# A Comparative Evaluation of Three Skin Color Detection Approaches

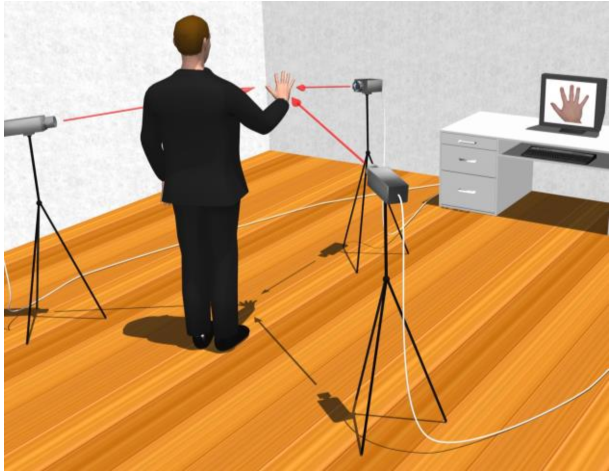
Dennis Jensch, Daniel Mohr, *Clausthal University*

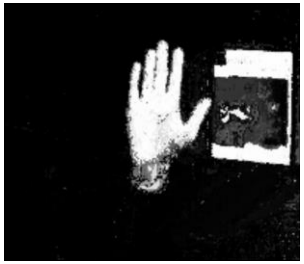
Gabriel Zachmann, *University of Bremen*

*VRAR 2012, Sep 2012, Düsseldorf*

# Motivation

- Long-term goal: marker-less hand tracking
  - Real-time estimation of pose and human hand
  - Skin segmentation helps to
    - localize hand very fast (**if robust**)
    - match hand temlates very fast





- Further applications
  - Detect person/adult images (e.g. filtering in search engines)
  - Face detection
  - Many more…

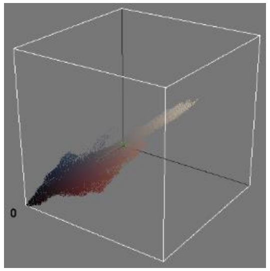
# Challenges of Skin Segmentation

- Different ethnic groups

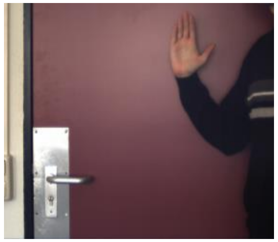


- Camera limitations






- Skin color in the background



- Illumination conditions


# Approaches Considered in the Following

- *RehgJones* [M. J. Jones and J. M. Rehg, IJCV 1999]
  - Learn skin color distribution from a manually labeled dataset
- *HybridClustering* [D. Mohr and G. Zachmann, CAIP 2007]
  - Combined color and image space clustering
  - Classification is done region-wise (opposed to pixel-wise)
- *NeuralGasColorClustering*
  - Inspired by HybridClustering with two modifications
  - Replace EM by Matrix Neural Gas
  - Replace the way the number of clusters is determined

# RehgJones

- Learn skin color distribution offline
  - Dataset randomly chosen from World Wide Web
    - ~ 1 billion pixels
  - Manually labeled as skin / non-skin
  - Color distributions for skin and non-skin

$$P(rgb|skin) = \frac{s[rgb]}{T_S} \qquad P(rgb|\neg skin) = \frac{s[\neg rgb]}{T_N}$$

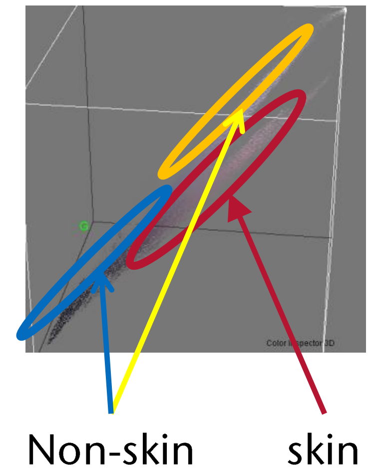
- Image classification:
  - Per pixel

$$\frac{P(rgb|skin)}{P(rgb|\neg skin)} \geq \Theta$$

  - $\Theta$ controls offset between *false positive* and *false negatives*

# HybridClustering

- Learn a rough skin direction vector offline
- Online classification:
  - Cluster the image in color space
    - Hierarchical EM
    - Smoothing of clusters in image space




- Classify image clusters as skin / non-skin
- Reproject to image space

\+ Keep image regions together

\- Depends on convergence behavior of EM

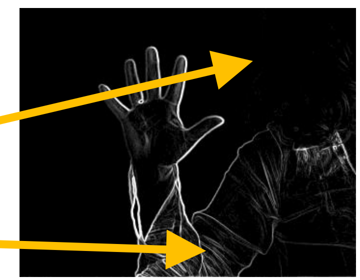
# NeuralGasColorClustering

- Tries to improve upon **HybridClustering**
- EM algorithm
  - sensitive to initialization
- Hierarchical clustering to determine number of clusters
  - Could choose wrong number of clusters
- Image edges as cluster quality measure
  - Is this really the best option?

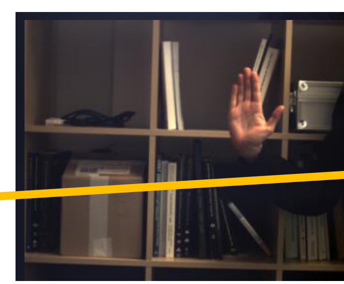
- **NeuralGasColorClustering**
- Matrix Neural Gas
  - Less sensitive to initialization
- Succesively test different number of clusters
  - Slower but expected to perform better
- Test 3 different measures
  - Border Length
  - Border Edges
  - Color Space Compactness

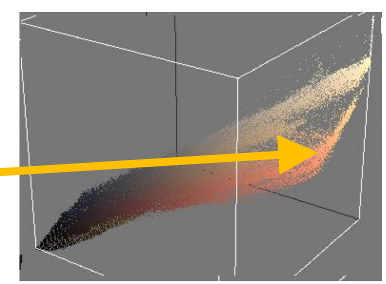
# Quality Measures for Cluster  in *NeuralGasColorClusters*

- Border Length
  - \+ Penalize unsharp borders
  - \- Penalizes long contours



- Border Edges
  - \+ Penalized edges across objects
  - \- Sensitive to edge noise and missing edges
- Color Space Compactness
  - \+ Penalized bad color distribution
  - \- Clusters can be distorted





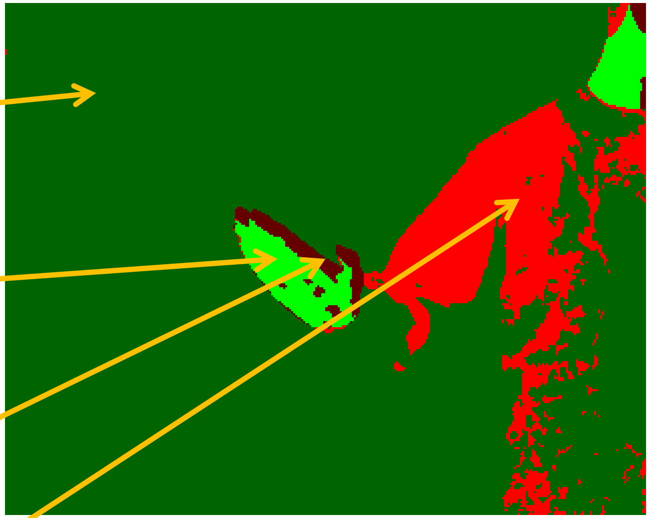
# Ground Truth Data

- 15 data sets
  - Background
    - Simple
    - Complex
    - Skin colored

  - Illumination: most images contain underexposed, normal exposed and overexposed regions

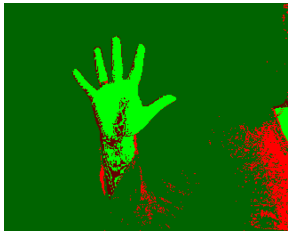
# Cases Possible after Segmentation

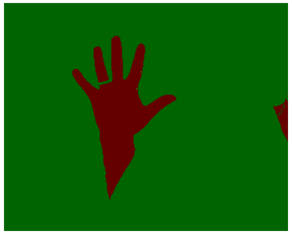
- Correctly classified pixels
  1. True Negatives (TN)
     - non-skin
  2. True Positives (TP)
     - skin
- Wrongly classified pixels
  3. False Negatives (FN)
     - skin classified as non-skin
  4. False Positives (FP)
     - non-skin classified as skin

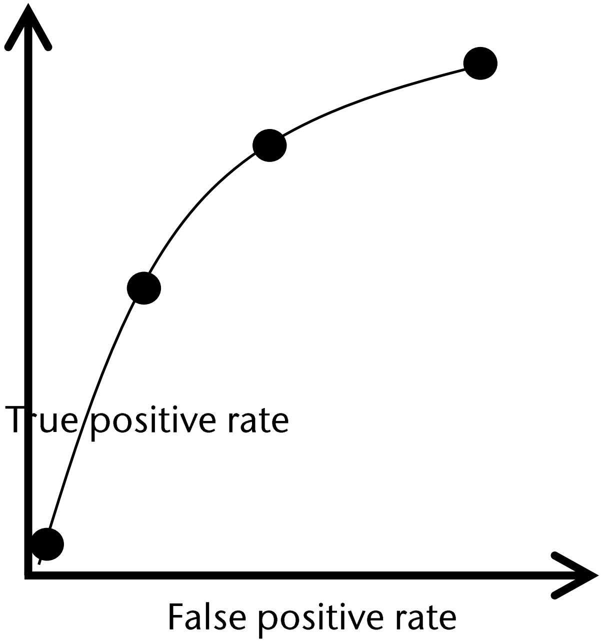


# The Receiver Operating Characteristic Curve

- ROC curve captures relation between
  - True Positive Rate = $\frac{TP}{TP+FN}$
  - False Positive Rate = $\frac{FP}{FP+TN}$

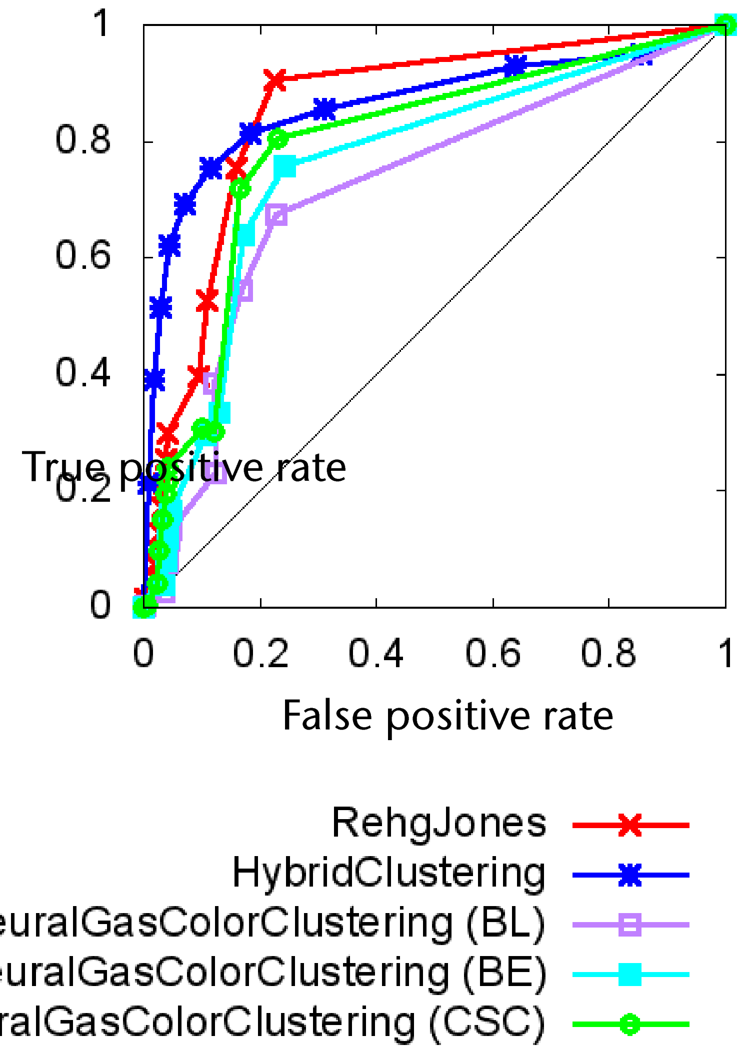






# Main Result: Overall Segmentation Quality

- *HybridClustering* performs best on average
- *NeuralGasColorSpaceClustering* surprisingly has worst quality
  - Color Space Compactness yields better result compared to the other cluster quality measures



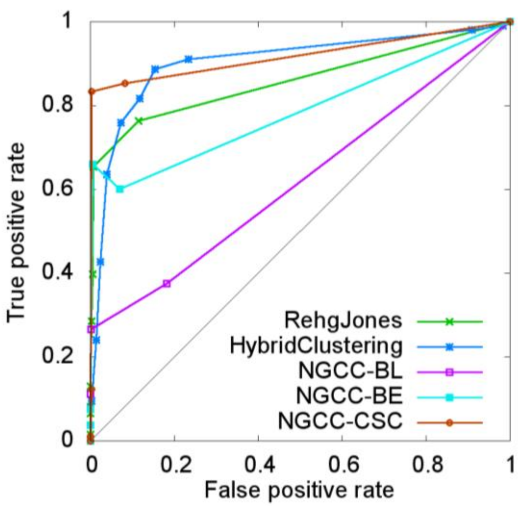
# The 3 most different data sets for detailed analysis

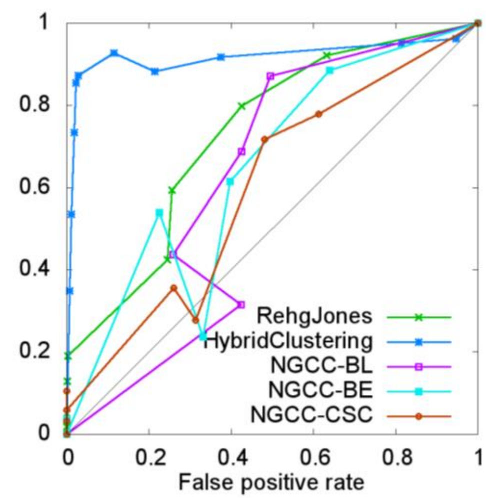

Simple background

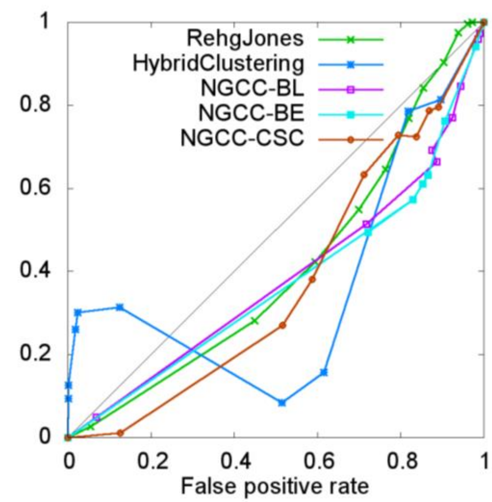

Complex background


Skin colored background

# Individual Segmentation Quality: Data Sets



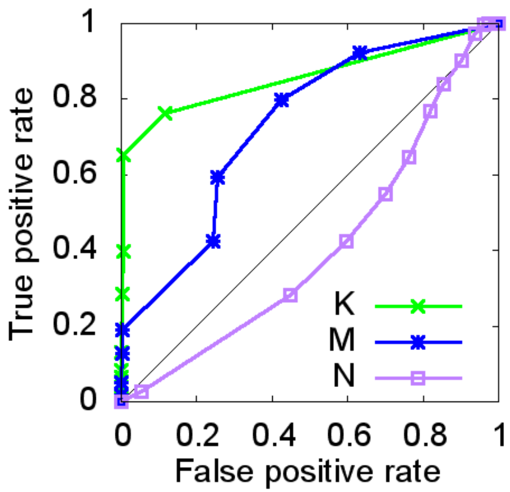



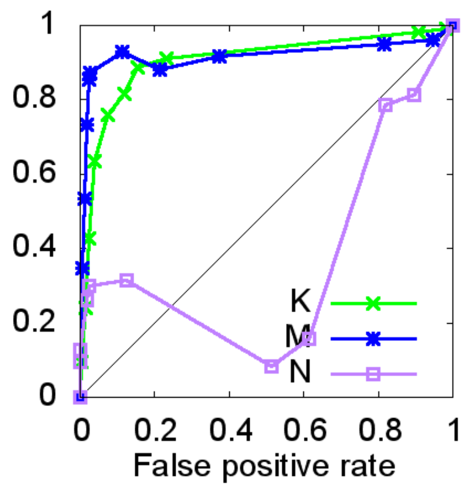

- *HybridClustering* yields best results with high acceptance threshold even for red-door dataset

# Individual Segmentation Quality: Approaches

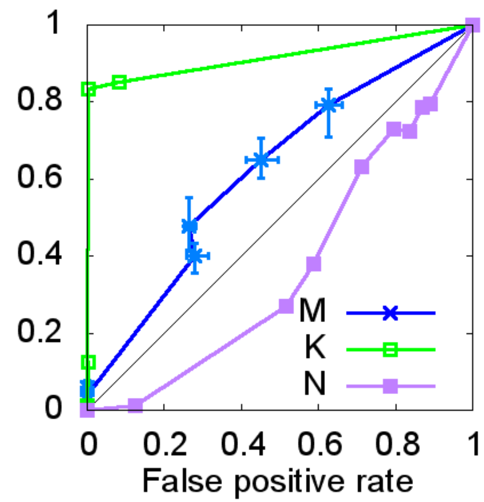
- *RehgJones*




- *HybridClustering*



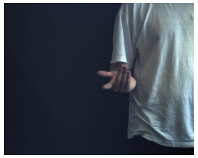

- *NeuralGasColorClustering*



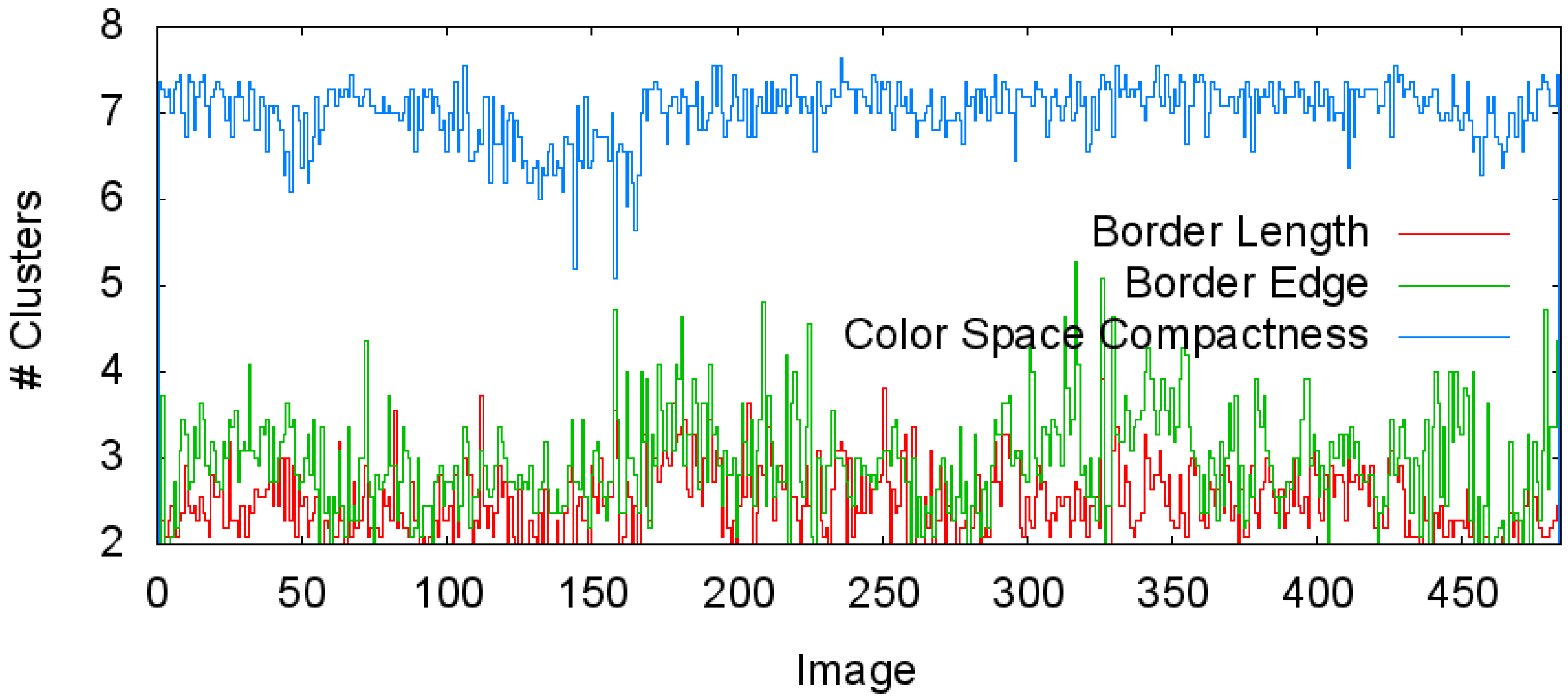

- Moderate variation between different data sets
- Except red-door dataset

- Moderate variation between different data sets

- High variation between different data sets



# Cluster Quality Measures




- Color Space Compactness yields by fast highest number of clusters

⟹ High number of clusters yields better segmentation results

⟹ Better use too many than not enough clusters

# Computation Time

| Approach | Time (ms) | Std. Dev (ms) |
|---|---:|---:|
| RehgJones | 1.23 | 0.06 |
| HybridClustering | 508 | 442 |
| NeuralGasColorClustering - BL | 45 013 | 2 458 |
| NeuralGasColorClustering - BE | 45 886 | 2 635 |
| NeuralGasColorClustering - CSC | 45 460 | 2 961 |

# Conclusion

- Compared the three skin segmentation approaches (*RehgJones*, *HybridClustering*, *NeuralGasColorClustering*
- Method of evaluation:
  - Ground truth dataset of about 500 images
  - ROC curve analysis
- Main result: *HybridClustering* performs best on average
- Detailed analysis reveals high variance between individual datasets
- Apparently, cluster-based segmentation algorithms better use too many cluster than too few

# Future Work

- Further investigate hypothesis about relation between number of clusters and overall segmentation quality
- Evaluate further skin segmentation approaches e.g. [Sigal et al., CVPR 2000]
- Extend ground truth dataset
- Integrate image space smoothing in *NeuralGasColorClustering*

# Thanks for your attention!

# Questions?