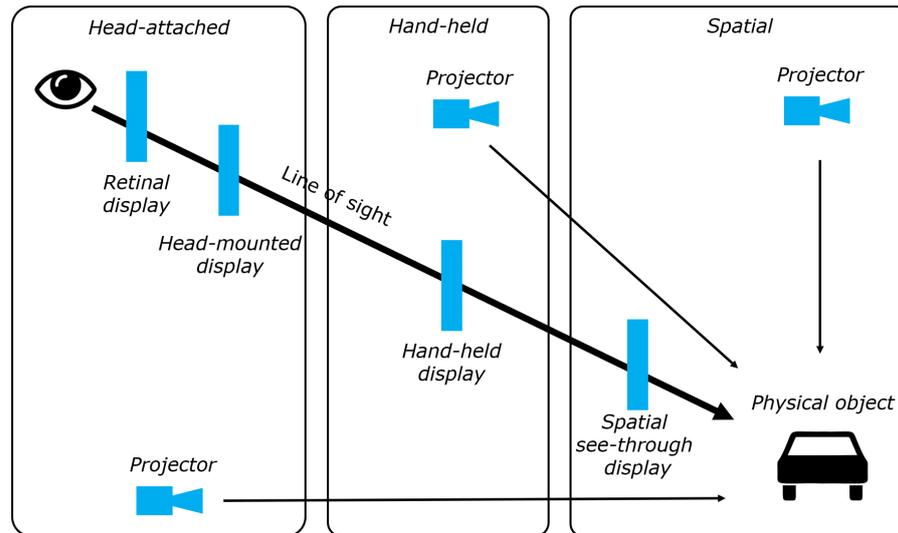
Figure 9.1: Visual display techniques and positioning adapted from Bimber and Raskar [286]

9.2 PROPERTIES OF SPATIAL AUGMENTED REALITY

Spatial augmented reality is a special AR approach that uses different display technologies for superimposing objects in reality and can "pave the way towards flexibility, higher efficiency, and new applications" [286, p. 1]. According to Bimber et al. [286, p. 7], "spatial displays detach the display technology from the user and integrate it into the environment." Bimber et al. call spatial augmented reality "new display paradigms exploit large spatially-aligned optical elements, such as mirror beam combiners, transparent screens, or holograms, as well as video projectors." It is a method of conveying digital information to workshop participants within a stationary context. "Spatial [...] displays overlay the real environment with computer graphics in such a way that the graphical images and the image of the real environment are visible at the same time. In contrast to head-attached or body-attached optical see-through displays, spatial displays generate images that are aligned with the physical environment. They do not follow the users' movements but rather support moving around them" [286]. "Target objects and users can move around in the environment, but the zone in which AR experiences take place is limited to the fields of view of both the fixed projector and supporting camera for tracking" [288].

As depicted in Figure 9.1, Bimber and Raskar show how projective SAR display techniques differ from head-attached and hand-held devices. They project digital information onto the real objects and thus superimpose the information onto the viewer. This information is view independent for multi-user setups when the virtual surface matches the physical surface - otherwise parallax effects are generated. Regarding the VMS framework hardware arrangement, an al-

most stationary setup applies, as the required infrastructure is available and projector positions do not have to vary heavily - in contrast to mobile SAR setups, which are discussed in the literature as well (see Kruijff [291]). The following opportunities and challenges of projected SAR can be found in the literature:

Opportunities of stationary SAR:

- Bright and full contrast projection (following Kruijff [291])

- High fidelity (following Kruijff [291])

- Large FoV of projected display (following Kruijff [291])

- Potentially single disparity plane (following Kruijff [291])

- Reduces or eliminates the need for computer monitors and screens, as the instructions [288] appear directly in the task space

- Reduces users' cognitive load when following work instructions due to the fact that there is no need for "attention switching" between work instructions and the task at hand

- Integrates into manual workflows by promoting a "no faults forward" policy to ensure and confirm correct execution of the preceding step

- Provides feedback on completed tasks for process improvement, traceability and unique digital IDs for build cycles

Challenges for stationary SAR:

- Surface-based distortions (following Kruijff [291]): "not possible to display mid-air virtual objects, but only 2D objects on a physical surface. Furthermore, surface-based distortions remain an issue, and the user could misunderstand the projected instructions" [287]. Surface-based distortions can lead to worse legibility

- No mid-air visualization capabilities (following Uva et al. [287])

- Shadows and clipping (following Krevelen et al. [292])

- Registration

- Ambiguities

- Multi-user support

9.3 RELATED WORK ON SPATIAL AUGMENTED REALITY IN MANUFACTURING

Research on augmented reality has received a great deal of attention over the past two decades. As Wang et al. state in their research review on AR in assembly [293], between 1990 and 2015 they evaluated 304 papers on this topic. This review gives a concise overview of the technical features, characteristics and wide range of applications for AR-based assembly systems. Wang et al. clustered the application papers in three main and twelve sub-categories:

- AR assembly guidance

- AR assembly training

- AR assembly simulation, design and planning

The VMS framework's research in the context of PV workshops contributes to the last category "AR assembly simulation, design and planning." Additionally, in the past few years research has focused on industrial augmented reality applications. Fraga-Lamas present a review on industrial AR systems for the Industry 4.0 Shipyard [294]. Their holistic review is not only limited to shipyard applications but also includes an overview on state-of-the-art AR technologies and applications in manufacturing including SAR applications. More generally speaking, Nee et al. present a holistic overview of general AR applications in 2012 [295]. More specifically, other studies on augmented reality application scenarios in automotive industry are conducted by Geissel [43] or Bliese [296].

With respect to spatial augmented reality, early studies refer to this domain as projective augmented reality. One of the early SAR examples of superimposing a calculator on a regular desk was presented by Wellner in 1991 [297]. Later studies by Wellner in 1994 showed a desk for virtual interaction as an "augmented environment for paper" [298]. Since then multiple use cases using SAR interaction technology have been presented, such as product design [299], interaction design [299], entertainment [111], museums [300], [301] and industrial use cases such as manual workstations [287], maintenance and shop floor assistance [302]. Volkswagen presents a whitepaper with use cases for Spatial Augmented Reality [299] in design and manufacturing on their website. Dedicated mobile projective SAR research is carried out in the MoSART projects by Cortes et al. [303] and AMP-D by Winkler et al. [304]. Both concepts share body-worn camera-projector systems to superimpose digital content onto objects in the physical environment. Even though there are multiple publications on SAR in manufacturing, the vast majority of papers present body-worn or hand-held AR applications due to their wide availability. The most relevant papers for projective SAR in the context of the VMS are presented below:

As early as 2003, Piper and Ishii proposed CADcast [305] - a "Method for Projecting Spatially Referenced Procedural Instructions." The system projects assembly instructions onto a workbench. They showed in a brief evaluation that task completion time is reduced by using SAR compared to paper-based methods.

Benko et al. present FoveAR [108], a combination of optical see-through displays with spatial augmented reality projections and ask the design space question about which content to display where. Using their prototype system FoveAR, they present an analysis of the interactive design space. In their description of their four exemplary use cases, they also discuss the advantages of this combination for displaying public and private content. Subsequently, the results of the design space evaluation are evaluated in the context of this publication.

Uva et al. present a journal paper on the effectiveness of SAR in manual workstations [287]. In this evaluation on an SAR workbench, assembly tasks on a motor cycle engine are carried out. For this study, Uva et al. use an augmented workbench discussed by the same authors in 2016 [306]. This workbench tracks and augments manual assembly activities. The authors investigate whether SAR is an effective support for the operators, while projecting technical instructions. Therefore task performance, task completion time and error rates have been measured. They come to the conclusion that "SAR presentation mode is significantly better than paper in completion times and error rates" [287] while using a manual assembly station. In particular, work tasks with high complexities show significantly lower completion times as there is a "substantial reduction in complexity" [287]. In contrast to the findings of Uva et al., the VMS framework does not limit the "main advantage of SAR" to "the reduction of error rates than to completion times" [287].

Sand et al. present a digital guidance and picking assembly assistance system supporting assembly workers at their workstations [307]. They provide operational information to untrained workers on assembling products without any previous knowledge. Even though there is no formal evaluation, they show that using SAR enables workers to perform the work tasks and that "projection-based instructions are able to assemble products faster and with a lower error rate, compared to the system based on smart glasses" [307].

Doshi et al. propose a projection-based SAR system to improve manual spot-welding precision and accuracy for automotive manufacturing [308]. They present a video-projection-based SAR system and evaluate this in a real automotive production plant. Overall, by using their system they determined a 52% reduction in standard deviations for manual spot-weld placements. They propose a set of best practice visual cues for spot-welding, such as circles, crosshairs and arrows in

different colors, which has also influenced the design space evalua-tion for the VMS.

Schwerdtfeger et al. developed an SAR system that employs laser projection systems [302] for quality assurance and maintenance. They provide insights into geometric, optical and implementation issues of laser-based projective SAR systems. In addition, Schwerdtfeger et al. discuss occlusion issues caused by the users and the objects them-selves. They conclude that their body-attached projector setup has to be faster and more accurate. They thus present a hybrid solution between fully mobile and completely stationary setups: A tripod-mounted SAR projector. This implementation later became commer-cially available as a derivative of this paper and is called "Werklicht." It is applied in the following design space evaluation and studies for the VMS framework.

Multiple publications are available on the overall performance us-ing SAR: Funk et al. concluded that using in-situ projection at assem-bly workstations results in a significant reduction in task completion times: "results show that assembling parts is significantly faster us-ing in-situ projection and locating positions is significantly slower using HMDs." [309]. Bimber and Raskas substantiate these findings with the "high level of consistency between real and virtual environ-ments." [286, p. 8], since "cognitive distance is high if users switched from the information space (the manual) to physical space." Kruijff et al. support these findings by stating that omitting SAR leads to the effect that users have to "understand the relationship between what is seen directly in the real world and what is shown on the screen when comparing both, which may require difficult mental rotation and scal-ing" [291]. Therefore, SAR is integrated in the VMS framework's key properties. Collaborative tasks in PV workshops are often coupled with a geometrical localization task, when one or multiple partici-pants want to show others certain product variants, mounting points, welding spots, mass points or similar geometric references.

Even though not all potential benefits are evaluated explicitly for the VMS framework, literature can be summarized as follows:

- Lower error rates (compare Uva et. al [287])

- Reduced task completion time (compare Funk et al. [309] and Uva et. al [287])

- Context-sensitive information: The reduction of mental work-load associated with a task because only the task-relevant infor-mation is displayed (compare Kim and Dey [310])

- Cognitive distance is high if users switch from information space (the manual) to physical space. Reduction of mental workload associated with these textual instructions is high when users try to memorize the sequence (compare Kim and Dey [310])
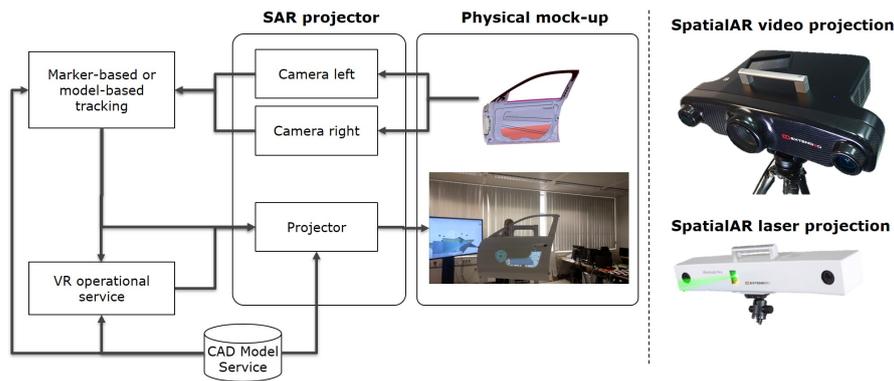
Figure 9.2: Left: Block diagram of SAR system applied in VMS; Right: Two different types of projective SAR: laser-based and video-based projective SAR systems

- The reduction of cognitive distance by minimizing the cognitive load in translating between virtual information spaces and the real world because the information space and the physical space coincide (compare Kim and Dey [310])

- Several publications discuss the effects of higher excitement [307] and motivation [289] when using SAR. Neumann and Majoros state "that AR is a candidate to produce better adherence to correct procedures by virtue of increasing motivation" [289]. For the VMS, this effect is not evaluated since long-term motivational effects vs. initial usage motivation have not been evaluated outside laboratory environments

- Reduction of ambiguities in human speech

Today, several companies offer commercially available projective SAR systems for manufacturing industry, such as Arkite (product: HMI), OPS Solutions (product: Lightguidesys) and EXTEND3D (product family: Werklicht).

9.4 SAR SYSTEM ARCHITECTURE

For all succeeding studies the same prototype setup is applied. The system consists of tracking hardware, projection hardware and special SAR software in order to realize the desired spatial augmentation of PMUs in the VMS. Two commercially available projective SAR systems have been integrated in the prototype system: Werklicht Video and Werklicht Pro manufactured by EXTEND3D (see Figure 9.2).

The system consists of multiple components:

- **Physical mock-ups** are superimposed with spatially aligned virtual contents. For the following studies, a physical door, a body-in-white and a main girder are used. The door and the body-in-

white have been primed and coated with white paint. For registration, all parts are equipped with photogrammetric markers for high-precision tracking.

- Both applied **SAR projection systems** are equipped with high resolution, industrial cameras, so that the markers on the PMU can be tracked properly. EXTEND3D's proprietary software offers multiple variants of tracking algorithms: Marker based tracking algorithms rely on photogrammetric marker detection. Using two cameras, the minimum amount of markers needed for registration is reduced to three, whereas having more markers within the FoV of the stereo-camera arrangement makes tracking results more reliable via bundle adjustment. For marker-less object tracking, Fraunhofer IGD (see Wuest et al. and Bockholt et al. [92], [311], [312]) technology is applied. Using this, SAR projection system calculates the relative pose of the PMU in real world coordinate frame. Real-time updates of the tracked pose are used to render the superimposed contents on the fly.

- **AR/VR operational service** is called veo. The features of the assembly simulation software are described in Section 6.3. The validation simulation software uses standardized ARVIDA RESTful services (see [132]) to exchange data with the EXTEND3D projection service. The PMU reference markers are aligned with the DMU coordiantes, so that the virtual and physical parts spatially match. In this way, the assembly simulation software is able to superimpose virtual CAD contents onto the PMU. The current assembly state can be projected onto the PMU. In PV workshops, the operator chooses which product parts of the workstation under assessment are superimposed and in what manner (wireframe, colored, outlines, etc.)

- The **CAD model service** is a centralized web-based file storage containing all CAD geometry in JT file format. Both the AR/VR operational service and the proprietary SAR projector software temporarily download the contents from the respective file storage.

All in all, the presented prototype system allows for real-time augmentation of physical objects in PV workshops. Both subsequent studies use this system architecture.

9.5   COMPARISON OF OST-HMD VS. PROJECTIVE SAR

Projection SAR and OST-HMD both have vastly differing properties. Current state-of-the-art AR HMD still struggle with limited resolution, excess of weight limited field of view, and a limited and fixed focal depth as well as bad ergonomics (see van Krevelen et al. [292]).

Figure 9.3: Application related comparison of SAR vs. OST-HMD

Figure 9.3 depicts a comparison of these two hardware classes. Even though Microsoft Hololens made considerable progress in terms of tracking, multiple practical aspects of using OST-HMDs are still present. When proposing 'FoveAR', Benko et al. [108] also see the limitations of both (OST-HMD and SAR) technologies and therefore combine both to form a prototype system which is "equally capable of providing spatially-registered view-dependent AR experiences as the near-eye displays." Hence they want to overcome the limitations on FoV and brightness. They also offer new benefits to the viewer, such as "private stereoscopic views, per-pixel controlled ambient lighting, surface shading effects not affected by head tracking lag or mismatch" [108].

Table 9.2 compares the most relevant technical properties of the OST-HMD hardware Microsoft Hololens with both applied projection-based SAR systems (Werklicht Pro and Werklicht Video)

For use in the VMS framework, the following properties are most relevant:

- **Body-attached device:** Whereas an OST-HMD must be worn at all times, the SAR devices are mounted on a tripod semi-statically. Therefore, no ergonomic issues and discomfort apply for SAR. Working with HMDs for an eight-hour shift in the VMS, acceptance is limited for PV workshops. "Industrial operators have to wear the HWD for long sessions and, for this main reason, industries do not well accept AR with these displays." [287]

- **Shared vs. private content:** OST-HMDs offer the possibility to display private content, so that other PV workshop participants cannot see it. Since PV workshops do not have to protect confidential information among workshop participants, only context sensitivity would be a reason for using such a feature. SAR projects all contents onto the PMU and therefore is public content for all users. Information that does not have to be shared with others is displayed on conventional notebooks respectively.

- **Amount of devices:** PV workshops can consist of up to 30 persons. If each individual uses OST-HMD, a massive amount of hardware has to be in place. This involves additional costs and

Table 9.2: Comparison of OST-HMD Hololens vs. Werklicht Pro vs. Werklicht Video

	Hololens [313]	**Werklicht Pro**	**Werklicht Video**
Visualization type	Display	XY Laser	Video projection
Resolution	2.3M total light points	Infinite	Up to 4k projector
Body-attached weight	579g	0g	0g
Horiz. FoV	<40°	60°	60°
Battery powered	Yes	No	No
User's Viewpoint	View-dependent	View-independent	View-independent
Visualization	Stereoscopic	Monoscopic	Monoscopic
AR devices mobility	Dynamic	Semi-static tripod	Semi-static tripod
Tracking	Inside-out 3D camera	Inside-out stereo camera	Inside-out stereo camera
Registration	Room Measurement	Marker-based or markerless	Marker-based or markerless
Arrangement	Head-worn	Spatial	Spatial
Users	Single user	Multi-user	Multi-user
Information amount	Limited by resolution	Approx. 40 symbols 30Hz	Limited by resolution
Disparity effects	No	Yes	Yes
Personal / Shared	Personal	Shared	Shared
Augmented object	Near-eye display	Directly on PMU	Directly on PMU
Optical occlusions	No	Yes	Yes
Ambiguities	Yes	Yes	Yes
Initial Costs	Low	Medium	Medium

logistics and maintenance overhead (loading, recharging, calibration, wiping, hygiene aspects, etc.)

- **View-dependent visualization:** Since OST-HMDs are single-user devices, view-dependent rendering for AR is the main benefit. SAR can also be used for view-dependent rendering but is not activated in VMS. All workshop participants consume the superimposed information, so viewpoint-dependent rendering is deactivated. Nevertheless, superimposed data matches perfectly for all participants when the virtual data lies on the PMU projection surface. For SAR "this is particularly visible for virtual objects far away from a projection surface (i.e., with big disparity)." [108]. Especially for mid-air visualizations, SAR has limited usability.

- **Optical occlusions:** Since OST-HMD are near-eye displays, no optical occlusions between digital content and the line of sight applies. Line of sight towards the PMU can still be occluded. Following Schwerdtfeger et al. [302] for projection SAR, two reasons for occlusion can apply. Either the users occlude the line of projections and therefore cast shadows onto the PMU, or objects occlude the line of projection. For complex PMUs and with protruding parts and concave geometry, self-occlusions apply. Therefore the proper placement of the SAR unit is crucial but not a hindrance factor for PV workshops.

These differences show why the SAR interaction design space is chosen for evaluation in the following sections. Uva et al. summarize SAR as possibly being the "optimal solution for the visualization of both instructions and technical information directly on the industrial workbench" [287]. This assumption is evaluated with regard to its design space for assembly assessment scenarios in VMS and PV workshops.

9.6 DESIGN SPACE EXPLORATION OF SPATIAL AUGMENTED REALITY

As shown in the literature review, the design space of SAR has not been described completely for the manufacturing industry. This study gives insights into the applicability of SAR for product validation scenarios. In the design space study, the following virtual contents are superimposed onto the workstation of the PMU:

- 2D Symbols & indicators

- Spatially aligned CAD geometry

- Alphanumeric assembly task instructions and technical information
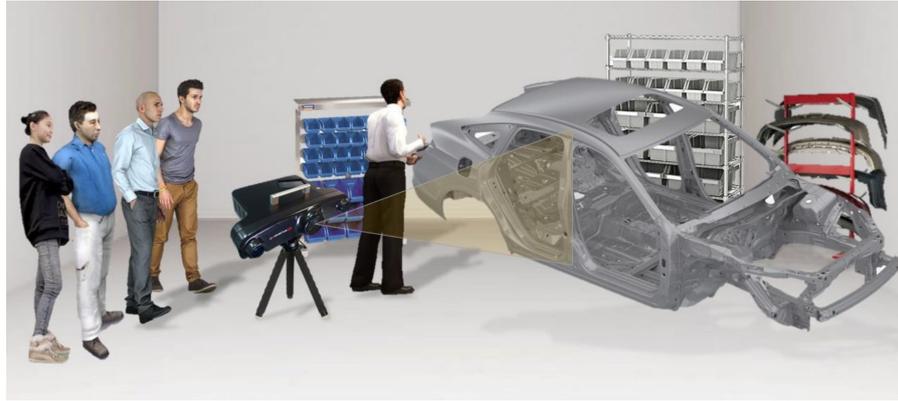
Figure 9.4: Schematic setup for design space evaluation for SAR applications

Typical product-related automotive PMUs are presented in Section 3.2.3. In the case of an automotive production workstation, additional types of PMUs are present:

- **Product:** All variants between one single physical assembly part and a complete set of parts of the entire product may be available. Physical parts are often outdated, since their digital representation is continuously optimized by the R&D department. Tolerances apply, especially in early physical prototypes.

- **Resources:** Multiple screw drivers, auxiliary materials, handling devices, automated devices, carriers, racks, boxes, etc. are available in PV workshops, since they are required to simulate assembly process.

9.6.1 *Setup, stimuli and design*

For this design space evaluation, a body-in-white, a car door and a main girder are available as physical mock-ups (see Figure 9.4). CAD data of all physical parts are available and used for registration. Spatial alignment of the PMUs to the SAR system is carried out with self-centered, magnetic markers, which are applied onto the surface of the PMUs. For all use cases, the SAR systems have been registered to the PMUs with sub-millimeter precision.

Derived from the SAR clusters in manufacturing as proposed by Neumann and Majoros [289], typical examples of PV workshops are projected using SAR. Results show the technical feasibility of the projection results using different specularities, colors, surfaces with different physical models and different contents (product, process, alphanumerics).

Figure 9.5: PMU quality inspection check by laser-based projective SAR to display potential offsets between DMU and PMU

9.6.2 *Results*

PMU/DMU quality checks: For inspection, discrimination, comparison and selection tasks, Figure 9.5 depicts the typical usage of SAR in the assessment of PMU quality. PMUs can be outdated, compared to DMUs. Figure 9.5 shows a weld spot of a chassis part. SAR projects a crosshair on the designated point on the PMU's surface. PV workshop participants can detect offsets between DMU and PMU. In addition to the crosshair, a numeric identification number is projected onto the PMU's surface in order to easily identify the spot's intended position.

Laser-based SAR is a good solution for such quality inspection tasks, since it offers high precision projection on a sub-millimeter basis and little information content, and the spatial information contained is located directly in the PMU's surface plane. Therefore, no disparity effects apply and all workshop participants have a view-independent benefit of the projected information.

Alignment tasks are often carried out in PV workshops, as parts must be mounted on a generic surface. Figure 9.6 shows an alignment task for bolts on a white car chassis. It is easy to see whether the projected DMU bolts are physically present and whether they are aligned correctly. If there is an offset between the projected bolts and the physical bolts, they can be physically altered or the deviation source between PMU and DMU can be found.

Similar to alignment tasks, **localization tasks** are very frequent in PMU-based PV workshops. One and the same assembly part can have multiple mounting locations in a product (see Figure 9.7). Therefore, SAR can help project the intended manufacturing sequence, and, even more importantly, the designated variant position within the product. For example, hundreds of plastic plugs must be inserted in the car body chassis, all having the same product IDs. To identify the correct position, SAR can visualize this position. Wireframes, outlines or col-
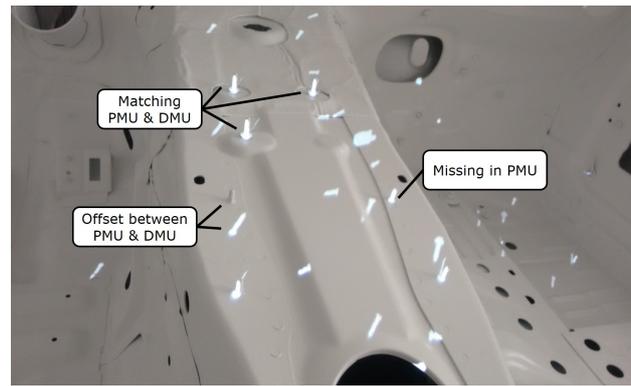
Figure 9.6: PMU alignment tasks using video-based projective SAR to display bolt positions on a body-in-white

Figure 9.7: Visual assembly support using video-based projective SAR to display assembly parts on a door

ored models of the designated assembly parts are projected onto the PMU's surface, using video-based projective SAR systems respectively.

Such **visual assembly support** projections work well in collaborative PV workshops, as long as the PMU's physical surface does not differ too much from the DMU's part depth. Figure 9.8 (left) shows an example of huge parallax effects, when projecting a bigger part onto the PMU. The DMU projection parts do not lie on the surface of the PMU. Therefore parallax effects from the participant's individual viewpoint are ambiguous. The entire off-surface projection contents only works for one specific view-point, namely close to the projection center (see Figure 9.8 middle). SAR video projection is also limited to the light power and reflection parameters of the objects. Even for white objects with lower projection intensities, contrast and thus readability are lowered drastically (see Figure 9.8 right). Comparing video-based projective SAR to laser-based projective SAR, the latter works better for generic colors and reflection parameters of PMU parts. As can be seen in Figure 9.5 and Figure 9.10, laser-based projections have better legibility even for black and partly reflective surfaces such as coated steel.
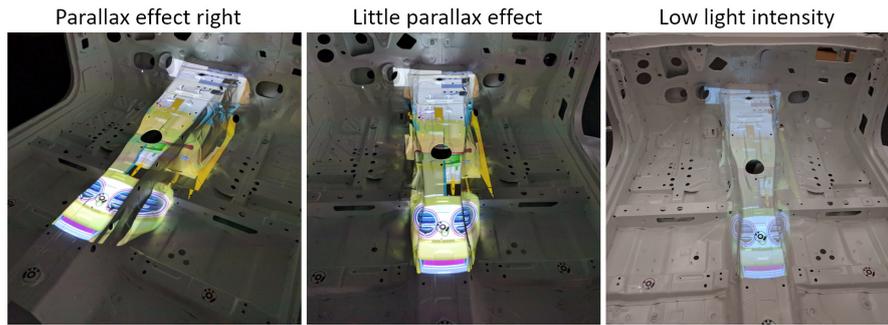
Figure 9.8: Visual assembly support with parallax effext (left), little parallax effect (middle) and low projection intensity (right) using video-based projective SAR. Projection a middle console onto a body-in-white

Figure 9.9: Alphanumeric assembly task information on generic PMUs using video-based projective SAR

Alphanumeric assembly instruction visualization is another possibility to support PV workshop participants who continuously have to check the process quality with multiple attributes. Projecting text for view-dependent AR is easier, since there is no need for a physical projection surface and no need to analyze the surface structure. Legibility of projected texts depends on many influencing parameters, such as environmental conditions, text style, material and shape of the target surface [314]. In contrast, for projective SAR the parallax effect causes disparities when projecting longer texts. Finding proper planar surfaces for legible alphanumeric texts is still in the research phase [314]. Figure 9.9 and Figure 9.10 show examples for projecting text onto generic non-planar surfaces. For instance, Figure 9.9 depicts three steps for determining the mounting location of the assembly parts by using crosshairs. This matches the use case visual assembly support. In addition, pictograms indicate the mounting directions. Textual information provides indications on which parts must be used.

Figure 9.10 depicts an example of laser-based projective SAR used for alphanumeric text projections. The XY-scanning unit of this laser projector only has a limited scanning frequency. The fewer symbols have to be projected, the better the projection quality is. In addition, the example also shows that parallax effects may influence the legibility of the overall assembly task instructions. As typical assembly task instructions have a length of at least 50 characters, this represents the

Figure 9.10: Alphanumeric assembly task information on generic PMUs using laser-based projective SAR

absolute limitations of a laser-projection based system. Even more, the projection of alphanumeric texts requires an overall strategy for view-independent legibility on generic surfaces.

9.6.3 *Discussion*

The presented results indicate the opportunities and limitations of projective SAR for use in VMS. Production engineers are able to choose whether they use OST-HMDs for their assessments or whether they use projective SAR. In general, OST-HMDs are single-user devices. Each user can have his own view of the virtual scene, whereas projective SAR can be used in collaborative situations out of the box. OST-HMD scenes must be synchronized in order to share the same VE. Therefore, the hardware and software has to be available multiple times. "Although HMDs provide an individual image plane for each participant of a multiple user session, users of large projection systems have to share the same image plane" [286, p. 297]. Each participant needs an OST-HMD hardware device which also multiplies the initial hardware investment and maintenance effort. Additionally, the registration quality of OST-HMDs may vary and configuration parameters, such as eye distance settings, must be adapted to each individual. This is more cumbersome than a passive device illuminating the PMUs within the VMS. Additionally, head-worn displays are heavy and cumbersome to wear all day long.

Generic tasks are compared below with respect to the three output devices considered:

PMU/DMU quality checks: Projective spatial augmented reality has proved capable of showing differences between PMU and DMU models. Product and process quality can be checked easily by projecting symbols, crosshairs for locations or outlines of assembly paths for designated locations. Therefore, no view-point specific AR is needed since parallax effects are small. The registration quality of SAR is more precise than current OST-HMD hardware (Microsoft Hololens). Therefore, tolerances in the PMU model can be detected more reliably than by using head-worn hardware.

Alignment tasks: For alignment tasks, high-precision registration of DMU and PMU is required. When using the room measurement of Hololens OST-HMD hardware, augmentation offsets between the physical item and augmentation is too big for assembly alignment tasks. Radkowski et al. present two additional calibration routines [315] to compensate these effects. Therefore, out of the box SAR is more suitable as a virtual 3D stencil.

Visual assembly support: Work piece tasks can be visualized using all three AR hardware devices. When projecting the manufacturing sequences onto the PMU, all workshop participants have a clear idea of which parts must be mounted next. Only for visualization of assembly paths in 3D space, OST-HMD hardware is highly recommended since manipulation takes place in open space inducing huge parallax effects for both SAR types. Reachability checks can be carried out more easily with SAR, since no body-attached hardware impedes the worker's movement.

Localization task: As described for the localization tasks, the projection of variants is a clear benefit of augmenting the PMU with additional information. For this generic task, all three AR variants also apply well, but SAR does not require all individuals to wear a head-mounted display. Therefore, all participants share the same digital information, even if they do not have or wear an OST-HMD. This supports the collaboration idea in workshop environments. Another use case mentioned in the requirements is the "assessment of logistics zone." All three AR devices can support this use case, sharing the same pros and cons as the localization tasks. In general, for logistics assessments on the assembly lines, AR technology overachieves the requirements in most cases. Simple pick-by-light or pick-by-beamer systems without 3D augmentation of parts are sufficient in most use cases.

Alphanumeric assembly instruction visualization: Assembly instructions must be visualized in some way for the workers in PV workshops since they must derive their actions from those assembly task descriptions. Legibility of SAR projected symbols and letters largely depends on the surface structure and the projection technology (see Di Donato et al. [314]). The fewer symbols are used, the better laser-based projective SAR works, otherwise video-based pro-

jective SAR outperforms. Video-based systems only work with limited colors and reflection parameters of assembly parts. The combination of the close spatial coupling between assembly parts and the corresponding alphanumeric assembly task description has proved to be helpful [287] but at this time there is no solution for surface-distortion compensation on generic surfaces PMUs (see Figure 9.10). Head-worn OST-HMDs are thus clearly the preferred choice when 2D non-spatial content must be coupled with 3D registered models.

9.6.4 *Summary*

The evaluation demonstrated benefits and drawbacks of all of the three AR hardware types compared:

Laser-based projective AR reaches high precision registration with sub-millimeter overlays and is highly independent of the PMU's surface color and reflection parameters. It can be natively used in collaborative workshop situations and does not disturb the participants in terms of required personal calibration routines, ergonomic matters, hygiene of shared body-attached devices and registration to the PMUs. Information density is limited by the XY scanning laser projection unit. Currently, 50 alphanumeric symbols can be projected at the same time. Laser-projectors are limited to one color.

Video-based projective AR shares most advantages of the laser-based systems, besides the independence of surfaces. It is limited to bright surfaces for good legibility. In contrast, it offers additional benefits in terms of projecting unlimited symbols simultaneously (limited by the resolution of the video projector) using all colors at the same time.

OST-HMDs offer individual views for each user. Each user requires his own device and all devices must be synchronized. In contrast to projective SAR systems, digital content can be visualized "mid-air" without having a projection surface. This bypasses the parallax effect of both projective SAR systems. Resolutions and FoV of current devices are still limited (reference device Microsoft Hololens v1). Ergonomic and practical restrictions apply, since users would have to wear the weight on their head for a long duration and recharge the batteries frequently throughout the day.

9.7 EVALUATION COLLABORATIVE TASK PERFORMANCE USING SAR

As presented in the design space evaluation, technical operators of the PV workshops have various options for visualizing the required digital content. However, no quantifiable answer has been given as to whether using AR in the VMS really speeds up work or whether other visualization possibilities are more efficient. Therefore, the most fre-

quent use case "localization task" (see SAR clusters in manufacturing proposed by Neumann and Majoros [289]) is evaluated in detail. PV workshops are collaborative environments in which all users must share knowledge in a highly interdisciplinary manner. Therefore, the division of labor and tasks is always a prerequisite. People must communicate with each other according to their assigned role (see Section 3.3.4 for participants and stakeholders of PV workshops). For instance, PV workshop participant 1 has to explain to everyone else where to attach a part on a certain cavity, bolt or spot of the car body chassis. This could represent one of the "localization tasks" mentioned above.

This study carries out an interdisciplinary, collaborative task between two users with unequal prior knowledge, measuring the performance and error rate of a localization task. Three different types of visualization scenarios are compared. Two scenarios are currently used in PV workshop situations, namely "verbal explanation" of assembly part locations and "visualization on a large screen." Verbal explanation is regarded as the baseline, since this is the most frequent way of communicating complex points. The third scenario uses 'spatial augmented reality' to communicate the localization tasks.

> **Publication**
>
> This evaluation was carried out for this doctoral thesis. Parts of this chapter were published in the paper on "Dual Reality for production verification workshops" by Otto et al. [80] for the peer-reviewed conference "CIRP Conference on Assembly Technology Systems 2016." The following evaluation has been extended and revised. The scenarios are described in more detail and the link to real world production validation scenarios explained in-depth.

9.7.1 *Study goal*

In real life, participants have unequal prior knowledge during the workshop sessions, since planners present their solution for the first time. The overall study goal is to find out whether there is any influence on the overall task completion time when using different methods for sharing prior knowledge on spatial circumstances.

Using SAR in collaborative workshop environments claims to reduce the efforts of interpreting spatially remote information (compare Kim and Dey [310]). Therefore, quantifying a typical use case of PV workshops provides insights into the effectiveness of applying such a new technology in PV workshops. Do visualization technologies (monoscopic displays and SAR) support the generation of a common spatial understanding in PMU-based workshop situations? As an

example, a "localization task" is carried out in three different scenarios. One of these scenarios uses spatial augmented reality technology.

9.7.2   *Methodology*

An abstracted collaborative workshop situation has been generated. Two participants had unequal prior knowledge and had to share knowledge on a certain spatial location as fast as they could, using three different methods to share this knowledge:

- Verbal explanations (baseline)

- 75" monoscopic screen

- Spatial augmented reality

The baseline scenario was verbal communication without the support of display devices. Both participants had to reach a common understanding on spatial relations of spatial points (e.g. for bolt or cavity locations) on a PMU as fast as possible.

9.7.3   *Experimental setup*

Figure 9.11 depicts the experimental setup. A car chassis wad placed in the middle of the 30 sq.m. lab room. Next to it, a 75" monoscopic display and the SAR projector were located. The SAR system was positioned in such a way that the digital content (3D-locations) could be projected properly inside the body-in-white.

Three persons were present during the study: The experimenter who led the experiment, a process expert (PE) and a technical operator (TO). The displays were arranged in such a way that both (PE) and (TO) could see the contents and both devices could be controlled properly by the (PE). The CAD workstation is operated by the PE and contains the DMU and a list of all spatial positions.

Figure 9.12 depicts the physical evaluation setup without the participants. For visualization of the CAD model, GOM Inspect Software has been utilized. In all three scenarios, the PE used this software to interact with the digital mock-up.

9.7.4   *Procedure and participants*

Having two active subjects is the simplest way of having a collaborative task with un-equal prior knowledge. In all three scenarios, the process expert (PE) must share his knowledge with the technical operator (TO). They both had to collaborate in order to solve the localization task. In the following, the 3D positions are called mounting spots for simplification.
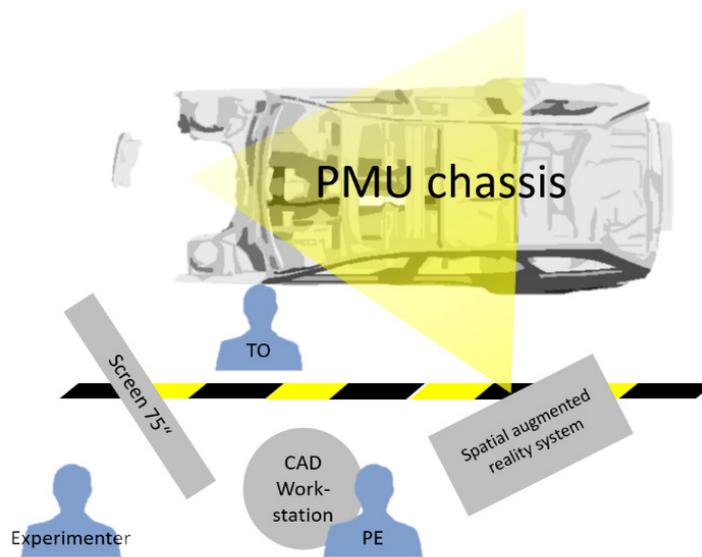
Figure 9.11: Study setup: A process expert (PE) explains spatial relations to the technical operator (TO) using three different methods: verbal description, support of display and SAR
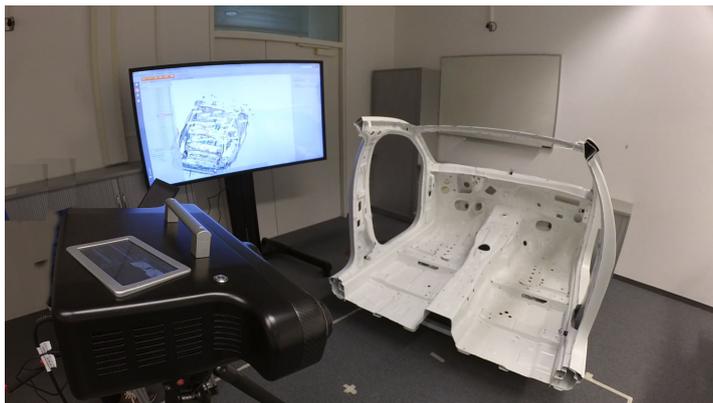
Figure 9.12: Physical evaluation setup without participants being present

Overall, 30 different mounting spots had been prepared and were presented in a randomized order. There were no repetitions of mounting spots within the ten runs. The experiment was carried out 11 times with 11 teams (22 participants). Each team carried out all three scenarios with 10 runs for each scenario. Therefore, a total of 330 data sets have been collected.

All participants took place on a voluntary basis, received no extra reward and were employees, interns or PhD students of the production planning departments.

The experimenter welcomed the team, consisting of two participants and assigned them randomly to their designated roles, PE or TO. Then the experimenter gave the (PE) a textual process description containing names of assembly parts. The (PE) must use this information to locate the mounting points of the assembly part in CAD model and guided the (TO), where the respective part must be assembled on the real chassis. Finally the (TO) must locate and reach the respective mounting spot on the PMU as fast as possible with his index finger. This sequence has been performed in three different scenarios: Using verbal explanations only, using a 75" screen with CAD data and using SAR, highlighting the respective parts directly on the PMU. For each scenario, the participants carried out the following sequence of tasks: Verbal explanation scenario:

1. The experiment leader reads out the next mounting point

2. PE localizes point in DMU by interacting with PC mouse in 3D model

3. PE verbally describes respective geometric mounting position to the TO

4. TO has to interpret this verbal information

5. TO has to reach the certain point with his index finger on the PMU

6. TO confirms the end of scenario by saying loudly "Check"

Scenario 75" screen:

1. The experiment leader reads out the next mounting point

2. PE localizes point in DMU by interacting with PC mouse in 3D model

3. PE moves the DMU viewpoint and verbally describes the respective mounting point position simultaneously

4. TO has to interpret this visual and verbal information

5. TO has to reach the mounting point with his index finger on the PMU

6. TO confirms the end of scenario by saying loudly "Check"

SAR explanation scenario:

1. The experiment leader reads out the next mounting point

2. PE localizes the point in a list of SAR points and activates it on SAR projector.

3. TO has to interpret this visual and verbal information

4. TO has to reach the projected mounting point with his index finger on the PMU

5. TO confirms the end of scenario by saying loudly "Check"

Time measurement began when the experiment leader passed the instructions to the (PE). The experiment ended when the TO touched the spatial reference point both have agreed on with his index finger. Below this is referred to as the task completion time. The experiment leader measured the overall task completion time and error rate. The experiment leader intentionally gave no explicit information on the trade-off between task solving speed and error rate.

9.7.5 *Results*

The results show a correlation between collaboration performance and the use of different visualization technologies.

Whereas task completion time of the verbal description scenario averages at 25.5s (SD = 19.5s), the same task can be achieved almost twice as fast (M = 14.0s, SD = 7.5s) by showing the DMU on a 75" screen (see Figure 9.13) to both participants. Using projective SAR technology leads to an additional reduction of almost 2.5 times (M = 6.0 s, SD = 2.1 s) in comparison to the monoscopic screen scenario and almost five times as fast as verbal description scenario. For both scenarios using shared visualization devices, results are error-free, whereas in the verbal description scenario one error occurred.

9.7.6 *Discussion*

Two major influencing factors for the significant reduction of task completion times are determined by observing the teams during their runs: During the verbal description scenario, collaboration is negatively influenced by differing mental coordinate systems (COS) of the users. For instance, the team members have not agreed on whether the word "left" refers to the car's, the PE's or the TO's COS, whereas
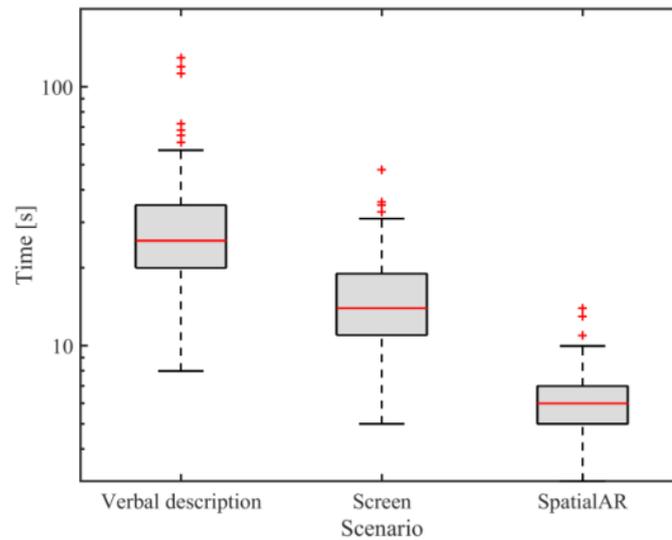
Figure 9.13: Evaluation on mean times and standard deviations (sorted by descending verbal description results)

this effect is eliminated when both participants can see a visualization. A second major impact factor is time-parallelization during task solving. For both visualization scenarios, the TO is anticipating the localization of the spatial position within the DMU while the PE is still searching for it.

As the results showed, only 1 out of 330 data sets failed. This indicates that the participants double-checked their results with higher task completion times rather than becoming faster at the expense of a higher error rate. In accordance with the results of Uva et al., "completion times can be significantly reduced, only if there is a substantial reduction in complexity due to SAR" [287]. This evaluation does not include quantitative performance measurements for all proposed visualization and interaction technologies.

9.8 CONCLUSION

In this chapter, SAR technology has been evaluated with respect to its applicability and performance for real life PV workshop situations.

The design space evaluation showed that SAR supports PV workshops in multiple use cases, such as PMU/DMU quality checks, alignment tasks and localization tasks. Integration of this technology as a standard method in PV workshops can efficiently help PMU buildups. Additional research must be carried out on how to project text on generic surfaces, so that multi user environments can benefit from alphanumeric information without experiencing distortion effects.

Overall, this chapter provided insights into practical design considerations for collaborative use of AR. The opportunities and drawbacks of SAR were defined in the design space evaluation, such as

distortion, parallax effects and generic projection of 2D content. Nevertheless, SAR visualization is a qualified method to be applied for many industrial use cases.

The **performance study** showed that using SAR for localization tasks as an additional visualization method in PMU-based workshops can speed up the overall PMU build-up. In addition, SAR tends to help reduce misunderstandings and generate higher common understanding of digital and physical mock-ups.

Thus, it can be concluded that SAR technology extends the VMS framework by an additional interaction technique and helps superimpose virtual models onto PMUs. Making adequate choices of methods and tools for PV workshop situations has a positive influence on the overall workshop performance. Therefore, SAR brings PMUs and DMUs closer together and enhances workshop efficiency.

OVERALL EVALUATION OF THE VIRTUAL MANUFACTURING STATION FRAMEWORK

This chapter presents an overall evaluation of the VMS framework for collaborative virtual production validation. This evaluation is based on contextual inquiry study including semi-structured expert interviews that were conducted stakeholders of the VMS framework. Drawing on the deficiencies found in the pre-evaluation (Chapter 4) and the key properties of the VMS (Chapter 5), this study shows the practical aspects of the VMS framework's implementation. In accordance with the research hypothesis formulated in Chapter 2, the expert interviews focused on whether DMU-based workshops can be carried out in a same quality, less time and with cost savings compared to hardware-based PV workshops".

10.1 PARTICIPANTS AND INTERVIEW PROCEDURE

Managers of the PV workshops were selected as interview partners since they had in-depth experience in organizing and carrying out PV workshops as well as thorough knowledge in the area of assembly planning. This qualified them as appropriate interview partners since they were also able to speak on behalf of their stakeholders. In addition, these participants were the most frequent customers of VMS since they participated in each PV workshop, next to the other participants, such as ergonomists, time experts, workers and many more. The expert interview partners' details can be found in Table 10.1. Overall, 1 hour 19min of audio was recorded and transcribed.

After attending a whole day of DMU-based workshops using the VMS, the experts were asked about the actual technical implementation of the VMS and its impact on planning quality, reduction of times and reduction of costs. Besides this, the experts were also encouraged to point out optimization potentials, advantages and disadvantages of the implementation details as well as to make suggestions for process improvements.

All expert interviews were carried out on a voluntary basis and in each participant's native language (all German). As an introduction, the experts were welcomed to the VMS. Each was shown a standardized demo of the 'Multi-Kinect system,' 'Spatial augmented reality system,' 'VR assembly simulation in the assessment software veo' and the 'floor visualization system' as depicted in Figure 10.1. This standardized demo took approximately 20 minutes. Audio files were recorded and afterwards transcribed using a text document tool. Fi-

Table 10.1: Detailed data of the expert interview on the interview partners

	Age	Gender	Position		Interview duration
Expert 1	41	masculine	Production manager	validation	18:44 min.
Expert 2	34	feminine	Production manager	validation	10:37 min.
Expert 3	52	masculine	Production manager	validation	12:03 min.
Expert 4	36	masculine	Production manager	validation	17:39 min.
Expert 5	44	masculine	Head of Production validation		20:39 min.
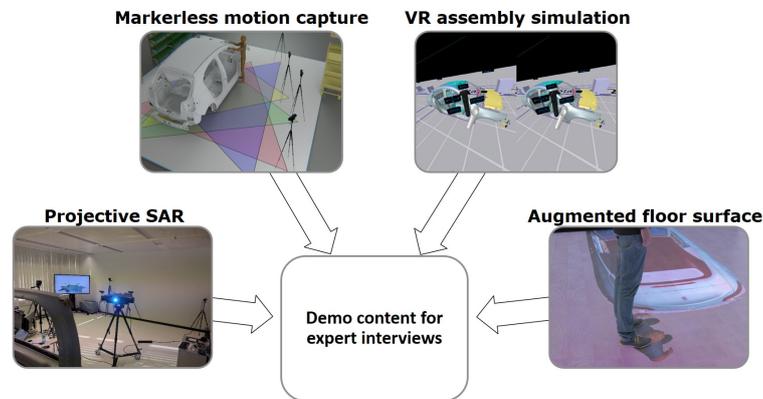
Figure 10.1: Demoed contents in advance of the expert interviews

nally, all interviews were qualitatively evaluated and coded in "MaxQDA 2018" software.

10.2 RESULTS

First, the interviewer asked the participants' opinions on whether the VMS could completely **replace physical prototypes** for PV workshops. All experts agreed that the current implementation of the VMS was not yet able to completely replace the hardware simulation of building up a vehicle. They confirmed that at least one out of several workshops needed to be hardware-based. Nevertheless, the vast majority of verification criteria could be simulated virtually. Experts found that there was a lack of options for the easy simulation of forces, haptics and flexible parts, for example while clipping or screwing. Expert 3 confirmed this by stating that he would need the "feeling of pressing in plugs" to be able to validate whether the assembly can be carried out correctly.

The interviewer asked about the **output possibilities and interaction methods** and how changes by true-to-scale visualization influence outcomes of PV workshops. Four out of five experts thought that there were good opportunities for applying VR glasses and Microsoft Kinect. Expert 1 provided multiple examples of simulations that take place in the vehicle interior, such as the assembly of grab handle consoles, sky panel assembly or front wall assembly in the virtual vehicle. Two experts pointed out the advantages of ergonomic assessments using VR glasses in combination with full-body tracking using Kinect, such as assessments of accessibility or necessary flexion of the worker while mounting a certain part. However, one of the five experts stated that he could not currently imagine for which actual topics VR glasses or Microsoft Kinect could be applied due to limited accuracy and precision. For expert 2, the possibility of digital "capturing, storing and evaluation of the walking paths" through the sensors integrated in the LED floor represented a major advantage of the VMS.

After demoing **spatial augmented reality** to the experts, the interviewer asked them how this technology could support PV workshops. Four experts stated that they would like to deploy projective SAR to visualize different variants or optional equipment onto the physical prototype. According to Expert 3, it is currently impossible to assess all variants or extra equipment, as only one physical mock-up in a particular variant can be considered during a PV workshop. Nevertheless, using SAR, components that have to be mounted in other variants can be assessed digitally. Expert 4 did not consider the use of spatial augmented reality to be necessary since he believed that it would be sufficient to have a non-available component displayed on a computer monitor or via a 'VR goggle.' Therefore he personally saw no benefit in using SAR.

Changing the scope to organizational topics, the interviewer asked whether workshop participants should be **empowered to use the assessment tools on their own** or whether they should rely on a specialist using the technology for them. All experts agreed that there should be a dedicated operator responsible for all simulation tools, whose main task is to support technical operations during the workshops. Expert 5 stated that in particular the PV managers should be familiar with technical possibilities as well, but "each workshop participant should be able to concentrate on his area of expertise."

In addition, Expert 2 expressed doubts that workers and craftsmen could take advantage of virtual assessments due to their 'lower affinity towards virtual assessments,' as they were accustomed to 'exclusively working with hardware prototypes' during their regular daily work. This would have a direct impact on the way they work with virtual simulations and on their trust in the assessment results.

The experts were asked which **additional features or data sets** were required in the **simulation environment** and what additional value

this would generate. All experts wished to integrate more dynamics, especially in the context of the assessment of walking paths, fusing digital recording and simulation with the help of the LED floor. This would also facilitate planning and validation of "shopping carts" or "material zones," which matches the findings in Section 8.4.1. The feature request of "simulating lifting equipment" was also mentioned by two experts. They wished to have a multi-axis kinematic model for proper validation of the overall handling device, for example to interactively obtain a sense of how a worker can manipulate the handling device with the attached parts. Expert 1 stated that for semi-active handling devices cycle-time diagrams are required in the context of human-robot collaboration.

Furthermore, Expert 3 remarked that he would like to carry out all assessments in one simulation software program. In addition, for time analysis he suggested using predetermined motion time systems such as methods-time Measurement (MTM) to be automatically derived by the holistic simulation as well as automatic deduction of ergonomic scales.

The interviewer asked the experts **how the usage of the VMS framework subjectively influenced the process of PV workshops.** By using VMS, the experts stated that there were opportunities for simplification and and future-oriented workshops with fewer PMUs available. Expert 3 confirmed that the VMS comes very close to PMU-based assessments. He also mentioned the advantage that product data is always up-to-date in the virtual domain, since product data changes occur continuously during the development and planning phases. The experts appreciated the possibility of parallel visualization of contents on the large high resolution displays to create a clearer understanding for workshop participants. Expert 5 also supported the idea that the number of participants could be reduced if the system could automatically derive ergonomic or walk path assessments, as ergonomics or MTM experts would no longer be necessary to be present at the workshops.

When asked about the **changed processes in production validation workshops due to the VMS,** all experts again emphasized that at least one PMU is required at the very end of the production preparation process. This is obligatory since assembly tests must ensure perfect haptics, simulation of forces and buildability. Nevertheless, DMU-based validation is sufficient or even outperforming the remaining verification criteria. In addition, workers from the assembly department continue to struggle with virtual assessments since they do not completely trust in sole DMU simulation results. However, all five experts agreed that even during PMU phases, the VMS can very well be a supplement to PMU. Different variants or special equipment, which are currently not available, can be viewed in the virtual domain or could be assessed with the help of the VMS. In addition, three experts

(experts 1, 3 and 4) wanted to change current PV workshop, which focus on production processes, so that they exclusively use virtual techniques. Expert 2 stated that the VMS enabled him to introduce new workshop types on urgent topics, so that smaller workshops could be held spontaneously without external operator support.

Discussing the impact of the VMS framework on a company-scope level, the interviewer asked the experts whether they could **save time or costs using markerless tracking** during the PV workshops instead of marker-based systems. Expert 4 clearly saw a time advantage in using marker-less tracking with the Multi-Kinect system. This allows processes to be digitally simulated and assessed quickly, since it is not necessary to put on cumbersome motion capture suits. It usually takes 30 minutes to put on these suits and to calibrate them properly, as the markers on the suit must be placed correctly and often multiple test runs must be carried out before the movements can be captured by marker-based systems. By using the VMS methods, all other experts also considered time savings throughout the workshops to be very probable, but they demanded technically flawless preparation of the workshop to achieve these time savings. The experts did comment on the topic of direct cost reductions. Overall, expert 5 was skeptical whether overall time could be saved. According to his statement, there will be no temporal difference, regardless of whether the process is performed virtually or on a PMU-basis. But he remarked that switching between variants was not possible in the physical domain and dismantling times of PMU vehicles before the actual start of the PV workshops could be reduced.

The interviewer wanted to know where **costs can be reduced by using the VMS** and in which stage of production, either in PV workshops themselves (through reduced time or number of participants), in prototype construction (through the reduction of physical prototypes) or in the factory (through better solutions and higher quality). The answers of the five experts differed vastly. Expert 1 saw a potential for cost savings in reduced PV workshop duration and therefore lower personnel costs. Three experts (Experts 2, 4 and 5) stated that the main potential for savings could be achieved by reducing the number of physical prototypes since hardware PMUs for the production validation workshops is expensive. Expert 4 weighed up the initial investment of the VMS in relation to the cost reductions for hardware prototypes. He came to the conclusion that the business case is clearly positive. Experts 2 and 5 agreed that the cost savings are noticeable in the target plant through better process solutions and higher product quality as larger product quality flaws are detected earlier. Expert 2 exemplified these cost savings when finding poor buildability of components for non-standardized processes earlier.

All experts agreed that by carrying out virtual PV workshops in the VMS, the resulting quality will be at least as high as by traditional

PMU-based workshops. In terms of quality, all five experts agreed that **more validation criteria can be validated using VMS methods**. They especially emphasized the importance of the possibility to visualize all product variants and extra equipment, which is currently not possible in PMU-based workshops.

As the interviewer discussed the **optimization potentials**, three experts mentioned that it would not be possible to determine forces such as how parts snap in physically. Moreover, expert 5 did not see a direct benefit of the VMS in late phases, when PMUs are available in parallel. With regard to improvement potentials, the experts gave multiple inputs: Expert 2 wanted to simplify the usability of the simulation software veo. For her, the software had to be as "simple as possible" and it had to run more fluently during the workshops. In addition, she mentioned that missing collision avoidance could be regarded as a software bug, for example when the avatar can walk through the virtual car. Since this is not possible in the physical domain, the expert was afraid that this might lead to rejection by the non-experienced workshop participants. Furthermore, expert 4 wanted to integrate more data sources such as simulation results of other tools, e.g. simulated walk paths.

10.3   DISCUSSION AND SUMMARY

The evaluation of the expert interviews shows that the managers of the PV workshops sometimes have different opinions about the individual topics regarding the VMS. However, it can be clearly said that the VMS will not be able to completely replace the hardware structures during PV workshops in the near future. However, the experts determined potential use of the VMS as a supplement to the PMU build-ups as it offers the possibility of validating more vehicle variants and verification criteria than before.

All experts confirmed that a constant quality of the assessment results by the VMS can be guaranteed. Cost savings were confirmed by the experts as well. Nevertheless, the experts did not agree on a single instance where cost savings can be attributed to. Potential cost savings were attributed to four causes: The reduction of time within PV workshops, the reduction of the number of physical prototypes and finding better solutions and achieving a higher planning quality for the factory. Most experts considered it likely that time will be saved in PV workshops, but also stated that it was difficult to quantify time reductions in highly dynamic and non-repeatable PV workshops.

CONCLUSION & OUTLOOK

As a major part of this doctoral thesis, a newly designed framework for virtual validation of manual assembly processes has been introduced and evaluated. In the domains of production engineering, computer supported collaborative work and human computer interaction, new fundamental research has been presented as well as the application of basic research to industrial practice. Especially in the domain of human computer interaction, fundamental research has been published such as studies on human perception. In the following sections, research results and practical usage of the proposed methods are summarized and discussed in terms of the research questions and research hypothesis presented in Section 2.2.

11.1 PRACTICAL EXPLOITATION OF THE VMS FRAMEWORK

The presented VMS framework has been implemented at Daimler AG in Ulm and Sindelfingen and is already widely used for production validation.

By now (2020), 15 multi-week PV workshops for ten car models have been carried out using the VMS framework's methods. International workshop groups from the U.S., South Africa, France, China and Germany have taken part in PV workshops in the VMS environment. Batch assessment allowed the efficient validation of entire final assembly lines in the respective countries. Each PV workshop chooses a typical model-mix of products that will be produced in the respective global final assembly line. The manufacturing sequence and precedence graphs are specifically adapted to factories around the world.

Methods of the proposed VMS framework are used on a daily basis for PV assessments. Depending on the status of the PDP, different methods are applied. The LHRDs are always used for collaborative public displays. They have proved to work reliably and to be helpful for collaboration on complex, heterogenous data. DHM assessments are carried out more frequently using the VMS as preparation lead times are minimized. Using the proposed methods, in PV workshops DHM animation can be carried out on-the-fly compared to virtual assessments afterwards. The LED floor is used in later phases of the PDP when walk path optimizations are carried out. For PMU-based workshops, projective spatial augmented reality is not in vast practical usage yet as data preparation and projection on generic surfaces

Video: Valtteri Bottas using the VMS: https://youtu.be/FUINhoVEji8

is still being researched even though PV workshop managers see the benefits of such a method.

Overall, workshop participants' feedback is enthusiastic and positive with respect to the presented VMS framework's methods and tools. They appreciate the massive amount of simultaneously presented data, markerless interaction capabilities and the space for large workshop groups.

11.2   RESEARCH QUESTIONS AND RESULTS

Overall, the research questions and the hypothesis are answered as follows:

> **Question 1 - Production Engineering**
>
> How is assembly validation presented in the literature and carried out in industrial practice? Which assessment criteria must be evaluated in the automotive production validation process?

In Chapter 3, a holistic literature review on automotive production, production planning and production validation has been presented. PV workshop contents, participants and the workshop characteristics have been described in-depth to understand the context of the VMS presented below. Additional insights are given on product-related, process-related, human-related, logistics-related and resource-related assessment criteria. These validation criteria not only apply to automotive final assembly but also to the manufacturing industry in general to assure proper production ramp-ups.

> **Question 2 - Production Engineering**
>
> Where are the deficiencies in current physical and virtual automotive verification processes, methods and tools? Which criteria can already be assessed in the virtual domain?

Based upon the literature review of current automotive final assembly, a contextual inquiry study with semi-structured expert interviews has been carried out to understand the state-of-the-art production validation processes, methods and tools. Supervision of PV workshops and interviews with experts in virtual technology and production planning, have helped reveal current drawbacks of PV processes in physical and digital validation. Hardware-based workshops in particular struggle with limited product parts availability, outdatedness of physical parts, high costs and limitations on assessing a wide variety of extra equipment. Therefore, all experts agree on the necessity to strengthen digital assessments as they assume DMU-based workshops to be quicker, enable assessment of all options and allow the instantaneous change of models. Nevertheless, they also provide insights into current limitations on using DMU-based assessments, namely too

little interaction between workshop participants, no interactive manipulation of parts, no holistic process simulations, a lack of digital human models and layouts and especially hard data preparation processes prior to PV workshops. These drawbacks in virtual assessment methods are in accordance with the findings in the literature review.

> **Question 3 - Human Computer Interaction & CSCW**
>
> Which requirements can be derived for a collaborative virtual assessment framework for the production validation of manual assembly tasks? What is the design space for a framework for virtual, mixed and car assemblies for batch assessments?

The requirements for a novel framework used in PV workshops have been derived from strategic, tactical and operative objectives. Tactical objectives comprise the reduction of available PMUs, achieving higher data maturity, enabling late changes and smooth ramp-ups of mixed model lines. On an operational level, this leads to the following consequences: more verification criteria can be assessed in the virtual domain, workshops are supported by interactive components, PPR data can be integrated for the first time, all methods are applicable throughout the PDP and the preparation effort for virtual assessments is reduced.

In order to reach these requirements, a novel framework for collaborative PV workshops is proposed: The "Virtual Manufacturing Station." As the literature review reveals, multiple works describe the design space for collaborative, virtual environments in co-located, multi-user workshop setups. These concepts are adopted to the automotive and, in general, to the manufacturing industry. By analyzing product details as well as stakeholders' and participants' needs, six key properties are proposed. All key properties support the concept of the VMS. Each key property is exemplified with the in-depth research studies carried out:

- **Collaborative Virtual Environment:** The collaboration features of the VMS were evaluated in two studies. The VR assembly simulation software presented drives the PV workshops. It is optimized for usage in collaborative situations so that all participants share the same virtual environment. Additionally, a spatial augmented reality study showed that the collaboration performance increased in terms of task completion time and decreased error rates.

- **True to scale output:** Having realized a large-scale LED MDE, multiple advances were made. Isometrically registered tracking setups are enabled by utilizing true-to-scale output devices. An in-depth evaluation showed that size estimations using a LED floor output device showing true-to-scale content is more accurate compared to relative-sized content.

- **Multi-User support:** This key property has been exemplified in ergonomic validations. Multiple PV workshop participants can be tracked simultaneously and instantaneously inside the tracking frustum and directly manipulate the digital human model.

- **Asymmetric and symmetric output for AR/VR assessments:** Each user can choose either public wall-sized displays for using the CVE or use private displays for individual viewpoint renderings of individual contents.

- **Integration of PMU&DMU:** All true-to-scale visualizations aim for an isometrically registered VMS environment. When tracking physical objects, such as assembly parts or tools in the VMS, their virtual representation behaves in the same way. In particular, the design space evaluation of spatial augmented reality provides in-depth insights into how the physical and virtual domain can be blended in the VMS.

- **Integration of natural user interfaces:** Markerless tracking systems allow instantaneous human full body tracking. Using this technology, the virtual domain can be manipulated in an intuitive manner, such as manipulating the digital human model for ergonomic assessments or mid-air gesture interfaces for disassembly of car parts.

All key properties are realized in the VMS, exemplified in multiple studies and are in active usage in PV workshops.

> **Question 4 - Human Computer Interaction**
>
> Which components are required in a VR batch assembly assessment simulation software and how can the performance and limitations of such a VR assembly assessment system be quantified?

The requirements for batch assembly assessment validation software are deduced in this doctoral thesis. Most importantly, PPR data have to be fused in one central simulation environment for the entire factory. In contrast to the evaluated state-of-the-art commercial software components, batch validation of hundreds of work places are in the focus of the presented simulation software. It unifies all interactable components, such as standardized protocols for markerless full-body tracking, seamless geometry projection for spatial augmented reality and fast switching between various product variants. Another key property is the ability to visualize content simultaneously on asymmetric output devices. The software presented is able to simultaneously show the CVE true to scale on LHRDs, the augmented floor surface as well as in VR.

Carrying out assembly simulations in VR raises the question about the overall system's performance. Therefore, a novel VR2A benchmark is proposed to quantify the overall VR system's limitations. VR2A can be universally applied for different VR environments, simulation software and VR hardware devices. As a standardized benchmark, it evaluates the overall assembly simulation system's capability to visually assess assembly processes and measures the overall achievement rate in terms of precision and spatial limitation. Since the VR2A benchmark is a standardized, open experiment design, the virtual scene is published under Creative Common licence, so that all third-party researchers can reuse this scene and evaluate their own VR system's performance. The VR2A score represents the overall ability of the system to visually assess clearances and small assembly part sizes without decomposing single influence parameters within the VR interaction cycle.

Using this VR2A score in combination with the presented simulation framework, assembly part sizes of 75mm and relative clearance of 109% are required to achieve an overall VR2A score of 70%. The evaluation showed that VR2A is a reliable benchmark for quantifying the system's performance and for revealing its limitations in assembly assessment.

> **Question 5 - Human Computer Interaction**
>
> How can a markerless, scalable tracking system be realized and what advantages of motion capture can be achieved? What are the limitations of markerless tracking systems and what tracking performance can be determined?

Animation of digital human models is crucial for assembly validation. Therefore, a scalable, markerless motion capture system has been presented and evaluated using multiple depth-cameras. The overall system has been developed to interactively manipulate digital human models on-the fly for PV workshops.

Kinect V2 is used to realize such a scalable, markerless, multi-depth camera system. The overall system's properties and limitations in markerless full-body motion capture were evaluated, showing that the tracking results of single sensors are error-prone and rotation-variant. Rotation-invariant joints of skeletons can be used for robust multi-depth camera fusion. A procedure of web-based data gathering, registration, filtering, temporal alignment and fusion is proposed in order to obtain skeletal tracking data for larger, scalable frustums. In an experimental setup, the proposed system revealed a mean accuracy error range from 9.6 mm (SD=4.5 mm) to 26.6 mm (SD=10 mm).

These results indicate that the overall system's performance is sufficient for ergonomic assessments, since DHM animation only requires a rough pose estimation in PV workshops. In a practical follow-up study, motion tracking results were applied to DHMs and compared

with physical assessments. Results show that nine out of eleven standardized working postures in the EAWS are feasible using the system, besides large bending over (>60° upper body bending) and lying flat on the floor. By using fused sensor data with the novel method, jitter, occlusions and rotation variance are reduced. Comparisons of the same physical and virtual arrangement show that the presented full-body motion capture system can be used for practical usage (see Figure 7.19) As a result, production related issues can be identified on-the-fly in the virtual domain, even though small deviations between PMU and DMU still remain. Overall, having a markerless motion capture system available in PV workshops is a helpful solution to reveal quality and ergonomic issues on-the-fly, while checking the overall product.

> **Question 6 - Human Computer Interaction**
>
> How do wall-sized displays and floor visualization displays influence spatial perception? Does the variation of interaction techniques have any influence on spatial perception and task performance?

The literature review on LHRDs reveals that these devices allow the display of more data details, multi-scale data, higher complexity and dimensionality and enable collaboration. These effects are leveraged in the VMS framework implementation. Nevertheless, the literature also indicates that there further reserach on LHRDs and their influence on spatial perception will be required, to which this doctoral thesis contributed.

Six generic use cases were presented where augmented floor surfaces may help in the context of the manufacturing industry and especially in PV workshops, namely virtual travels with interactive maps, self-navigation in virtual space, size estimation, virtual true-to-scale stencil, true-to-scale work place layout assessments and interactive walk path assessments.

Two setups for VMS framework's true to scale output devices were proposed, implemented and evaluated. First, a large-scale floor projection system for true-to-scale digital work place layouts are proposed. Since practical usage of this prototype system has revealed considerable benefits in PV workshops, a second LED-based setup is created with 32 sq.m. of LED walls and 54 sq.m. of LED floor.

Two studies using this LED floor gave insights into spatial perception and task completion time in PV workshops: The first experiment presented insights on spatial perception and size estimations using different modes of perception. Different scenarios showing true-to-scale (absolute) and relative-scale visualizations reveal significant differences in size estimations. Size judgment accuracy is better using absolute visualization scenarios, whereas task completion time increases using the LED floor, both compared to the relative-sized vi-

sualizations in the baseline scenario. As various use cases in PV workshops depend on reliable spatial estimations of humans, the LED floor proved to be a proper interface and to be a helpful tool for visualization of virtual true-to-scale contents. Second, a practical experiment was carried out, where the LED floor was used as a virtual stencil. Using the LED floor as a virtual stencil, the task completion time decreased by a factor of almost two and the number of errors was reduced from 75 errors to zero errors.

These studies showed that using LHRDs in PV workshops directly help solve certain tasks more reliably (spatial estimations) as well as faster in terms of task completion time (virtual stencil). Practically, slight increases in task completion time while performing spatial estimations are negligible in practical PV workshops as the overall time percentage of size estimations ranks low compared to the benefit of finding high quality solutions.

> **Question 7 - Human Computer Interaction & CSCW**
>
> What is the design space for using projection spatial augmented reality in co-located PMU and DMU-based environments and how does computer-mediated communication influence workshop performance in terms of task completion time and errors?

In a design space study, the usage of projective spatial augmented reality was evaluated for the manufacturing industry. Five different generic tasks were evaluated, namely PMU/DMU checks, alignment tasks, visual assembly support, localization tasks and alphanumeric assembly instruction visualizations. These generic tasks represent typical validation criteria within PV workshops. The design space evaluation revealed the limitations of SAR and showed the need for additional research on how to project text onto generic surfaces, so that multi-user environments can benefit from alphanumeric information without dealing with distortion effects. The VMS integrated this technology to efficiently support PMU build-ups, even though some drawbacks of non-view-dependent SAR were made clear in the design space evaluation, such as distortion, parallax effects and generic projection of textual information.

Applying projective SAR in a performance study showed task completion time decreases when it is used in PMU-based workshops for localization tasks. Hence, in physical PV workshops collaboration efficiency can be increased by using SAR visualizations. Compared to two other error-prone baseline visualization scenarios, no errors could be found using the SAR scenario.

> **Question 8 - Production Engineering**
>
> What implications are there regarding time, costs and quality compared to hardware based workshops? Can integrated virtual environments help production planning to achieve goals more quickly and reliably during the planning and validation of automotive manual end assembly?

In a contextual inquiry study including a semi-structured expert interview, PV workshop managers responded to the overall realization results of the VMS. All experts were stakeholders organizing PV workshops themselves. 1h 19min of transcripts revealed the advantages and drawbacks compared to hardware-based PV workshops.

Results showed that currently not all verification tasks can be achieved by sole usage of the VMS framework's virtual assessment methods, but it helps reduce the overall amount of required PMUs especially in the early phases of the PDP. Novel interaction techniques, such as full-body markerless motion capture and VR assessments, are able to reduce the overall required amount of time in PV workshops and increase the quality of the overall assessment validity. Novel interaction capabilities such as the augmented floor surface shorten the time spent on assessing the walk paths.

The experts mentioned three fields of cost savings but did not quantify them: More efficient production validation, reduction of physical prototypes in early phases and even benefits in production itself due to better solutions from production planning and higher product quality.

Overall, in Section 2.2 the research hypothesis has been formulated as follows:

> **Research hypothesis**
>
> Utilizing collaborative virtual environments in production validation workshops for manual assembly tasks, verification criteria can be assessed in same quality, less time and less costs compared to hardware-based workshops.

Summarizing answers to the research questions, the research hypothesis can be approved. The VMS framework represents a collaborative, co-located virtual environment for production validation workshops. All of the presented studies directly or indirectly support the research hypothesis even though each study contributes to singular aspects regarding the VMS key properties. Each study contributes to a specific research question and focuses on an exemplary use case of the PV workshop.

Additionally, the final expert interviews (see Chapter 10), the daily practical usage of the VMS framework and the broad support of the VMS implementation revealed the benefits of shifting PV workshops

from the physical to the virtual domain. More verification criteria can be assessed now using DMUs, such as ergonomic assessments and walk path layouts, and strengthen the collaboration performance. By using the VMS's methods, such as LHRDs, tracking and SAR, product quality increases and errors are reduced when constantly building up the product with the methods presented. Besides one study (size estimations), all experiments indicate that task completion times are lowered using the proposed virtual methods. Thus the overall PV workshop execution time can also be reduced. On a strategic company level, these effects of higher quality and reduced task completion time result in cost reductions compared to hardware-based PV workshops as fewer physical prototypes are required, higher quality product parts can be found and process quality is increased throughout all phases of the PDP.

11.3 OUTLOOK

By now, the VMS is already widely used for production validation. However, multiple research domains have not been integrated. One can differentiate short-term and long-term research requirements in the context of the VMS.

In the **short term research**, interaction within the VMS framework is still an ongoing topic for adding realistic animation to the virtual simulation models. PV requires the completion of more verification tasks solely using DMUs. Hence, the simulation environment must be enriched with additional features, such as collision detection, avoidance and flexibilization of parts in the presented batch validation system. Even though research has already presented many algorithms for collision avoidance and flexibilization of DMU assembly parts, they are still not widely used in real-time interaction scenarios. In addition, the simulation of haptic feedback is required in VMS's assessments to give users a realistic feeling of assembly parts.

In the **long term**, the overall research goal is to achieve a fully automatic deduction of a holistic, dynamic simulation scene by using generative models. The so-called "cyberphysical production equivalence" of real production processes must be accomplished without any additional authoring effort (compare Stork [100]). By now, DMUs used in VMS framework use simplified kinematics of product representations. All complex animation is injected by the user's interaction using tracking systems and human intelligence. Artificial Intelligence (AI) approaches may help establish generative motion models. In order to reach a state of cyberphysical equivalence, additional research has to be carried out in the area of collision-free path planning, avoidance of starting collisions, automatic deduction of product precedence trees, AI human reasoning and generative DHM animation for natural motions and standardized kinematics of DMUs. When

all these components are available, holistic simulation of manual assembly processes will be achieved without manual interaction. The resulting simulation model may be used for higher quality validation in the VMS framework. The VMS users' tasks will change to the adaption of simulation parameters in order to optimize simulation models instead of dedicated 3D interaction. Therefore, the overall planning and assessment effort will be drastically reduced since detailed work plans are automatically deduced by the generative simulation model. Simulation results only have to be validated and optimized.

In future, having such a holistic assembly simulation may lead to a higher automation degree in final assembly. If the simulation is capable to adaptively react to the complex production program and to unexpected situations in final assembly, human flexibility can be incrementally replaced by automated machinery. This leads to a shift of manual labor to automated processes and changes the paradigms on how planning and production is carried out.

BIBLIOGRAPHY

[1] Hoda A. ElMaraghy and Waguih H. ElMaraghy. "Variety, Complexity and Value Creation." In: *Enabling Manufacturing Competitiveness and Economic Sustainability*. Ed. by Michael F. Zaeh. Cham: Springer International Publishing, 2014, pp. 1–7. DOI: 10.1007/978-3-319-02054-9_1.

[2] Behzad Saberi. "The Role of the Automobile Industry in the Economy of Developed Countries." In: *International Robotics & Automation Journal* 4.3 (May 17, 2018). DOI: 10.15406/iratj.2018.04.00119.

[3] International Organization of Motor Vehicle Manufacturers. *2019 Statistics | OICA*. 2019. URL: http://www.oica.net/category/production-statistics/2019-statistics/ (visited on 03/29/2020).

[4] International Organization of Motor Vehicle Manufacturers. *1999 Statistics | OICA*. 1999. URL: http://www.oica.net/category/production-statistics/1999-statistics/ (visited on 01/14/2019).

[5] Dr. Sebastian Wedeniwski. *The Mobility Revolution in the Automotive Industry*. 00000. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015. URL: http://link.springer.com/10.1007/978-3-662-47788-5 (visited on 01/07/2016).

[6] Felix Kuhnert and Christopher Stürmer. "Five Trends Transforming the Automotive Industry." In: *PWC* (2018). URL: https://www.pwc.com/hu/hu/kiadvanyok/assets/pdf/five_trends_transforming_the_automotive_industry.pdf.

[7] Daimler AG. *CASE - Intuitive Mobilität*. Jan. 2019. URL: https://www.daimler.com/case/ (visited on 01/20/2019).

[8] JP Morgan. *Driving into 2025: The Future of Electric Vehicles | J.P. Morgan*. Aug. 10, 2018. URL: https://www.jpmorgan.com/global/research/electric-vehicles (visited on 01/17/2019).

[9] SAE Mobilus. *J3016B: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles - SAE International*. June 15, 2018. URL: https://www.sae.org/standards/content/j3016_201806/ (visited on 11/10/2019).

[10] McKinsey & Company. *Automotive Revolution – Perspective towards 2030*. 00000. URL: http://informationresearch.e.corpintra.net/ior/fileadmin/studien/studien_2016/McKinsey_automotive_2030.pdf (visited on 02/10/2017).

[11] K. Feldmann and S. Slama. "Highly Flexible Assembly – Scope and Justification." In: *CIRP Annals* 50.2 (Jan. 1, 2001), pp. 489–498. DOI: 10.1016/S0007-8506(07)62987-4.

[12] H. ElMaraghy, G. Schuh, W. ElMaraghy, F. Piller, P. Schönsleben, M. Tseng, and A. Bernard. "Product Variety Management." In: *CIRP Annals - Manufacturing Technology* 62.2 (2013), pp. 629–652. DOI: 10.1016/j.cirp.2013.05.007.

[13] M. Johnny Rungtusanatham and Fabrizio Salvador. "From Mass Production to Mass Customization: Hindrance Factors, Structural Inertia, and Transition Hazard." In: *Production and Operations Management* 17.3 (May 6, 2008). 00054, pp. 385–396. DOI: 10.3401/poms.1080.0025.

[14] Ingrid Göpfert, David Braun, and Matthias Schulz, eds. *Automobillogistik*. Wiesbaden: Springer Fachmedien Wiesbaden, 2017. URL: http://link.springer.com/10.1007/978-3-658-11103-8 (visited on 11/14/2016).

[15] Nestor Andres Arteaga Martin, Thomas Bär, and Rainer Stark. "A Reference Framework for Manual Assembly Simulation." In: *Smart Product Engineering*. Ed. by Michael Abramovici and Rainer Stark. Lecture Notes in Production Engineering. Springer Berlin Heidelberg, 2013, pp. 293–302. URL: http://dx.doi.org/10.1007/978-3-642-30817-8%E2%80%97%2029.

[16] Progenium. *Produktportfolios Der Fünf Großen Deutschen Autohersteller in Deutschland: 1990 Bis 2014 Das Dilemma Mit Der Vielfalt. München, PDF*. 2014. URL: https://docplayer.org/15484141-Produktportfolios-der-fuenf-grossen-deutschen-autohersteller-in-deutschland-1990-bis-2014-das-dilemma-mit-der-vielfalt-muenchen-24.html (visited on 01/19/2019).

[17] PROGENIUM. *Das Dilemma Mit Der Vielfalt*. URL: https://www.progenium.com/wp-content/uploads/2016/04/PROGENIUM-Das-Dilemma-mit-der-Vielfalt.pdf.

[18] Paul G. Maropoulos and Dariusz Ceglarek. "Design Verification and Validation in Product Lifecycle." In: *CIRP Annals-Manufacturing Technology* 59.2 (2010), pp. 740–759.

[19] Yoram Koren. *The Global Manufacturing Revolution: Product-Process-Business Integration and Reconfigurable Systems*. Vol. 80. Wiley Series in Systems Engineering and Management. John Wiley & Sons, 2010.

[20] Henry Ford and Samuel Crowther. *My Life and Work*. Doubleday, Page & company, 1922. 311 pp. URL: http://archive.org/details/mylifeandwork00crowgoog.

[21] George Chryssolouris. *Manufacturing Systems: Theory and Practice*. Springer Science & Business Media, Mar. 9, 2013. 431 pp. URL: https://www.springer.com/de/book/9780387256832.

[22] H. Kinutani. "Modular Assembly in Mixed-Model Production at Mazda." In: *Transforming Automobile Assembly*. Ed. by Koichi Shimokawa, Ulrich Jürgens, and Takahiro Fujimoto. Berlin, Heidelberg: Springer Berlin Heidelberg, 1997, pp. 94–108. URL: http://link.springer.com/10.1007/978-3-642-60374-7_8 (visited on 08/03/2016).

[23] UGS PLM Software. *General Assembly Manufacturing*. White Paper. Plano, TX: UGS PLM Sof tware, 2007, p. 15. URL: https://www.plm.automation.siemens.com/zh_cn/Images/11560_tcm78-49841.pdf.

[24] Uwe Achterholt, Roland Schmid, and Fiona Sheridan. "Momentum: KPMG's Global Auto Executive Survey 2009." In: (2009), p. 44. URL: https://www.kpmg.de/docs/20090101_Global_Auto_Executive_Survey_2009.pdf.

[25] Julian Weber. *Automotive Development Processes: Processes for Successful Customer Oriented Vehicle Development*. 00073. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. URL: http://link.springer.com/10.1007/978-3-642-01253-2 (visited on 08/19/2016).

[26] Antonino Gomes de Sá and Gabriel Zachmann. "Virtual Reality as a Tool for Verification of Assembly and Maintenance Processes." In: *Computers & Graphics* 23.3 (June 1, 1999). 00000, pp. 389–403. DOI: 10.1016/S0097-8493(99)00047-3.

[27] Waldemar Walla. "Standard- Und Modulbasierte Digitale Rohbauprozesskette : Frühzeitige Produktbeeinflussung Bezüglich Produktionsanforderungen Im Karosserierohbau Der Automobilindustrie." KIT Scientific Publishing, 2017. 205 pp. URL: https://publikationen.bibliothek.kit.edu/1000061838 (visited on 02/10/2019).

[28] John Brooke. "SUS-A Quick and Dirty Usability Scale." In: *Usability evaluation in industry* 189.194 (1996), pp. 4–7.

[29] James R. Lewis. "Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies." In: *International Journal of Human-Computer Interaction* 14.3-4 (2002), pp. 463–488.

[30] Bob G. Witmer and Michael J. Singer. "Measuring Presence in Virtual Environments: A Presence Questionnaire." In: *Presence* 7.3 (1998), pp. 225–240.

[31] T. Bär, M. Riedl, B. Kärcher, H. Hämmerle, H. Beesten, M. Pfeil, and F. Damrath. *Physikbasierte Simulation Zur Unterstützung Der Virtuellen Inbetriebnahme*. 00000. 2014.

[32]   Sebastian Süß, Dominik Hauf, Anton Strahilov, and Christian Diedrich. "Standardized Classification and Interfaces of Complex Behaviour Models in Virtual Commissioning." In: *Procedia CIRP*. The Sixth International Conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV2016) 52 (2016). 00000, pp. 24–29. DOI: 10.1016/j.procir.2016.07.049.

[33]   Achille Mbang Sama. *Holistic integration of product, process and resources integration in the automotive industry using the example of car body design and production : product design, process modeling, IT implementation and potential benefits*. 2008. DOI: 10.5445/KSP/1000008181.

[34]   T. Deviprasad and T. Kesavadas. "Virtual Prototyping of Assembly Components Using Process Modeling." In: *Journal of Manufacturing Systems* 22.1 (Jan. 2003), pp. 16–27. DOI: 10.1016/S0278-6125(03)90002-1.

[35]   C.I.R.P., ed. *Dictionary of Production Engineering/Wörterbuch Der Fertigungstechnik/Dictionnaire Des Techniques de Production Mechanique Vol IV*. 00000. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. URL: http://link.springer.com/10.1007/978-3-642-12007-7 (visited on 08/04/2016).

[36]   John Stark. "Product Lifecycle Management." In: *Product Lifecycle Management*. 00696. Cham: Springer International Publishing, 2015, pp. 1–29. URL: http://link.springer.com/10.1007/978-3-319-17440-2_1 (visited on 08/04/2016).

[37]   *VDI 2221 - Systematic Approach to the Development and Design of Technical Systems and Products*. Mar. 1993. URL: https://www.beuth.de/de/technische-regel/vdi-2221/973992 (visited on 02/10/2019).

[38]   Munro & Associates, Inc. *History of Lean Design® | Munro & Associates, Inc.* 00000. 2016. URL: http://leandesign.com/history-of-lean-design/ (visited on 08/18/2016).

[39]   Bruno Lotter and Hans-Peter Wiendahl, eds. *Montage in der industriellen Produktion*. 00127. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. URL: http://link.springer.com/10.1007/978-3-642-29061-9 (visited on 04/15/2016).

[40]   Geoffrey Boothroyd, Peter Dewhurst, and Winston A. Knight. *Product Design for Manufacture and Assembly, Third Edition*. 00003. CRC Press, Dec. 8, 2010. 710 pp. Google Books: W2FDCcVPBcAC.

[41]   Christoph Runde, Florin Girbacia, and Eugen Butila. "Virtual and Augmented Environments for Concurrent Engineering: Concurrent Virtual Engineering." In: *Concurrent Engineering Approaches for Sustainable Product Development in a Multi-Disciplinary Environment*. Ed. by Josip Stjepandić, Georg Rock, and Cees

Bil. 00000. London: Springer London, 2013, pp. 849–860. URL: http://link.springer.com/10.1007/978-1-4471-4426-7_72 (visited on 01/12/2016).

[42] Geoffrey Boothroyd. "Product Design for Manufacture and Assembly." In: *Computer-Aided Design* 26.7 (July 1, 1994). 01912, pp. 505–520. DOI: 10.1016/0010-4485(94)90082-5.

[43] Oliver Geißel. "AMMU Automotive Mixed Mock-Up: Konzeption Einer Neuen Entwicklungsplattform Für Die Automobilindustrie." 00003. 2012. URL: http://elib.uni-stuttgart.de/opus/volltexte/2012/7191/ (visited on 12/02/2015).

[44] Jörg Roscher. "Bewertung von Flexibilitätsstrategien Für Die Endmontage in Der Automobilindustrie." 00026. Universität Stuttgart, 2008.

[45] G. Lay and E. Schirrmeister. *Sackgasse Hochautomatisierung?* Ed. by Karlsruhe Fraunhofer-Institut für Systemtechnik und Innovationsforschung. Mitteilungen Aus Der Produktionsinnovationserhebung. PI-Mitteilungen. Karlsruhe: ISI, 2000. 12 pp.

[46] Bruno Lotter. *Manufacturing Assembly Handbook - 1st Edition*. 1st ed. Elsevier, 1989. URL: https://www.elsevier.com/books/manufacturing-assembly-handbook/lotter/978-0-408-03561-3 (visited on 02/15/2019).

[47] József Váncza. "Production Networks." In: *CIRP Encyclopedia of Production Engineering*. Ed. by The International Academy for Produ, Luc Laperrière, and Gunther Reinhart. 00000. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 1–8. URL: http://link.springer.com/10.1007/978-3-642-35950-7_16829-1 (visited on 08/05/2016).

[48] Peter Butala and Khumbulani Mpofu. "Assembly Systems." In: *CIRP Encyclopedia of Production Engineering*. Ed. by The International Academy for Produ, Luc Laperrière, and Gunther Reinhart. 00000. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 1–5. URL: http://link.springer.com/10.1007/978-3-642-35950-7_16789-1 (visited on 08/05/2016).

[49] Lars Weyand. "Risikoreduzierte Endmontageplanung am Beispiel der Automobilindustrie." 2010. DOI: http://dx.doi.org/10.22028/D291-22651.

[50] Company Ford. *Innovation: 100 Years of the Moving Assembly Line | Ford Motor Company*. 2019. URL: http://corporate.ford.com/innovation/100-years-moving-assembly-line.html (visited on 02/15/2019).

[51] Audi Media Center. *The Home of the New A3 Sedan – Audi's Expands Its Site in Győr into a Fully-Fledged Car Plant*. 2019. URL: https://www.audi-mediacenter.com:443/en/audi-expands-its-hungarian-site-in-gyor-into-a-fully-fledged-car-factory-2988/the-home-of-the-new-a3-sedan-audis-expands-its-site-in-gyor-into-a-fully-fledged-car-plant-2999 (visited on 01/26/2019).

[52] Verein Deutscher Ingenieure VDI. *VDI-Richtlinie: VDI 2860 Blatt 1 Montage- und Handhabungstechnik - Handhabungsfunktionen, Handhabungseinrichtungen, Begriffe, Definitionen, Symbole.* URL: /richtlinie/vdi_2860-montage_und_handhabungstechnik_handhabungsfunktionen_handhabungseinrichtungen_begriffe_definitionen/ (visited on 01/27/2019).

[53] DIN 8593-0 DIN. *DIN 8593-0 Manufacturing Processes Joining - Part 0: General; Classification, Subdivision, Terms and Definitions.* Sept. 2003. URL: https://www.beuth.de/de/norm/din-8593-0/65031206 (visited on 01/27/2019).

[54] Christian Küber, Engelbert Westkämper, Bernd Keller, and Hans-Friedrich Jacobi. "Planning Method for the Design of Flexible as Well as Economic Assembly and Logistics Processes in the Automotive Industry." In: *Procedia CIRP*. Research and Innovation in Manufacturing: Key Enabling Technologies for the Factories of the Future - Proceedings of the 48th CIRP Conference on Manufacturing Systems 41 (Jan. 1, 2016). 00000, pp. 556–561. DOI: 10.1016/j.procir.2015.12.038.

[55] Standard. *DIN IEC 60050-351:2014-09, Internationales Elektrotechnisches Wörterbuch_- Teil_351: Leittechnik (IEC_60050-351:2013)*. Beuth Verlag GmbH, Sept. 2014. DOI: 10.31030/2159569.

[56] T. Fujimoto. "What Do You Mean by Automation Ratio? Definitions by the Japanese Auto Makers." In: *Transforming Automobile Assembly*. Ed. by Koichi Shimokawa, Ulrich Jürgens, and Takahiro Fujimoto. 00000. Berlin, Heidelberg: Springer Berlin Heidelberg, 1997, pp. 61–69. URL: http://link.springer.com/10.1007/978-3-642-60374-7_5 (visited on 08/03/2016).

[57] Walter Eversheim. *Organisation in Der Produktionstechnik Band 4: Fertigung Und Montage*. 2nd ed. VDI-Buch. Berlin Heidelberg: Springer-Verlag, 1989. URL: //www.springer.com/de/book/9783642648007 (visited on 01/27/2019).

[58] Günther Pawellek. *Ganzheitliche Fabrikplanung*. 00081. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. URL: http://link.springer.com/10.1007/978-3-662-43728-5 (visited on 01/04/2016).

[59]   Giuseppe Stecca. "Production Planning." In: *CIRP Encyclopedia of Production Engineering*. Ed. by The International Academy for Produ, Luc Laperrière, and Gunther Reinhart. 00000. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 1–4. URL: http://link.springer.com/10.1007/978-3-642-35950-7_6569-4 (visited on 08/05/2016).

[60]   Claus-Gerold Grundig. *Fabrikplanung: Planungssystematik - Methoden - Anwendungen*. Hanser, 2000. 319 pp. Google Books: 0WFhPQAACAAJ.

[61]   Mahmut Gülesjn. "10 - Process Planning in Design and Manufacturing Systems." In: *Expert Systems*. Ed. by Cornelius T. Leondes. Burlington: Academic Press, Jan. 1, 2002, pp. 327–380. DOI: 10.1016/B978-012443880-4/50054-5.

[62]   Hoda A. ElMaraghy. "Evolution and Future Perspectives of CAPP." In: *CIRP Annals* 42.2 (Jan. 1, 1993), pp. 739–751. DOI: 10.1016/S0007-8506(07)62537-2.

[63]   The International Academy for Production Engineering, Luc Laperrière, and Gunther Reinhart, eds. *CIRP Encyclopedia of Production Engineering*. 00000. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. URL: http://link.springer.com/10.1007/978-3-642-20617-7 (visited on 01/18/2016).

[64]   Bruno Lotter. *Wirtschaftliche Montage: ein Handbuch für Elektrogerätebau und Feinwerktechnik*. VDI-Verlag, 1992. 458 pp. Google Books: LC8JHAAACAAJ.

[65]   Marko Hartel and Bruno Lotter. "Planung und Bewertung von Montagesystemen." In: *Montage in der industriellen Produktion* (2006), pp. 407–432. DOI: 10.1007/3-540-36669-5_13.

[66]   Martin Manns and Nestor Andres Arteaga Martin. "A Vagueness Measure for Concurrent Engineering of Manual Assembly Lines." In: *Enabling Manufacturing Competitiveness and Economic Sustainability*. Ed. by Michael F. Zaeh. Springer International Publishing, Jan. 1, 2014, pp. 131–135. URL: http://link.springer.com/chapter/10.1007/978-3-319-02054-9_22 (visited on 06/21/2014).

[67]   Arbeitswissenschaft. *Ergonomic Assessment Worksheet (EAWS)*. 00000. URL: http://www.iad.tu-darmstadt.de/forschung_15/methodenundlabore/ergonomic_assessment_worksheet_eaws.de.jsp (visited on 07/13/2016).

[68]   Philipp Agethen, Felix Gaisbauer, Michael Otto, and Enrico Rukzio. "Interactive Simulation for Walk Path Planning within the Automotive Industry." In: *Proceedings of the 51st CIRP Conference on Manufacturing Systems*. CIRP CMS 2018. Vol. 72. Stockholm, Sweden, Jan. 1, 2018, pp. 285–290. DOI: 10.1016/j.procir.2018.03.223.

[69]   Nima Rafibakhsh. "Automated Assembly Planning: From CAD Model to Virtual Assembly Process." In: (2017).

[70]   Ming C. Leu, Hoda A. ElMaraghy, Andrew YC Nee, Soh Khim Ong, Michele Lanzetta, Matthias Putz, Wenjuan Zhu, and Alain Bernard. "CAD Model Based Virtual Assembly Simulation, Planning and Training." In: *CIRP Annals* 62.2 (2013), pp. 799–822.

[71]   Merriam-Webster. *Definition of Mock–Up*. URL: http://www.merriam-webster.com/dictionary/mock%E2%80%93up (visited on 01/07/2016).

[72]   Carlos Enrique Gorbea Diaz. "Vehicle Architecture and Life-cycle Cost Analysis." 00008. Technische Universität München, 2011. URL: http://mediatum.ub.tum.de/?id=1108157 (visited on 08/04/2016).

[73]   Gerd F. Kamiske and Jörg-Peter Brauer. *Qualitätsmanagement von A bis Z: wichtige Begriffe des Qualitätsmanagements und ihre Bedeutung*. 00012. Hanser, 2011. 402 pp.

[74]   Fan Dai. *Virtual Reality for Industrial Applications*. 00124. Heidelberg: Springer Science & Business Media, 2008. 177 pp.

[75]   Abhishek Seth, Judy M. Vance, and James H. Oliver. "Virtual Reality for Assembly Methods Prototyping: A Review." In: *Virtual Reality* 15.1 (Jan. 22, 2010). 00095, pp. 5–20. DOI: 10.1007/s10055-009-0153-y.

[76]   Andreas Syska. *Produktionsmanagement: Das A - Z wichtiger Methoden und Konzepte für die Produktion von heute*. Springer-Verlag, Aug. 15, 2006. 192 pp.

[77]   M. Mousavi, F. A. Aziz, and N. Ismail. "Opportunities and Constraints of Virtual Reality Application in International and Domestic Car Companies of Malaysia." In: *2012 UKSim 14th International Conference on Computer Modelling and Simulation*. 2012 UKSim 14th International Conference on Computer Modelling and Simulation. Mar. 2012, pp. 273–277. DOI: 10.1109/UKSim.2012.46.

[78]   Karl-Josef Wack, Steffen Strassburger, and Thomas Bär. "Virtual Validation of Material Provisioning in Assembly in the Automotive Industry." In: *Proceedings of the 44th International CIRP Conference on Manufacturing Systems. Madison, Wisconsin, USA*. 00001. 2011. URL: http://www.strassburger-online.de/papers/cirp44_virtual%20validation%20of%20material.pdf (visited on 08/03/2016).

[79]   *3P – Prozess Der Produktionsvorbereitung | Volkswagen*. 00000. URL: http://www.volkswagen-poznan.pl/de/3p-prozess-der-produktionsvorbereitung (visited on 04/04/2017).

[80]   Michael Otto, Michael Prieur, Philipp Agethen, and Enrico Rukzio. "Dual Reality for Production Verification Workshops: A Comprehensive Set of Virtual Methods." In: *Proceedings of the 6th CIRP Conference on Assembly Technologies and Systems*. CIRP CATS 2016. Vol. 44. CIRP CATS 2016. Gothenburg, Sweden, 2016, pp. 38–43. DOI: 10.1016/j.procir.2016.02.140.

[81]   Jan C. Aurich, Dirk Ostermayer, and C. H. Wagenknecht. "Improvement of Manufacturing Processes with Virtual Reality-Based CIP Workshops." In: *International Journal of Production Research* 47.19 (2009), pp. 5297–5309. URL: http://www.tandfonline.com/doi/abs/10.1080/00207540701816569 (visited on 06/25/2014).

[82]   J. C. Aurich, H. Hagen, D. Ostermayer, and M. Bertram. "VR-Unterstützter KVP-Workshop." In: *wt online* 95 (). URL: http://www-scheuer.informatik.uni-kl.de/~bertram/papers/VR_KVP.pdf (visited on 06/25/2014).

[83]   ISO Organisation Internationale de Normalisation. *ISO 9000:2015(En), Quality Management Systems — Fundamentals and Vocabulary*. URL: https://www.iso.org/obp/ui/#iso:std:iso:9000:ed-4:v1:en (visited on 02/17/2019).

[84]   Osman Balci. "Verification, Validation, and Certification of Modeling and Simulation Applications." In: *Proceedings of the 35th Conference on Winter Simulation: Driving Innovation*. Winter Simulation Conference, 2003, pp. 150–158.

[85]   Robert Meißner. *Digitale Absicherung der Montagetauglichkeit: ein Beitrag zur Integration von Produktentwicklung und Montageplanung*. Otto-von-Guericke-Univ., Lehrstuhl für Maschinenbauinformatik, 2010. 189 pp. Google Books: ZvDDZwEACAAJ.

[86]   He Wang, Xiaowei Zhu, Hui Wang, S. Jack Hu, Zhongqin Lin, and Guanlong Chen. "Multi-Objective Optimization of Product Variety and Manufacturing Complexity in Mixed-Model Assembly Systems." In: *Journal of Manufacturing Systems* 30.1 (Jan. 2011). 00039, pp. 16–27. DOI: 10.1016/j.jmsy.2011.03.002.

[87]   Annki Falck, Roland Örtengren, and Dan Högberg. "Early Risk Identification and Cost-Benefit Analyses through Ergonomics Simulation." In: (2009).

[88]   Karl-Josef Wack, Franz Otto, Martin Manns, and Steffen Strassburger. "Efficient Preparation of Digital Production Validation." In: (). 00000.

[89]   Ulrich Seiffert and Gotthard Rainer. *Virtuelle Produktentstehung Für Fahrzeug Und Antrieb Im KFZ: Prozesse, Komponenten, Beispiele Aus Der Praxis*. Springer, 2008.

[90]   Christoph Picker. *Prospektive Zeitbestimmung Für Nicht Wertschöpfende Montagetätigkeiten*. 00010. 2007. URL: http://dx.doi.org/10.17877/DE290R-883 (visited on 05/22/2016).

[91]   REFA. *REFA Methodenlehre der Betriebsorganisation, Datenermittlung*. 00000. München: Fachbuchverlag Leipzig, July 28, 1997. 456 pp.

[92]   Ulrich Bockholt, Thomas Bochtler, Volker Enderlein, Manuel Olbrich, Michael Otto, Michael Prieur, Richard Sauerbier, Roland Stechow, Andreas Wehe, and Hans-Joachim Wirsching. "Motion Capturing." In: *Web-basierte Anwendungen Virtueller Techniken*. Ed. by Werner Schreiber, Konrad Zürl, and Peter Zimmermann. Berlin, Heidelberg: Springer Vieweg, Berlin, Heidelberg, 2017, pp. 219–262. DOI: 10.1007/978-3-662-52956-0_5.

[93]   Sören Kunst. *Konzeption eines Modells zur Bewertung der Wirtschaftlichkeit von Virtual Reality Systemen in der Digitalen Fabrik*. Apr. 23, 2006. 127 pp. Google Books: 4VZoAQAAQBAJ.

[94]   Stefan H Thomke. "Simulation, Learning and R&D Performance: Evidence from Automotive Development." In: *Research Policy* 27.1 (May 1, 1998). 00176, pp. 55–74. DOI: 10.1016/S0048-7333(98)00024-9.

[95]   Martin Manns and Nestor Andres Arteaga Martin. "Improving A* Walk Trajectories with B-Splines and Motion Capture for Manual Assembly Verification." In: *Procedia CIRP*. 9th CIRP Conference on Intelligent Computation in Manufacturing Engineering - CIRP ICME '14 33 (Jan. 1, 2015), pp. 364–369. DOI: 10.1016/j.procir.2015.06.083.

[96]   Reuters. *Tesla Is Skipping Traditional Manufacturing Steps for Model 3*. Apr. 24, 2017. URL: http://www.newsweek.com/tesla-skips-traditional-manufacturing-steps-model-3-588503 (visited on 04/28/2017).

[97]   Z. Szalavári, D. Schmalstieg, A. Fuhrmann, and M. Gervautz. ""Studierstube": An Environment for Collaboration in Augmented Reality." In: *Virtual Real.* 3.1 (Mar. 1998). 00342, pp. 37–48. DOI: 10.1007/BF01409796.

[98]   Paul Milgram and Fumio Kishino. "A Taxonomy of Mixed Reality Visual Displays." In: *IEICE TRANSACTIONS on Information and Systems* 77.12 (1994). 02041, pp. 1321–1329. URL: http://search.ieice.org/bin/summary.php?id=e77-d_12_1321 (visited on 05/18/2015).

[99]   Joshua Lifton and Joseph A. Paradiso. "Dual Reality: Merging the Real and Virtual." In: *Facets of Virtual Environments*. Ed. by Fritz Lehmann-Grube and Jan Sablatnig. Red. by Ozgur Akan,

Paolo Bellavista, Jiannong Cao, Falko Dressler, Domenico Ferrari, Mario Gerla, Hisashi Kobayashi, Sergio Palazzo, Sartaj Sahni, Xuemin (Sherman) Shen, Mircea Stan, Jia Xiaohua, Albert Zomaya, and Geoffrey Coulson. Vol. 33. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 12–28. URL: http://link.springer.com/10.1007/978-3-642-11743-5_2 (visited on 09/15/2015).

[100]   Andre Stork. "Visual Computing Challenges of Advanced Manufacturing and Industrie 4.0 [Guest Editors' Introduction]." In: *IEEE Computer Graphics and Applications* 35.2 (Mar. 2015), pp. 21–25. DOI: 10.1109/MCG.2015.46.

[101]   Monica Bordegoni, Umberto Cugini, Giandomenico Caruso, and Samuele Polistina. "Mixed Prototyping for Product Assessment: A Reference Framework." In: *International Journal on Interactive Design and Manufacturing (IJIDeM)* 3.3 (July 12, 2009), pp. 177–187. DOI: 10.1007/s12008-009-0073-9.

[102]   Hiroshi Ishii. "Tangible Bits: Beyond Pixels." In: *Proceedings of the 2nd International Conference on Tangible and Embedded Interaction*. ACM, 2008, pp. xv–xxv. URL: http://dl.acm.org/citation.cfm?id=1347392 (visited on 06/12/2014).

[103]   "Collaborative Virtual Environments." In: *Encyclopedia of Multimedia*. Ed. by Borko Furht. 00000. Boston, MA: Springer US, 2008, pp. 74–75. URL: http://www.springerlink.com/index/10.1007/978-0-387-78414-4_235 (visited on 12/15/2016).

[104]   Dave Snowdon, Elizabeth F. Churchill, and Alan J. Munro. "Collaborative Virtual Environments: Digital Spaces and Places for CSCW: An Introduction." In: *Collaborative Virtual Environments*. 00081. Springer, 2001, pp. 3–17. URL: http://link.springer.com/chapter/10.1007/978-1-4471-0685-2_1 (visited on 12/15/2016).

[105]   Mark Billinghurst, Ivan Poupyrev, Hirokazu Kato, and Richard May. "Mixing Realities in Shared Space: An Augmented Reality Interface for Collaborative Computing." In: *IEEE International Conference on Multimedia and Expo (III)*. Citeseer, 2000, pp. 1641–1644.

[106]   Christian Mai, Sarah Bartsch, and Lea Rieger. "Evaluating Shared Surfaces for Co-Located Mixed-Presence Collaboration | Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia." In: MUM 2018: Proceedings of the 17th International Conference on Mobile and Ubiquitous. 2018, pp. 1–5. DOI: https://doi.org/10.1145/3282894.3282910.

[107]   Eike Langbehn, Eva Harting, and Frank Steinicke. "Shadow-Avatars: A Visualization Method to Avoid Collisions of Physically Co-Located Users in Room-Scale VR." In: IEEE Workshop on Everyday Virtual Reality. 2018.

[108]   Hrvoje Benko, Eyal Ofek, Feng Zheng, and Andrew D. Wilson. "Fovear: Combining an Optically See-through near-Eye Display with Projector-Based Spatial Augmented Reality." In: *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. ACM, 2015, pp. 129–135.

[109]   Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. "ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users." In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017, pp. 4021–4033.

[110]   Michael Otto, Michael Prieur, and Enrico Rukzio. "Using Scalable, Interactive Floor Projection for Production Planning Scenario." In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*. ACM ITS 2014. Dresden, Germany: Association for Computing Machinery, Nov. 16, 2014, pp. 363–368. DOI: 10.1145/2669485.2669547.

[111]   Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. "RoomAlive: Magical Experiences Enabled by Scalable, Adaptive Projector-Camera Units." In: *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*. UIST '14. New York, NY, USA: ACM, 2014, pp. 637–644. DOI: 10.1145/2642918.2647383.

[112]   Julian Seifert, David Dobbelstein, Dominik Schmidt, Paul Holleis, and Enrico Rukzio. "From the Private into the Public: Privacy-Respecting Mobile Interaction Techniques for Sharing Data on Surfaces." In: *Personal and Ubiquitous Computing* 18.4 (2014), pp. 1013–1026.

[113]   Ricardo Langner, Ulrich von Zadow, Tom Horak, Annett Mitschick, and Raimund Dachselt. "Content Sharing Between Spatially-Aware Mobile Phones and Large Vertical Displays Supporting Collaborative Work." In: *Collaboration Meets Interactive Spaces*. Springer, 2016, pp. 75–96.

[114]   Christian Winkler, Markus Löchtefeld, David Dobbelstein, Antonio Krüger, and Enrico Rukzio. "SurfacePhone: A Mobile Projection Device for Single-and Multiuser Everywhere Tabletop Interaction." In: (2014).

[115]   Ferran Argelaguet, Andre Kunert, Alexander Kulik, and Bernd Froehlich. "Improving Co-Located Collaboration with Show-through Techniques." In: *3D User Interfaces (3DUI), 2010 IEEE Symposium On*. 00014. IEEE, 2010, pp. 55–62. URL: http://ieeexplore.ieee.org/abstract/document/5444719/ (visited on 08/29/2017).

[116]   Steve Ballmer. *CES 2010: A Transforming Trend – The Natural User Interface*. Mar. 18, 2010. URL: https://www.huffingtonpost.com/steve-ballmer/ces-2010-a-transforming-t_b_416598.html (visited on 01/09/2019).

[117]   Daniel Wigdor and Dennis Wixon. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. 1st. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011.

[118]   JoshB. *What Is the Natural User Interface? (Book Excerpt)*. URL: http://nui.joshland.org/2010/03/what-is-natural-user-interface-book.html (visited on 01/09/2019).

[119]   Gabriel Zachmann and Alexander Rettig. "Natural and Robust Interaction in Virtual Assembly Simulation." In: *Eighth ISPE International Conference on Concurrent Engineering: Research and Applications (ISPE/CE2001)*. Vol. 1. 00080. 2001, pp. 425–434. URL: http://cgvr.cs.uni-bremen.de/papers/ce2001/ce2001.pdf (visited on 08/28/2017).

[120]   Bernhard Preim and Raimund Dachselt. *Interaktive Systeme: Band 2: User Interface Engineering, 3D-Interaktion, Natural User Interfaces*. Springer-Verlag, Apr. 2, 2015. 786 pp. Google Books: nwPSBwAAQBAJ.

[121]   Tomasz Miaskiewicz and Kenneth A. Kozar. "Personas and User-Centered Design: How Can Personas Benefit Product Design Processes?" In: *Design Studies* 32.5 (Sept. 1, 2011), pp. 417–430. DOI: 10.1016/j.destud.2011.03.003.

[122]   Joe Cecil and Atipol Kanchanapiboon. "Virtual Engineering Approaches in Product and Process Design." In: *The International Journal of Advanced Manufacturing Technology* 31.9-10 (2007), pp. 846–856.

[123]   Martin Manns and Néstor Andrés Arteaga Martín. "Automated DHM Modeling for Integrated Alpha-Numeric and Geometric Assembly Planning." In: *Smart Product Engineering* (2013), pp. 325–334. DOI: 10.1007/978-3-642-30817-8_32.

[124]   SangSu Choi, Kiwook Jung, and Sang Do Noh. "Virtual Reality Applications in Manufacturing Industries: Past Research, Present Findings, and Future Directions." In: *Concurrent Engineering* 23.1 (2015), pp. 40–63.

[125]   Holger Graf, Gino Brunetti, and André Stork. "A Methodology Supporting the Preparation of 3D-CAD Data for Design Reviews in VR." In: *DS 30: Proceedings of DESIGN 2002, the 7th International Design Conference, Dubrovnik*. 2002, pp. 489–496.

[126]   Joseph J. LaViola Jr. "A Discussion of Cybersickness in Virtual Environments." In: *ACM SIGCHI Bulletin* 32.1 (2000), pp. 47–56.

[127]   Matthias Karl and Johannes Behr. "Visualization as a Service (VAAS)." May 2017. URL: https://www.prostep-ivip-symposium.org/fileadmin/Veranstaltungen/symposium17/Presentations/Presentation_Karl-BMW_Behr-Fraunhofer-IGD.pdf (visited on 11/03/2018).

[128]   3DInteractive. *Technology - 3DInteractive*. URL: https://www.3dinteractive.de/about-3dinteractive/technology/ (visited on 11/05/2018).

[129]   NetAllied Systems GmbH. *NetAllied Systems GmbH*. URL: https://netallied.de/dmu4all.html (visited on 11/05/2018).

[130]   Johannes Behr, Christophe Mouton, Samuel Parfouru, Julien Champeau, Clotilde Jeulin, Maik Thöner, Christian Stein, Michael Schmitt, Max Limper, Miguel de Sousa, Tobias Alexander Franke, and Gerrit Voss. "webVis/Instant3DHub: Visual Computing As a Service Infrastructure to Deliver Adaptive, Secure and Scalable User Centric Data Visualisation." In: *Proceedings of the 20th International Conference on 3D Web Technology*. Web3D '15. New York, NY, USA: ACM, 2015, pp. 39–47. DOI: 10.1145/2775292.2775299.

[131]   AutomationML e.V. *AutomationML*. URL: http://www.automationml.org/o.red.c/home.html (visited on 04/23/2018).

[132]   Werner Schreiber, Konrad Zürl, and Peter Zimmermann, eds. *Web-Basierte Anwendungen Virtueller Techniken: Das ARVIDA-Projekt – Dienste-Basierte Software-Architektur Und Anwendungsszenarien Für Die Industrie*. Springer Vieweg, 2017. URL: //www.springer.com/de/book/9783662529553 (visited on 09/23/2018).

[133]   Infor. *Engineering BOM for Assembly Production*. 2018. URL: https://docs.infor.com/ln/10.5/en-us/lnolh/help/ti/onlinemanual/000519.html (visited on 04/23/2018).

[134]   ISO. *ISO 14306:2012 - Industrial Automation Systems and Integration – JT File Format Specification for 3D Visualization*. 00000. Dec. 2012. URL: https://www.iso.org/standard/60572.html (visited on 10/07/2017).

[135]   A2Mac1. *Benchmarking Contents*. Apr. 2019. URL: https://portal.a2mac1.com/benchmarking-database/ (visited on 04/07/2019).

[136]   Liang Gong, Jonatan Berglund, Dennis Saluäär, and Björn Johansson. "A Novel VR Tool for Collaborative Planning of Manufacturing Process Change Using Point Cloud Data." In: *Procedia CIRP*. Manufacturing Systems 4.0 – Proceedings of the 50th CIRP Conference on Manufacturing Systems 63 (2017). 00000, pp. 336–341. DOI: 10.1016/j.procir.2017.03.089.

[137]   Erik Lindskog. "Towards Realistic Visualisation of Production Systems." Fraunhofer Chalmers Centre, 2014.

[138]   E. Shellshear, R. Berlin, and J.S. Carlson. "Maximizing Smart Factory Systems by Incrementally Updating Point Clouds." In: *IEEE Computer Graphics and Applications* 35.2 (Mar. 2015), pp. 62–69. DOI: 10.1109/MCG.2015.38.

[139]   M. Heyer and B. Brüderlin. "Visibility-Guided Rendering for Visualizing Very Large Virtual Reality Scenes." In: *Proc. 1st GI Workshop Virtual and Augmented Reality (VR/AR 04), Shaker Verlag*. 2004, pp. 163–172.

[140]   André Antakli, Pablo Alvarado Moya, Beat Brüderlin, Ulrich Canzler, Holger Dammertz, Volker Enderlein, Jürgen Grüninger, Andreas Harth, Hilko Hoffmann, Eduard Jundt, Peter Keitler, Felix Leif Keppmann, Roland Krzikalla, Sebastian Lampe, Alexander Löffler, Julian Meder, Michael Otto, Frieder Pankratz, Sebastian Pfützner, Matthias Roth, Richard Sauerbier, Werner Schreiber, Roland Stechow, Johannes Tümler, Christian Vogelgesang, Oliver Wasenmüller, Andreas Weinmann, Jochen Willneff, Hans-Joachim Wirsching, Ingo Zinnikus, and Konrad Zürl. "Virtuelle Techniken und Semantic-Web." In: *Webbasierte Anwendungen Virtueller Techniken*. Ed. by Werner Schreiber, Konrad Zürl, and Peter Zimmermann. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, pp. 17–116. DOI: 10.1007/978-3-662-52956-0_2.

[141]   B. Brüderlin, M. Heyer, and S. Pfützner. "Interviews3D: A Platform for Interactive Handling of Massive Data Sets." In: *IEEE Computer Graphics and Applications* 27.6 (Nov. 2007), pp. 48–59. DOI: 10.1109/MCG.2007.153.

[142]   M. Heyer and B. Brüderlin. "Hardware-Supported Occlusion Culling for Visualizing Extremely Large Models." In: *Proc. 3, Paderborner Workshop, Paderborn Augmented & Virtual Reality in Der Produktentstehung, Hanser Verlag*. 2004, pp. 39–49.

[143]   Russell M. Taylor, Thomas C. Hudson, Adam Seeger, Hans Weber, Jeffrey Juliano, and Aron T. Helser. "VRPN: A Device-Independent, Network-Transparent VR Peripheral System." In: *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. ACM, 2001, pp. 55–61.

[144]   Don B. Chaffin. "Human Motion Simulation for Vehicle and Workplace Design." In: *Human Factors and Ergonomics in Manufacturing & Service Industries* 17.5 (2007), pp. 475–484. DOI: `10.1002/hfm.20087`.

[145]   *MakeHuman | Open Source Tool for Making 3d Characters*. URL: `http://www.makehuman.org/` (visited on 05/02/2018).

[146]   H-Anim. *ISO/IEC FCD 19774:200x – Humanoid Animation Index*. URL: `http://h-anim.org/Specifications/H-Anim200x/ISO_IEC_FCD_19774/` (visited on 03/17/2015).

[147]   Daniel Roth, Jean-Luc Lugrin, Julia Buser, Gary Bente, Arnulph Fuhrmann, and Marc Erich Latoschik. "A Simplified Inverse Kinematic Approach for Embodied VR Applications." In: *2016 IEEE Virtual Reality (VR)*. 2016 IEEE Virtual Reality (VR). Greenville, SC, USA: IEEE, Mar. 2016, pp. 275–276. DOI: `10.1109/VR.2016.7504760`.

[148]   Eiji Hayashi, Manuel Maas, and Jason I. Hong. "Wave to Me: User Identification Using Body Lengths and Natural Gestures." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. 00008. New York, NY, USA: ACM, 2014, pp. 3453–3462. DOI: `10.1145/2556288.2557043`.

[149]   Ivan E. Sutherland. "A Head-Mounted Three Dimensional Display." In: *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*. ACM, 1968, pp. 757–764.

[150]   D.A. Bowman and R.P. McMahan. "Virtual Reality: How Much Immersion Is Enough?" In: *Computer* 40.7 (July 2007). 00216, pp. 36–43. DOI: `10.1109/MC.2007.257`.

[151]   Helena Van Kerrebroeck, Malaika Brengman, and Kim Willems. "Escaping the Crowd: An Experimental Study on the Impact of a Virtual Reality Experience in a Shopping Mall." In: *Computers in Human Behavior* 77 (Dec. 1, 2017), pp. 437–450. DOI: `10.1016/j.chb.2017.07.019`.

[152]   Kate E. Laver, Stacey George, Susie Thomas, Judith E. Deutsch, and Maria Crotty. "Virtual Reality for Stroke Rehabilitation." In: *Cochrane Database of Systematic Reviews* 2 (2015). DOI: `10.1002/14651858.CD008349.pub3`.

[153]   Abbie Brown and Tim Green. "Virtual Reality: Low-Cost Tools and Resources for the Classroom." In: *TechTrends* 60.5 (2016), pp. 517–519.

[154]   Zahira Merchant, Ernest T. Goetz, Lauren Cifuentes, Wendy Keeney-Kennicutt, and Trina J. Davis. "Effectiveness of Virtual Reality-Based Instruction on Students' Learning Outcomes in K-12 and Higher Education: A Meta-Analysis." In: *Computers & Education* 70 (Jan. 1, 2014), pp. 29–40. DOI: `10.1016/j.compedu.2013.07.033`.

[155] P. Zimmermann and A. G. Volkswagen. "Virtual Reality Aided Design: A Survey of the Use." In: 2008.

[156] Stig Ottosson. "Virtual Reality in the Product Development Process." In: *Journal of Engineering Design* 13.2 (June 1, 2002), pp. 159–172. DOI: 10.1080/09544820210129823.

[157] Glyn Lawson, Davide Salanitri, and Brian Waterfield. "Future Directions for the Development of Virtual Reality within an Automotive Manufacturer." In: *Applied Ergonomics*. Transport in the 21st Century: The Application of Human Factors to Future User Needs 53 (Mar. 1, 2016), pp. 323–330. DOI: 10.1016/j.apergo.2015.06.024.

[158] Leif P. Berg and Judy M. Vance. "Industry Use of Virtual Reality in Product Design and Manufacturing: A Survey." In: *Virtual Reality* 21.1 (Mar. 1, 2017), pp. 1–17. DOI: 10.1007/s10055-016-0293-9.

[159] Andreas Dünser, Karin Steinbügl, Hannes Kaufmann, and Judith Glück. "Virtual and Augmented Reality As Spatial Ability Training Tools." In: *Proceedings of the 7th ACM SIGCHI New Zealand Chapter's International Conference on Computer-Human Interaction: Design Centered HCI*. CHINZ '06. 00073. New York, NY, USA: ACM, 2006, pp. 125–132. DOI: 10.1145/1152760.1152776.

[160] A. Stork, N. Sevilmis, D. Weber, D. Gorecky, C. Stahl, M. Loskyll, and F. Michel. "Enabling Virtual Assembly Training in and beyond the Automotive Industry." In: *2012 18th International Conference on Virtual Systems and Multimedia (VSMM)*. 2012 18th International Conference on Virtual Systems and Multimedia (VSMM). Sept. 2012, pp. 347–352. DOI: 10.1109/VSMM.2012.6365944.

[161] Charlene Jennett, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. "Measuring and Defining the Experience of Immersion in Games." In: *International journal of human-computer studies* 66.9 (2008), pp. 641–661.

[162] Stephen R. Ellis. "Presence of Mind: A Reaction to Thomas Sheridan's "Further Musings on the Psychophysics of Presence"." In: *Presence: Teleoperators & Virtual Environments* 5.2 (1996), pp. 247–259.

[163] Markus Funk, Thomas Kosch, Scott W. Greenwald, and Albrecht Schmidt. "A Benchmark for Interactive Augmented Reality Instructions for Assembly Tasks." In: *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*. MUM '15. New York, NY, USA: ACM, 2015, pp. 253–257. DOI: 10.1145/2836041.2836067.

[164]	John Brooke. "SUS: A Retrospective." In: *Journal of usability studies* 8.2 (2013), pp. 29–40.

[165]	Diederick C. Niehorster, Li Li, and Markus Lappe. "The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research." In: *i-Perception* 8.3 (June 1, 2017), p. 2041669517708205. DOI: `10.1177/2041669517708205`.

[166]	Shinji Umeyama. "Least-Squares Estimation of Transformation Parameters between Two Point Patterns." In: *IEEE Transactions on pattern analysis and machine intelligence* 13.4 (1991), pp. 376–380.

[167]	Lars Fritzsche. "Ergonomics Risk Assessment with Digital Human Models in Car Assembly: Simulation versus Real Life." In: *Human Factors and Ergonomics in Manufacturing & Service Industries* 20.4 (July 1, 2010), pp. 287–299. DOI: `10.1002/hfm.20221`.

[168]	John F. Lockett, Ernst Assmann, Rush Green, Matthew P. Reed, Ulrich Raschke, and Jean-Pierre Verriest. "Digital Human Modeling Research and Development User Needs Panel." In: *SAE Transactions* 114 (2005), pp. 886–890. URL: `https://www.jstor.org/stable/44682509` (visited on 08/21/2018).

[169]	Lars Fritzsche, Ricardo Jendrusch, Wolfgang Leidholdt, Sebastian Bauer, Thomas Jäckel, and Attila Pirger. "Introducing Ema (Editor for Manual Work Activities) – A New Tool for Enhancing Accuracy and Efficiency of Human Simulations in Digital Production Planning." In: *Digital Human Modeling*. Ed. by Vincent G. Duffy. Lecture Notes in Computer Science 6777. Springer Berlin Heidelberg, 2011, pp. 272–281.

[170]	Lars Hanson, Dan Högberg, Johan S. Carlson, Robert Bohlin, Erik Brolin, Niclas Delfs, Peter M\a ardberg, Gustafsson Stefan, Ali Keyvani, and Ida-Märta Rhen. "IMMA–Intelligently Moving Manikins in Automotive Applications." In: *Third International Summit on Human Simulation (ISHS2014)*. 2014.

[171]	R. Dörner, W. Broll, P. Grimm, and B. Jung. *Virtual Und Augmented Reality (VR / AR): Grundlagen Und Methoden Der Virtuellen Und Augmentierten Realität*. Springer Berlin Heidelberg, 2014. URL: `http://books.google.de/books?id=aXckBAAAQBAJ`.

[172]	Michael M. Otto, Philipp Agethen, Florian Geiselhart, Michael Rietzler, Felix Gaisbauer, and Enrico Rukzio. "Presenting a Holistic Framework for Scalable, Marker-Less Motion Capturing: Skeletal Tracking Performance Analysis, Sensor Fusion Algorithms and Usage in Automotive Industry." In: *Journal of Virtual Reality and Broadcasting, JVRB* 3.13 (2016), p. 10. DOI: `10.20385/1860-2037/13.2016.3`.

[173]   Michael Otto, Philipp Agethen, Florian Geiselhart, and Enrico Rukzio. "Towards Ubiquitous Tracking: Presenting a Scalable, Markerless Tracking Approach Using Multiple Depth Cameras." In: *Proceedings of the EuroVR 2015*. EuroVR 2015. Best Industrial Paper Award. Lecco, Italy: European Association for Virtual Reality and Augmented Reality, 2015. URL: https://www.eurovr-association.org/wp-content/uploads/2019/11/06112017_Proceedings-1.pdf.

[174]   Florian Geiselhart, Michael Otto, and Enrico Rukzio. "On the Use of Multi-Depth-Camera Based Motion Tracking Systems in Production Planning Environments." In: *Proceedings of the 48th CIRP Conference on Manufacturing Systems*. CIRP CMS 2015. Vol. 41. CIRP CMS. Ischia, Italy: Elsevier, 2015, pp. 759–764. DOI: 10.1016/j.procir.2015.12.088.

[175]   Zhidong Xiao, Hammadi Nait-Charif, and Jian J. Zhang. "Automatic Estimation of Skeletal Motion from Optical Motion Capture Data." In: *Motion in Games*. International Workshop on Motion in Games. Springer, Berlin, Heidelberg, June 14, 2008, pp. 144–153. DOI: 10.1007/978-3-540-89220-5_15.

[176]   Huiyu Zhou and Huosheng Hu. "Human Motion Tracking for Rehabilitation—A Survey." In: *Biomedical Signal Processing and Control* 3.1 (2008), pp. 1–18.

[177]   Gary Bishop, Greg Welch, and B. Danette Allen. "Tracking: Beyond 15 Minutes of Thought." In: *SIGGRAPH Course Pack* 11 (2001).

[178]   Sarvenaz Salehi, Gabriele Bleser, Norbert Schmitz, and Didier Stricker. "A Low-Cost and Light-Weight Motion Tracking Suit." In: *2013 IEEE 10th International Conference on Ubiquitous Intelligence and Computing and 2013 IEEE 10th International Conference on Autonomic and Trusted Computing*. IEEE, 2013, pp. 474–479.

[179]   Francesco Caputo, Alessandro Greco, D'Amato Egidio, Immacolata Notaro, and Stefania Spada. "IMU-Based Motion Capture Wearable System for Ergonomic Assessment in Industrial Environment." In: *International Conference on Applied Human Factors and Ergonomics*. Springer, 2018, pp. 215–225.

[180]   Daniel Roetenberg, Henk Luinge, and Per Slycke. "Xsens MVN: Full 6DOF Human Motion Tracking Using Miniature Inertial Sensors." In: *Xsens Motion Technologies BV, Tech. Rep* 1 (2009).

[181]   Suya You, Ulrich Neumann, and Ronald Azuma. "Hybrid Inertial and Vision Tracking for Augmented Reality Registration." In: *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*. IEEE, 1999, pp. 260–267.

[182]   Advanced Realtime Tracking GmbH. *Motion Capture - Markers & Targets - Products - ART Advanced Realtime Tracking*. 2019. URL: https://ar-tracking.com/products/markers-targets/motion-capture/#c328 (visited on 05/20/2019).

[183]   Advanced Realtime Tracking GmbH. *DTrack2 User Manual v2.9.0: System User Manual ARTtrack, TRACKPACK & DTrack*. 00000. 2013.

[184]   Jamie Shotton, Ross Girshick, Andrew Fitzgibbon, Toby Sharp, Mat Cook, Mark Finocchio, Richard Moore, Pushmeet Kohli, Antonio Criminisi, Alex Kipman, and Andrew Blake. "Efficient Human Pose Estimation from Single Depth Images." In: *Trans. PAMI* (2012). URL: http://research.microsoft.com/apps/pubs/default.aspx?id=174951.

[185]   Jamie Shotton, Toby Sharp, Alex Kipman, Andrew Fitzgibbon, Mark Finocchio, Andrew Blake, Mat Cook, and Richard Moore. "Real-Time Human Pose Recognition in Parts from Single Depth Images." In: *Communications of the ACM* 56.1 (2013), pp. 116–124. URL: http://dl.acm.org/citation.cfm?id=2398381 (visited on 09/13/2016).

[186]   Antonio Sementille, Luís Escaramuzi Lourenço, José Remo Brega, and Ildeberto Rodello. "A Motion Capture System Using Passive Markers." In: Jan. 1, 2004, pp. 440–447. DOI: 10.1145/1044588.1044684.

[187]   Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. "Realtime Multi-Person 2d Pose Estimation Using Part Affinity Fields." In: *arXiv preprint arXiv:1611.08050* (2016).

[188]   N. Hasler, B. Rosenhahn, T. Thormahlen, M. Wand, J. Gall, and H.-P. Seidel. "Markerless Motion Capture with Unsynchronized Moving Cameras." In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference On*. 00143. 2009, pp. 224–231. DOI: 10.1109/CVPR.2009.5206859.

[189]   Nils Hasler, Carsten Stoll, Martin Sunkel, Bodo Rosenhahn, and H.-P. Seidel. "A Statistical Model of Human Pose and Body Shape." In: *Computer Graphics Forum*. Vol. 28. Wiley Online Library, 2009, pp. 337–346.

[190]   Klaus D. Tönnies. *Grundlagen Der Bildverarbeitung*. Informatik : Bildverarbeitung. 00113. München and Boston [u.a.]: Pearson Studium, 2005.

[191]   Marcin Grzegorzek, Christian Theobalt, Reinhard Koch, and Andreas Kolb, eds. *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*. Red. by David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos,

Doug Tygar, Moshe Y. Vardi, and Gerhard Weikum. Vol. 8200. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. URL: http://srr.cir-mcs.e.corpintra.net/book/10.1007/978-3-642-44964-2 (visited on 09/18/2014).

[192]   J. Sell and P. O'Connor. "The Xbox One System on a Chip and Kinect Sensor." In: *Micro, IEEE* 34.2 (Mar. 2014). 00008, pp. 44–53. DOI: 10.1109/MM.2014.9.

[193]   Microsoft. *Kinect for Windows Features*. 00000. 2015. URL: http://www.microsoft.com/en-us/kinectforwindows/meetkinect/features.aspx.

[194]   Zhengyou Zhang. "Microsoft Kinect Sensor and Its Effect." In: *MultiMedia, IEEE* 19.2 (2012). 00873, pp. 4–10. DOI: 10.1109/MMUL.2012.24.

[195]   Chen Qian, Xiao Sun, Yichen Wei, Xiaoou Tang, and Jian Sun. "Realtime and Robust Hand Tracking from Depth." In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014, pp. 1106–1113.

[196]   Christian Schönauer and Hannes Kaufmann. "Wide Area Motion Tracking Using Consumer Hardware." In: *The International Journal of Virtual Reality* 12.1 (2013). 00009, pp. 1–9. URL: http://publik.tuwien.ac.at/files/PubDat%E2%80%97%20219247.pdf.

[197]   M. Martínez-Zarzuela, M. Pedraza-Hueso, F. J. Díaz-Pernas, D. González-Ortega, and M. Antón-Rodríguez. "Indoor 3D Video Monitoring Using Multiple Kinect Depth-Cameras." In: *The International Journal of Multimedia And Its Applications* 1 (Mar. 12, 2014), pp. 61–77. DOI: 10.5121/ijma.2014.6105.

[198]   Andy Wilson. *RoomAlive Toolkit*. GitHub, 2015. URL: https://github.com/Kinect/RoomAliveToolkit (visited on 01/11/2015).

[199]   Michael Rietzler, Florian Geiselhart, Janek Thomas, and Enrico Rukzio. "FusionKit: A Generic Toolkit for Skeleton, Marker and Rigid-Body Tracking." In: *Proceedings of the 8th ACM SIGCHI Symposium on Engineering Interactive Computing Systems*. EICS '16. 00000. New York, NY, USA: ACM, 2016, pp. 73–84. DOI: 10.1145/2933242.2933263.

[200]   Felix Leif Keppmann, Tobias Käfer, Steffen Stadtmüller, René Schubotz, and Andreas Harth. "High Performance Linked Data Processing for Virtual Reality Environments." In: *Proceedings of the ISWC 2014 Posters & Demonstrations Track a Track within the 13th International Semantic Web Conference, ISWC 2014, Riva Del Garda, Italy, October 21, 2014*. 00005. 2014, pp. 193–196. URL: http://ceur-ws.org/Vol-1272/paper%E2%80%97%20149.pdf.

[201]   D. Alex Butler, Shahram Izadi, Otmar Hilliges, David Molyneaux, Steve Hodges, and David Kim. "Shake'n'sense: Reducing Interference for Overlapping Structured Light Depth Cameras." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '12. New York, NY, USA: ACM, 2012, pp. 1933–1936. DOI: 10.1145/2207676.2208335.

[202]   F. Faion, P. Ruoff, A. Zea, and U.D. Hanebeck. "Recursive Bayesian Calibration of Depth Sensors with Non-Overlapping Views." In: *2012 15th International Conference on Information Fusion (FUSION)*. 2012 15th International Conference on Information Fusion (FUSION). 2012, pp. 757–762.

[203]   A. Maimone and H. Fuchs. "Encumbrance-Free Telepresence System with Real-Time 3D Capture and Display Using Commodity Depth Cameras." In: *2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR). 00102. 2011, pp. 137–146. DOI: 10.1109/ISMAR.2011.6092379.

[204]   Andrew Wilson and Hrvoje Benko. "Combining Multiple Depth Cameras and Projectors for Interactions on, above and between Surfaces." In: *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology*. UIST '10. 00303. New York, NY, USA: ACM, 2010, pp. 273–282. DOI: 10.1145/1866029.1866073.

[205]   Kai Berger, Kai Ruhl, Christian Brümmer, Yannic Schröder, Alexander Scholz, and Marcus Magnor. "Markerless Motion Capture Using Multiple Color-Depth Sensors." In: *Proc. Vision, Modeling and Visualization (VMV) 2011*. 00093. Oct. 2011, pp. 317–324.

[206]   Licong Zhang, Jürgen Sturm, Daniel Cremers, and Dongheui Lee. "Real-Time Human Motion Tracking Using Multiple Depth Cameras." In: *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference On*. IEEE, 2012, pp. 2389–2395. URL: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6385968 (visited on 03/27/2015).

[207]   Szymon Rusinkiewicz and Marc Levoy. "Efficient Variants of the ICP Algorithm." In: *Proceedings of the Third Intl. Conf. on 3D Digital Imaging and Modeling*. 02356. 2001, pp. 145–152. URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.23.9388 (visited on 12/10/2013).

[208]   F. Pomerleau, S. Magnenat, F. Colas, Ming Liu, and R. Siegwart. "Tracking a Depth Camera: Parameter Exploration for Fast ICP." In: *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 00058. 2011, pp. 3824–3829. DOI: 10.1109/IROS.2011.6094861.

[209]   Stylianos Asteriadis, Anargyros Chatzitofis, Dimitrios Zarpalas, Dimitrios S. Alexiadis, and Petros Daras. "Estimating Human Motion from Multiple Kinect Sensors." In: *Proceedings of the 6th International Conference on Computer Vision / Computer Graphics Collaboration Techniques and Applications*. MIRAGE '13. 00016. New York, NY, USA: ACM, 2013, 3:1–3:6. DOI: 10.1145/2466715.2466727.

[210]   Seongmin Baek and Myunggyu Kim. "Dance Experience System Using Multiple Kinects." In: *International Journal of Future Computer and Communication* 4.1 (2015). 00003, pp. 45–49. DOI: 10.7763/IJFCC.2015.V4.353.

[211]   S. Kaenchan, P. Mongkolnam, B. Watanapa, and S. Sathienpong. "Automatic Multiple Kinect Cameras Setting for Simple Walking Posture Analysis." In: *2013 International Computer Science and Engineering Conference (ICSEC)*. 2013 International Computer Science and Engineering Conference (ICSEC). Sept. 2013, pp. 245–249. DOI: 10.1109/ICSEC.2013.6694787.

[212]   Maurizio Caon, Yong Yue, Julien Tscherrig, Elena Mugellini, and Omar Abou Khaled. "Context-Aware 3D Gesture Interaction Based on Multiple Kinects." In: AMBIENT 2011, The First International Conference on Ambient Computing, Applications, Services and Technologies. 00026. Oct. 23, 2011, pp. 7–12. URL: http://www.thinkmind.org/index.php?view=article&articleid=ambient_2011_1_20_60044 (visited on 10/30/2013).

[213]   N. A. Azis, H. Choi, and Y. Iraqi. "Substitutive Skeleton Fusion for Human Action Recognition." In: *2015 International Conference on Big Data and Smart Computing (BIGCOMP)*. 2015 International Conference on Big Data and Smart Computing (BIGCOMP). Feb. 2015, pp. 170–177. DOI: 10.1109/35021BIGCOMP.2015.7072828.

[214]   A. Kitsikidis, K. Dimitropoulos, S. Douka, and N. Grammalidis. "Dance Analysis Using Multiple Kinect Sensors." In: *2014 International Conference on Computer Vision Theory and Applications (VISAPP)*. 2014 International Conference on Computer Vision Theory and Applications (VISAPP). Vol. 2. Jan. 2014, pp. 789–795.

[215]   Kwok-Yun Yeung, Tsz-Ho Kwok, and Charlie C. L. Wang. "Improved Skeleton Tracking by Duplex Kinects: A Practical Approach for Real-Time Applications." In: *Journal of Computing and Information Science in Engineering* 13.4 (Oct. 16, 2013). 00012, pp. 041007–041007. DOI: 10.1115/1.4025404.

[216]   Jasper Brekelmans. *Multi Sensor Alpha*. URL: https://brekel.com/multi-sensor/ (visited on 05/21/2019).

[217]   Q. Wang, G. Kurillo, F. Ofli, and R. Bajcsy. "Evaluation of Pose Tracking Accuracy in the First and Second Generations of Microsoft Kinect." In: *2015 International Conference on Healthcare Informatics*. 2015 International Conference on Healthcare Informatics. Oct. 2015, pp. 380–389. DOI: 10.1109/ICHI.2015.54.

[218]   L. Yang, L. Zhang, H. Dong, A. Alelaiwi, and A. E. Saddik. "Evaluating and Improving the Depth Accuracy of Kinect for Windows V2." In: *IEEE Sensors Journal* 15.8 (Aug. 2015), pp. 4275–4285. DOI: 10.1109/JSEN.2015.2416651.

[219]   S. Obdrzalek, G. Kurillo, F. Ofli, R. Bajcsy, E. Seto, H. Jimison, and M. Pavel. "Accuracy and Robustness of Kinect Pose Estimation in the Context of Coaching of Elderly Population." In: *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 00121. 2012, pp. 1188–1193. DOI: 10.1109/EMBC.2012.6346149.

[220]   Karen Otte, Bastian Kayser, Sebastian Mansow-Model, Julius Verrel, Friedemann Paul, Alexander U. Brandt, and Tanja Schmitz-Hübsch. "Accuracy and Reliability of the Kinect Version 2 for Clinical Measurement of Motor Function." In: *PloS one* 11.11 (2016), e0166532.

[221]   Rudolph Emil Kalman. "A New Approach to Linear Filtering and Prediction Problems." In: *Journal of basic Engineering* 82.1 (1960), pp. 35–45.

[222]   Ari Shapiro, Andrew Feng, Ruizhe Wang, Hao Li, Mark Bolas, Gerard Medioni, and Evan Suma. "Rapid Avatar Capture and Simulation Using Commodity Depth Sensors." In: *Computer Animation and Virtual Worlds* 25.3-4 (2014), pp. 201–211. DOI: 10.1002/cav.1579.

[223]   Hao Li, Etienne Vouga, Anton Gudym, Linjie Luo, Jonathan T. Barron, and Gleb Gusev. "3D Self-Portraits." In: *ACM Trans. Graph.* 32.6 (Nov. 2013), 187:1–187:9. DOI: 10.1145/2508363.2508407.

[224]   T. Zhao, S. Li, K. N. Ngan, and F. Wu. "3-D Reconstruction of Human Body Shape From a Single Commodity Depth Camera." In: *IEEE Transactions on Multimedia* 21.1 (Jan. 2019), pp. 114–123. DOI: 10.1109/TMM.2018.2844087.

[225]   Konstantina Kilteni, Raphaela Groten, and Mel Slater. "The Sense of Embodiment in Virtual Reality." In: *Presence: Teleoperators and Virtual Environments* 21.4 (2012), pp. 373–387.

[226] H. Haggag, M. Hossny, S. Nahavandi, and D. Creighton. "Real Time Ergonomic Assessment for Assembly Operations Using Kinect." In: *2013 UKSim 15th International Conference on Computer Modelling and Simulation (UKSim)*. 2013 UKSim 15th International Conference on Computer Modelling and Simulation (UKSim). 00013. Apr. 2013, pp. 495–500. DOI: 10.1109/UKSim.2013.105.

[227] Marco Bortolini, Mauro Gamberi, Francesco Pilati, and Alberto Regattieri. "Automatic Assessment of the Ergonomic Risk for Manual Manufacturing and Assembly Activities through Optical Motion Capture Technology." In: *Procedia CIRP*. 51st CIRP Conference on Manufacturing Systems 72 (Jan. 1, 2018), pp. 81–86. DOI: 10.1016/j.procir.2018.03.198.

[228] Lynn McAtamney and E. Nigel Corlett. "RULA: A Survey Method for the Investigation of Work-Related Upper Limb Disorders." In: *Applied ergonomics* 24.2 (1993), pp. 91–99.

[229] Sue Hignett and Lynn McAtamney. "Rapid Entire Body Assessment (REBA)." In: *Applied ergonomics* 31.2 (2000), pp. 201–205.

[230] National Institute for Occupational Safety and Health Division of Biomedical and Behavioral Science NIOSH. *Work Practices Guide for Manual Lifting*. Cincinnati, OH: U.S. Dept. of Health and Human Services, Public Health Service, Centers for Disease Control, National Institute for Occupational Safety and Health, Division of Biomedical and Behavioral Science, 1981. 183 pp.

[231] E. OCCHIPINTI. "OCRA: A Concise Index for the Assessment of Exposure to Repetitive Movements of the Upper Limbs." In: *Ergonomics* 41.9 (Sept. 1, 1998), pp. 1290–1311. DOI: 10.1080/001401398186315. pmid: 9754032.

[232] K. Schaub and G. Winter. "Design Check. A screening method to assess physical stress." In: *Der Orthopade* 31.10 (Oct. 2002), pp. 987–996. DOI: 10.1007/s00132-002-0353-6. pmid: 12376873.

[233] Peter van der Meulen and Andreas Seidl. "Ramsis – The Leading Cad Tool for Ergonomic Analysis of Vehicles." In: *Digital Human Modeling*. Ed. by Vincent G. Duffy. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2007, pp. 1008–1017.

[234] Karlheinz G. Schaub, Jens Mühlstedt, Benjamin Illmann, Sebastian Bauer, Lars Fritzsche, Torsten Wagner, Angelika C. Bullinger-Hoffmann, and Ralph Bruder. "Ergonomic Assessment of Automotive Assembly Tasks with Digital Human Modelling and the 'Ergonomics Assessment Worksheet' (EAWS)." In: *International Journal of Human Factors Modelling and Simu-*

lation 3.3-4 (Jan. 1, 2012), pp. 398–426. DOI: `10.1504/IJHFMS.2012.051581`.

[235] Pierre Plantard, Antoine Muller, Charles Pontonnier, Georges Dumont, Hubert P. H. Shum, and Franck Multon. "Inverse Dynamics Based on Occlusion-Resistant Kinect Data: Is It Usable for Ergonomics?" In: *International Journal of Industrial Ergonomics* 61 (Sept. 2017). 00000, pp. 71–80. DOI: `10.1016/j.ergon.2017.05.010`.

[236] Peter Peltonen, Esko Kurvinen, Antti Salovaara, Giulio Jacucci, Tommi Ilmonen, John Evans, Antti Oulasvirta, and Petri Saarikko. "It's Mine, Don't Touch!: Interactions at a Large Multi-Touch Display in a City Centre." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2008, pp. 1285–1294.

[237] Mary Czerwinski, Greg Smith, Tim Regan, Brian Meyers, George G. Robertson, and Gary K. Starkweather. "Toward Characterizing the Productivity Benefits of Very Large Displays." In: *Interact*. Vol. 3. 2003, pp. 9–16.

[238] Christopher Andrews, Alex Endert, Beth Yost, and Chris North. "Information Visualization on Large, High-Resolution Displays: Issues, Challenges, and Opportunities." In: *Information Visualization* 10.4 (2011), pp. 341–355.

[239] Philipp Agethen, Michael Otto, Stefan Mengel, and Enrico Rukzio. "Using Marker-Less Motion Capture Systems for Walk Path Analysis in Paced Assembly Flow Lines." In: *Proceedings of the 6th CIRP Conference on Learning Factories*. CIRP CLF 2016. Vol. 54. 2016, pp. 152–157. DOI: `10.1016/j.procir.2016.04.125`.

[240] Philipp Agethen, Michael Otto, Felix Gaisbauer, and Enrico Rukzio. "Presenting a Novel Motion Capture-Based Approach for Walk Path Segmentation and Drift Analysis in Manual Assembly." In: *Proceedings of the Sixth International Conference on Changeable, Agile, Reconfigurable and Virtual Production*. CIRP CARV 2016. Vol. 52. CIRP CARV 2016. Jan. 1, 2016, pp. 286–291. DOI: `10.1016/j.procir.2016.07.048`.

[241] Michael Otto, Eva Lampen, Philipp Agethen, Mareike Langohr, Gerald Masan, and Enrico Rukzio. "Evaluation on Perceived Sizes Using Large-Scale Augmented Floor Visualization Devices." In: *Proceedings of the 8th ACM International Symposium on Pervasive Displays*. ACM PerDis 2019. Palermo, Italy: Association for Computing Machinery, June 12, 2019, pp. 1–7. DOI: `10.1145/3321335.3324951`.

[242] Michael Otto, Eva Lampen, Philipp Agethen, Gabriel Zachmann, and Enrico Rukzio. "Using Large-Scale Augmented Floor Surfaces for Industrial Applications and Evaluation on Perceived Sizes." In: *Personal and ubiquitous computing, PAUC*. Theme Issue on Pervasive Displays tbd.2020 (2020), p. 28. DOI: 10.1007/s00779-020-01380-9.

[243] Mark Weiser. "The Computer for the 21st Century." In: *Scientific american* 265.3 (1991), pp. 94–105.

[244] Lucia Terrenghi, Aaron Quigley, and Alan Dix. "A Taxonomy for and Analysis of Multi-Person-Display Ecosystems." In: *Personal and Ubiquitous Computing* 13.8 (Nov. 1, 2009), p. 583. DOI: 10.1007/s00779-009-0244-5.

[245] Lars Lischke, Jürgen Grüninger, Khalil Klouche, Albrecht Schmidt, Philipp Slusallek, and Giulio Jacucci. "Interaction Techniques for Wall-Sized Screens." In: *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*. ITS '15. New York, NY, USA: ACM, 2015, pp. 501–504. DOI: 10.1145/2817721.2835071.

[246] Fernando Garcia-Sanjuan, Javier Jaen, and Vicente Nacher. "Toward a General Conceptualization of Multi-Display Environments." In: *Frontiers* 3.20 (2016), p. 1.

[247] Yvonne Rogers and Siân Lindley. "Collaborating around Vertical and Horizontal Large Interactive Displays: Which Way Is Best?" In: *Interacting with Computers* 16.6 (2004), pp. 1133–1152.

[248] Xiaojun Bi and Ravin Balakrishnan. "Comparing Usage of a Large High-Resolution Display to Single or Dual Desktop Displays for Daily Work." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009, pp. 1005–1014.

[249] Lars Lischke, Sven Mayer, Katrin Wolf, Niels Henze, Albrecht Schmidt, Svenja Leifert, and Harald Reiterer. "Using Space: Effect of Display Size on Users' Search Performance." In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '15. New York, NY, USA: ACM, 2015, pp. 1845–1850. DOI: 10.1145/2702613.2732845.

[250] Lars Lischke, Sven Mayer, Jan Hoffmann, Philipp Kratzer, Stephan Roth, Katrin Wolf, and Pawe\l Woniak. "Interaction Techniques for Window Management on Large High-Resolution Displays." In: *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*. MUM '17. New York, NY, USA: ACM, 2017, pp. 241–247. DOI: 10.1145/3152832.3152852.

[251] Christopher Andrews, Alex Endert, and Chris North. "Space to Think: Large High-Resolution Displays for Sensemaking." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2010, pp. 55–64.

[252] Daniel Vogel and Ravin Balakrishnan. "Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays." In: *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*. ACM, 2005, pp. 33–42.

[253] Shahzad Malik, Abhishek Ranjan, and Ravin Balakrishnan. "Interacting with Large Displays from a Distance with Vision-Tracked Multi-Finger Gestural Input." In: *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*. ACM, 2005, pp. 43–52.

[254] Nobuchika Sakata, Fumihiro Sato, Tomu Tominaga, and Yoshinori Hijikata. "Extension of Smartphone by Wearable Input/Output Interface with Floor Projection." In: *International Conference on Collaboration Technologies*. Springer, 2017, pp. 168–181.

[255] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. De-Fanti. "Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE." In: *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*. ACM, 1993, pp. 135–142.

[256] Claudio Pinhanez. "The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces." In: *Ubicomp 2001: Ubiquitous Computing*. Springer, 2001, pp. 315–331. URL: http://link.springer.com/chapter/10.1007/3-540-45427-6_27 (visited on 06/12/2014).

[257] Thomas Augsten, Konstantin Kaefer, René Meusel, Caroline Fetzer, Dorian Kanitz, Thomas Stoff, Torsten Becker, Christian Holz, and Patrick Baudisch. "Multitoe: High-Precision Interaction with Back-Projected Floors Based on High-Resolution Multi-Touch Input." In: *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology*. ACM, 2010, pp. 209–218. URL: http://dl.acm.org/citation.cfm?id=1866064 (visited on 08/29/2014).

[258] Jörg Müller, Dieter Eberle, and Constantin Schmidt. "Base-Lase: An Interactive Focus+Context Laser Floor." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. New York, NY, USA: ACM, 2015, pp. 3869–3878. DOI: 10.1145/2702123.2702246.

[259] Nicholas Sheep Dalton. "TapTiles: LED-Based Floor Interaction." In: *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*. ACM, 2013, pp. 165–174.

URL: http://dl.acm.org/citation.cfm?id=2512800 (visited on 06/12/2014).

[260] Tobi Delbrück, Adrian M. Whatley, Rodney Douglas, Kynan Eng, Klaus Hepp, and Paul FMJ Verschure. "A Tactile Luminous Floor for an Interactive Autonomous Space." In: *Robotics and autonomous systems* 55.6 (2007), pp. 433–443. URL: http://www.sciencedirect.com/science/article/pii/S0921889007000231 (visited on 06/12/2014).

[261] Jorge D. Camba, Jeff Feng, and Eunsook Kwon. "Interactive Experiences in Public Spaces: A Novel Floor Display System Based on Luminous Modular Units." In: *Hawaii International Conference on System Sciences 2018*. Hawaii International Conference on System Sciences 2018. Jan. 3, 2018. DOI: 10.24251/HICSS.2018.040.

[262] Jo Vermeulen, Kris Luyten, Karin Coninx, Nicolai Marquardt, and Jon Bird. "Proxemic Flow: Dynamic Peripheral Floor Visualizations for Revealing and Mediating Large Surface Interactions." In: *Human-Computer Interaction*. Springer, 2015, pp. 264–281.

[263] Kaj Grønbaek, Ole S. Iversen, Karen Johanne Kortbek, Kaspar Rosengreen Nielsen, and Louise Aagaard. "IGameFloor: A Platform for Co-Located Collaborative Games." In: *Proceedings of the International Conference on Advances in Computer Entertainment Technology*. ACE '07. 00000. New York, NY, USA: ACM, 2007, pp. 64–71. DOI: 10.1145/1255047.1255061.

[264] Juliano M. Franz, Joseph W. Malloch, Derek Reilly, and Vinicius Sanches. "SpaceHopper: Bounce Your Way to Galactic Domination." In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI EA '18. New York, NY, USA: ACM, 2018, D211:1–D211:4. DOI: 10.1145/3170427.3186509.

[265] Dominik Schmidt, Raf Ramakers, Esben W. Pedersen, Johannes Jasper, Sven Köhler, Aileen Pohl, Hannes Rantzsch, Andreas Rau, Patrick Schmidt, and Christoph Sterz. "Kickables: Tangibles for Feet." In: *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2014, pp. 3143–3152.

[266] Ben Heller, Terry Senior, and Jon Wheat. "The Smartfloor: A Large Area Force-Measuring Floor for Investigating Dynamic Balance and Motivating Exercise." In: *Procedia engineering* 72 (2014), pp. 226–231.

[267]   Marianne Graves Petersen, Peter Gall Krogh, Martin Ludvigsen, and Andreas Lykke-Olesen. "Floor Interaction HCI Reaching New Ground." In: *CHI'05 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2005, pp. 1717–1720.

[268]   Alvin W. Law, Jessica W. Ip, Benjamin V. Peck, Yon Visell, Paul G. Kry, and Jeremy R. Cooperstock. "Multimodal Floor for Immersive Environments." In: *ACM SIGGRAPH 2009 Emerging Technologies*. ACM, 2009, p. 16.

[269]   Dominik Schmidt, Johannes Frohnhofen, Sven Knebel, Florian Meinel, Mariya Perchyk, Julian Risch, Jonathan Striebel, Julia Wachtel, and Patrick Baudisch. "Ergonomic Interaction for Touch Floors." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. New York, NY, USA: ACM, 2015, pp. 3879–3888. DOI: 10.1145/2702123.2702254.

[270]   Leyard Inc. *Leyard TWA Series*. URL: /en/products/led-video-walls/twa/ (visited on 09/24/2018).

[271]   Anastasia Bezerianos and Petra Isenberg. *Perceptual Affordances of Wall-Sized Displays for Visualization Applications*. European Union, 2013.

[272]   William Epstein. "Attitudes of Judgment and the Size-Distance Invariance Hypothesis." In: *Journal of experimental Psychology* 66.1 (1963), p. 78.

[273]   William Epstein and Krishan D. Broota. "Attitude of Judgment and Reaction Time in Estimation of Size at a Distance." In: *Perception & Psychophysics* 18.3 (1975), pp. 201–204.

[274]   Mark Wagner. "The Metric of Visual Space." In: *Perception & psychophysics* 38.6 (1985), pp. 483–495.

[275]   William S. Cleveland and Robert McGill. "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods." In: *Journal of the American statistical association* 79.387 (1984), pp. 531–554.

[276]   Justin Talbot, Vidya Setlur, and Anushka Anand. "Four Experiments on the Perception of Bar Charts." In: *IEEE transactions on visualization and computer graphics* 20.12 (2014), pp. 2152–2160.

[277]   Jack M. Loomis and Joshua M. Knapp. "Visual Perception of Egocentric Distance in Real and Virtual Environments." In: *Virtual and adaptive environments* 11 (2003), pp. 21–46.

[278]   Rebekka S. Renner, Boris M. Velichkovsky, and Jens R. Helmert. "The Perception of Egocentric Distances in Virtual Environments - A Review." In: *ACM Comput. Surv.* 46.2 (Dec. 2013), 23:1–23:40. DOI: 10.1145/2543581.2543590.

[279] Jonathan W. Kelly, Lucia A. Cherep, and Zachary D. Siegel. "Perceived Space in the HTC Vive." In: *ACM Trans. Appl. Percept.* 15.1 (July 2017), 2:1–2:16. DOI: 10.1145/3106155.

[280] J. Scott Armstrong and Fred Collopy. "Error Measures for Generalizing about Forecasting Methods: Empirical Comparisons." In: *International journal of forecasting* 8.1 (1992), pp. 69–80.

[281] J. Martin Bland and Douglas G. Altman. "Measuring Agreement in Method Comparison Studies." In: *Statistical methods in medical research* 8.2 (1999), pp. 135–160.

[282] F. Doil, W. Schreiber, T. Alt, and C. Patron. "Augmented Reality for Manufacturing Planning." In: *Proceedings of the Workshop on Virtual Environments 2003*. EGVE '03. 00108. New York, NY, USA: ACM, 2003, pp. 71–76. DOI: 10.1145/769953.769962.

[283] K. Pentenrieder, C. Bade, F. Doil, and P. Meier. "Augmented Reality-Based Factory Planning - an Application Tailored to Industrial Needs." In: *6th IEEE and ACM International Symposium on Mixed and Augmented Reality, 2007. ISMAR 2007*. 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, 2007. ISMAR 2007. 00059. Nov. 2007, pp. 31–42. DOI: 10.1109/ISMAR.2007.4538822.

[284] Michele Fiorentino, Rafael Radkowski, Antonio Boccaccio, and Antonio Emmanuele Uva. "Magic Mirror Interface for Augmented Reality Maintenance: An Automotive Case Study." In: *Proceedings of the International Working Conference on Advanced Visual Interfaces*. AVI '16. 00000. New York, NY, USA: ACM, 2016, pp. 160–167. DOI: 10.1145/2909132.2909268.

[285] Pierre Fite-Georgel. "Is There a Reality in Industrial Augmented Reality?" In: *2011 10th Ieee International Symposium on Mixed and Augmented Reality*. IEEE, 2011, pp. 201–210.

[286] Oliver Bimber and Ramesh Raskar. *Spatial Augmented Reality: Merging Real and Virtual Worlds*. 00838. CRC press, 2005. URL: https://books.google.de/books?hl=de&lr=&id=JEu3BgAAQBAJ&oi=fnd&pg=PP1&dq=spatial+augmented+reality&ots=NegseiH3V2&sig=ZwxinJkpCDDTwhLBhzjQNWMS6s0 (visited on 06/06/2017).

[287] Antonio E. Uva, Michele Gattullo, Vito M. Manghisi, Daniele Spagnulo, Giuseppe L. Cascella, and Michele Fiorentino. "Evaluating the Effectiveness of Spatial Augmented Reality in Smart Manufacturing: A Solution for Manual Working Stations." In: *The International Journal of Advanced Manufacturing Technology* 94.1-4 (Jan. 1, 2018), pp. 509–521. DOI: 10.1007/s00170-017-0846-4.

[288] theAreaFB. *What Is Projection or Spatial Augmented Reality?* Nov. 18, 2017. URL: http : / / thearea . org / projection - augmented - reality/ (visited on 01/01/2019).

[289] U. Neumann and A. Majoros. "Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance." In: *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180)*. Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180). Mar. 1998, pp. 4–11. DOI: 10 . 1109/VRAIS.1998.658416.

[290] L. De Marchi, A. Ceruti, N. Testoni, A. Marzani, and A. Liverani. "Use of Augmented Reality in Aircraft Maintenance Operations." In: 9064 (Mar. 1, 2014), p. 906412. DOI: 10.1117/12. 2044889.

[291] Ernst Kruijff, J. Edward Swan, and Steven Feiner. "Perceptual Issues in Augmented Reality Revisited." In: *2010 IEEE International Symposium on Mixed and Augmented Reality*. 2010 IEEE International Symposium on Mixed and Augmented Reality. Oct. 2010, pp. 3–12. DOI: 10.1109/ISMAR.2010.5643530.

[292] D. W. F. Van Krevelen and Ronald Poelman. "A Survey of Augmented Reality Technologies, Applications and Limitations." In: *International journal of virtual reality* 9.2 (2010), p. 1.

[293] X. Wang, S. K. Ong, and A. Y. C. Nee. "A Comprehensive Survey of Augmented Reality Assembly Research." In: *Advances in Manufacturing* 4.1 (Mar. 1, 2016), pp. 1–22. DOI: 10 . 1007/ s40436-015-0131-4.

[294] Paula Fraga-Lamas, Tiago M. Fernández-Caramés, Óscar Blanco-Novoa, and Miguel A. Vilar-Montesinos. "A Review on Industrial Augmented Reality Systems for the Industry 4.0 Shipyard." In: *IEEE Access* 6 (2018), pp. 13358–13375.

[295] A. Y. C. Nee, S. K. Ong, G. Chryssolouris, and D. Mourtzis. "Augmented Reality Applications in Design and Manufacturing." In: *CIRP Annals - Manufacturing Technology* 61.2 (2012). 00106, pp. 657–679. DOI: 10.1016/j.cirp.2012.05.010.

[296] Björn Bliese. "Ein gebrauchstaugliches Augmented Reality-System für geometrische Analysen in der Produktentstehung." In: (2019). DOI: http://dx.doi.org/10.18419/opus-10271.

[297] Pierre Wellner. "DigitalDesk Calculator: Tangible Manipulation on a Desk Top Display." In: Proceedings of the 4th Annual ACM Symposium on User Interface Software and Technology. ACM, 1991. URL: https : / / dl . acm . org / citation . cfm?id=120785 (visited on 01/01/2019).

[298]   Pierre David Wellner. *Interacting with Paper on the DigitalDesk*. UCAM-CL-TR-330. University of Cambridge, Computer Laboratory, 1994. URL: https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-330.html (visited on 01/01/2019).

[299]   Volkswagen AG. *Volkswagen Group Use of Spatial Augmented Reality in the Automotive Industry*. 00000. Nov. 14, 2016. URL: http://www.volkswagenag.com/content/vwcorp/content/en/innovation/Virtual_technologies/Use_of_spatial_augmented_reality_in_the_automotive_industry.html (visited on 11/14/2016).

[300]   Jiayi Jin. "Spatial Augmented Reality (SAR) for Museum Learning Merging Augmented Reality Based Features in Educational Space." In: *Museum Management and Curatorship* (Aug. 1, 2017).

[301]   O. Bimber, F. Coriand, A. Kleppe, E. Bruns, S. Zollmann, and T. Langlotz. "Superimposing Pictorial Artwork with Projected Imagery." In: *IEEE MultiMedia* 12.1 (Jan. 2005), pp. 16–26. DOI: 10.1109/MMUL.2005.9.

[302]   Björn Schwerdtfeger, Daniel Pustka, Andreas Hofhauser, and Gudrun Klinker. "Using Laser Projectors for Augmented Reality." In: *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology*. VRST '08. New York, NY, USA: ACM, 2008, pp. 134–137. DOI: 10.1145/1450579.1450608.

[303]   Guillaume Cortes, Eric Marchand, Guillaume Brincin, and Anatole Lécuyer. "MoSART: Mobile Spatial Augmented Reality for 3D Interaction With Tangible Objects." In: *Frontiers in Robotics and AI* 5.93 (2018).

[304]   Christian Winkler, Julian Seifert, David Dobbelstein, and Enrico Rukzio. "Pervasive Information through Constant Personal Projection: The Ambient Mobile Pervasive Display (AMP-D)." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2014, pp. 4117–4126.

[305]   Ben Piper and Hiroshi Ishii. "CADcast: A Method for Projecting Spatially Referenced Procedural Instructions." In: *Technical Report, MIT Media Lab* (2001).

[306]   Antonio Emmanuele Uva, Michele Fiorentino, Michele Gattullo, Marco Colaprico, Maria F. de Ruvo, Francescomaria Marino, Gianpaolo F. Trotta, Vito M. Manghisi, Antonio Boccaccio, Vitoantonio Bevilacqua, and Giuseppe Monno. "Design of a Projective AR Workbench for Manual Working Stations." In: *Augmented Reality, Virtual Reality, and Computer Graphics*. Ed. by Lucio Tommaso De Paolis and Antonio Mongelli. Vol. 9768. 00000. Cham: Springer International Publishing, 2016, pp. 358–367. DOI: 10.1007/978-3-319-40621-3_25.

[307]   Oliver Sand, Sebastian Büttner, Volker Paelke, and Carsten Röcker. "smARt.Assembly – Projection-Based Augmented Reality for Supporting Assembly Workers." In: International Conference on Virtual, Augmented and Mixed Reality. 00000. Springer International Publishing, July 17, 2016, pp. 643–652. DOI: `10.1007/978-3-319-39907-2_61`.

[308]   Ashish Doshi, Ross T. Smith, Bruce H. Thomas, and Con Bouras. "Use of Projector Based Augmented Reality to Improve Manual Spot-Welding Precision and Accuracy for Automotive Manufacturing." In: *The International Journal of Advanced Manufacturing Technology* 89.5-8 (Mar. 1, 2017), pp. 1279–1293. DOI: `10.1007/s00170-016-9164-5`.

[309]   Markus Funk, Thomas Kosch, and Albrecht Schmidt. "Interactive Worker Assistance: Comparing the Effects of in-Situ Projection, Head-Mounted Displays, Tablet, and Paper Instructions." In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016, pp. 934–939.

[310]   SeungJun Kim and Anind K. Dey. "Simulated Augmented Reality Windshield Display as a Cognitive Mapping Aid for Elder Driver Navigation." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 00113. ACM, 2009, pp. 133–142. URL: `http://dl.acm.org/citation.cfm?id=1518724` (visited on 06/05/2017).

[311]   Harald Wuest. "Efficient Line and Patch Feature Characterization and Management for Real-Time Camera Tracking." 00007. phd. Darmstadt: Technische Universität, Nov. 21, 2008. URL: `http://tuprints.ulb.tu-darmstadt.de/1190/` (visited on 10/17/2016).

[312]   Harald Wuest and Didier Stricker. "Tracking of Industrial Objects by Using CAD Models." In: *JVRB - Journal of Virtual Reality and Broadcasting* 4(2007).1 (2007). 00037. URL: `http://nbn-resolving.de/urn:nbn:de:0009-6-11595`.

[313]   mattzmsft. *HoloLens Hardware Details - Mixed Reality*. URL: `https://docs.microsoft.com/en-us/windows/mixed-reality/hololens-hardware-details` (visited on 01/04/2019).

[314]   Michele Di Donato, Michele Fiorentino, Antonio E. Uva, Michele Gattullo, and Giuseppe Monno. "Text Legibility for Projected Augmented Reality on Industrial Workbenches." In: *Computers in Industry* 70 (June 2015). 00004, pp. 70–78. DOI: `10.1016/j.compind.2015.02.008`.

[315]   Rafael Radkowski and Sravya Kanunganti. "Augmented Reality System Calibration for Assembly Support with the Microsoft Hololens." In: *ASME 2018 13th International Manufac-

turing Science and Engineering Conference. American Society of
Mechanical Engineers Digital Collection, 2018.