Segmentation of Distinct Homogeneous Color Regions in Images

Daniel Mohr

Clausthal University, Germany

mohr@in.tu-clausthal.de

CAIP'07, Aug 2007, Vienna, Austria

Motivation

- 3D hand tracking
  - skin color to detect hands in 2D-images
- Model online skin color distrib. changes
- Good initialization needed
- Jones & Rehg not bad, but false detection rate high at different lighting conditions
  - more robust skin color initialization method needed

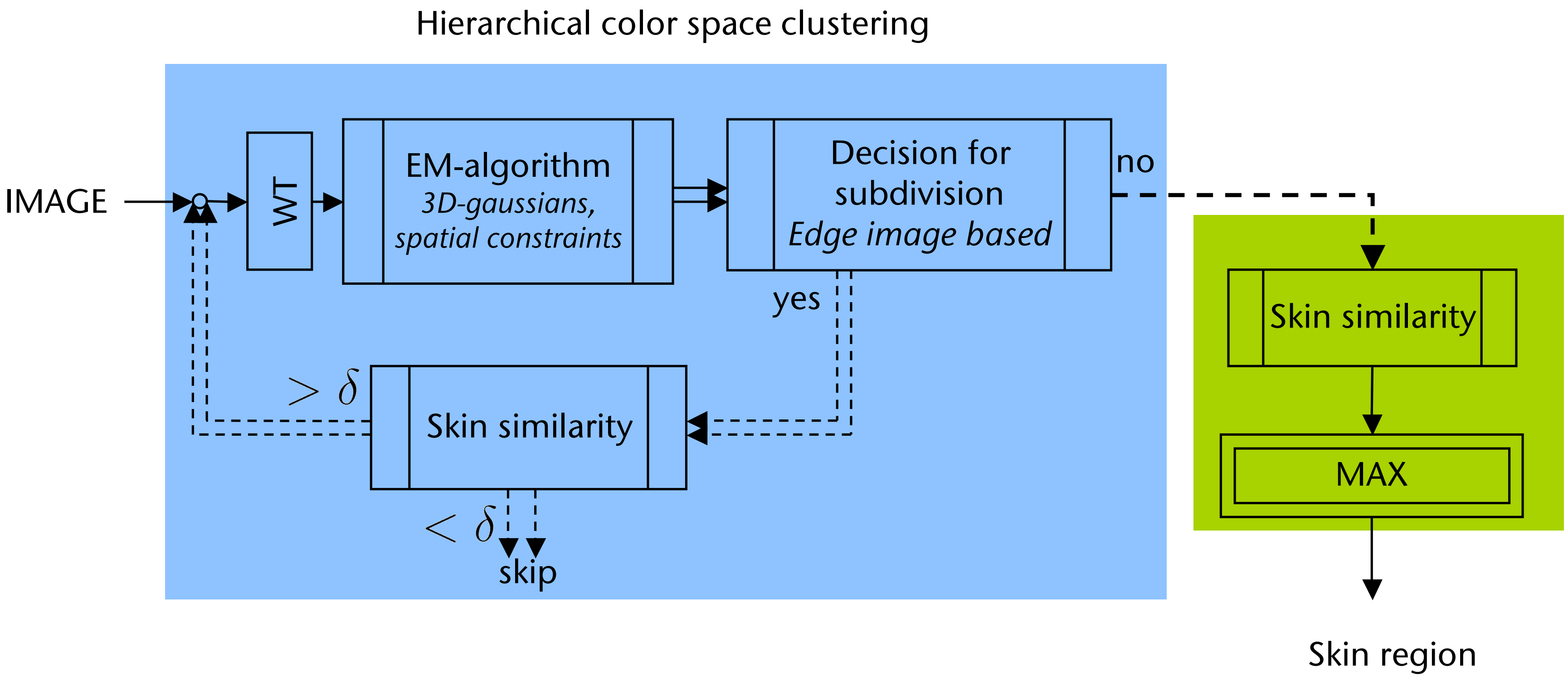
Related Work

- M. J. Jones and J. M. Rehg: Statistical Color Models with Application to Skin Detection, 1999
- M. Wimmer and B. Radig: Adaptive skin color classificator, 2005
- R.L. Hsu, M. Abdel-Mottaleb, A.K. Jain: Face detection in color images, 2002
- J. Yang, W. Lu, A.Waibel : Skin-color modeling and adaptation, 1998
- L. Sigal, S.Sclaroff, V.Athitsos: Skin color-based video segmentation under time-varying illumination, 2004

Overview of our method

- Partially image segmentation
  - Segmentation through color space clustering
    - model object colors as 3D-Gaussians
    - parameter estimation through EM-Algorithm
    - additional spatial constraints
  - Hierarchical subdivision of image
    - subdivision stopping criterion: edge-based
    - select regions: estimate of regions object color similarity
- After segmentation, identify object region(s)
  - Object color similarity

Overview of our method



Hierarchical Clustering

- Number of clusters/image regions a priori unknown
- Divisive approach
  - Process only subregions similar to object color,
    image pixels visited fewer times compared to full segmentation
  - Agglomerative clustering can have quadratic complexity

Original image

Partial segmentation

Full segmentation

Which region to process/skip?

- Based on "skin color similarity" of region
- Variance of image color distribution differs

  →influences object color distribution

Examples

Image 1

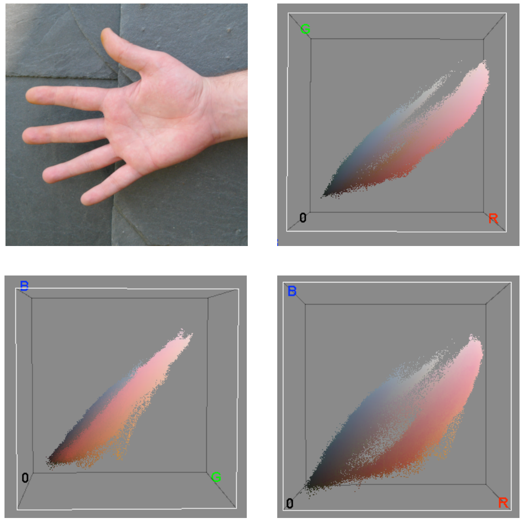
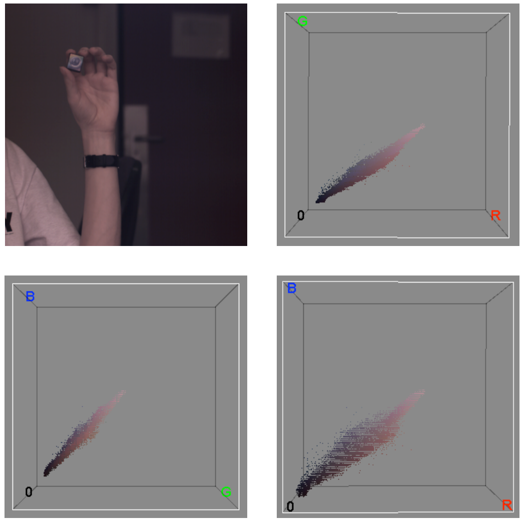


Image 2



Skin color similarity

- Whitening transformed values

$$\tilde{\ }: \mathbf{m} \vdash \left(U \cdot S^{-\frac{1}{2}} \cdot U^T\right)(\mathbf{m} - \bar{\mathbf{m}})$$

with $\left[U, S, U^T\right] = \text{svd}\left(Cov\left(\mathcal{I}_{\mathcal{C}}\right)\right)$

and $\bar{m} = E\left(\mathcal{I}_{\mathcal{C}}\right)$

- Determine skin direction vector
  - small data set of 15 images taken under different illumination
  - skin hand labeled
- Skin similarity measure
  - compare image color region direction vector $\tilde{\mathbf{m}}_i$ with skin direction vector $\tilde{\mathbf{m}}_s$

$$\alpha_i = \frac{\tilde{\mathbf{m}}_s \cdot \tilde{\mathbf{m}}_i}{\parallel \tilde{\mathbf{m}}_s \parallel \cdot \parallel \tilde{\mathbf{m}}_i \parallel \cdot \parallel \tilde{\mathbf{m}}_s - \tilde{\mathbf{m}}_i \parallel}$$

Stopping criterion

- Based on edge image
- Good color segmentation in image space has its borders on image edges
- Quality measure
  - image region $\mathcal{C}$ is splitted into two disjoint regions $\mathcal{C}_1$ and $\mathcal{C}_2$
  - $\mathcal{B}_i = \{\mathbf{x} \mid \mathbf{x} \in \mathcal{C}_i \ \wedge \ \exists \mathbf{x}_j \in \mathcal{N}_3(\mathbf{x}), \ \mathbf{x}_j \notin \mathcal{C}_i\}$,
    $\mathcal{N}_3(\mathbf{x})$ is the 3x3 neighborhood of $\mathbf{x}$ in image space
  - Edge distance map $D(\mathbf{x}_i)$ *(next slide)*

$$\text{split} \quad \Leftrightarrow \quad \frac{1}{|\mathcal{B}_1| + |\mathcal{B}_2|} \sum_{x_i \in \mathcal{B}_1 \cup \mathcal{B}_2} D(\mathbf{x}_i) > \delta$$

Edge distance map $D(\mathcal{I})$

- Generate edge intensity image $C(\mathcal{I})$ through laplace operator
- For each image pixel $\mathbf{x}_j$, store nearest, distance weighted edge intensity

$$D(\mathbf{x}_i) = \max_{\mathbf{x}_j \in \mathcal{N}_k(\mathbf{x}_i)} \frac{C(\mathbf{x}_j)}{\|\mathbf{x}_i - \mathbf{x}_j\| + 1}$$

- $\mathcal{N}_k(\mathbf{x}_i)$ is the $k \times k$ neighborhood of pixel $\mathbf{x}_i$
- Normalize $D(\mathcal{I})$ to [0;1]

Initialization of the EM-algorithm

- EM converges only to local maximum
- Good approximation for initialization needed
- Maximum color difference criterion
  - cluster with respect to pixel's hue value
  - fuzzy-k-means: 1D-data, two clusters
  - avoid clustering in cyclic space
    - Find minimum density hue value and shift it to zero
    - perform fuzzy-k-means

Spatial Constraints

- Use edge distance map $D(\mathcal{I})$
- No edge near pixel

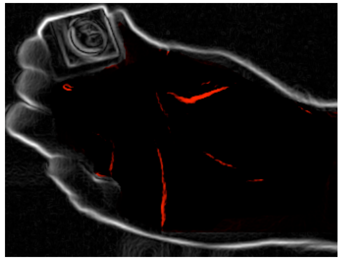
  $\Rightarrow$ pixel belongs to same region
- Edge near pixel

  $\Rightarrow$ maybe region border, maybe not
- For each pixel, new pixel probability to belong to a cluster is calculated through interpolation between
  - pixel probability
  - average pixel neighborhood probability

$$p_{new}(\mathbf{x}|\theta_k) = D(\mathbf{x})p(\mathbf{x}|\theta_k) + (1 - D(\mathbf{x}))\bar{p}(\mathbf{x}|\theta_k)$$

Original image

Edge image



Results

- Image Resolution 250x250
- Spatial Constraints: pixel neighborhood size 3x3
- Edge distance map: filter size 5x5
- System: Athlon 64 X2, 2.0 GHz
- Segmentation process of an image about 0.5 sec.

Results


Indoor, neon light, white skin
Indoor, daylight, white skin
Indoor, daylight, white skin
Outdoor, daylight, dark skin
Original image
Our approach
M.J. Jones and J.M. Rehg

Video

Conclusions

- Segmentation of object with homogeneous color region
- Application to skin segmentation
- EM-algorithm with spatial constraints for color space clustering
- Small training dataset to determine rough direction
- Identify region most similar to object color/skin

Future Work

- Extend to model-based approach
  - average image background vary strong from gray, difficult to identify correct region
- Take model into account to improve stopping criterion
  - e.g. image with poor edges, subdivide not enough or to often, following higher false detection rate
- Extend color distribution model
- Combine multiple images (video sequence) for better color model estimation

Acknowledgment

- Gabriel Zachmann, Clausthal University
- DFG grant DFG grant ZA292/1-1