# Multiple Face Segmentation and Tracking Based on Robust Hausdorff Distance Matching

Chang-Woo Park*, Young-Ouk Kim*, Ha-Gyeong Sung* and Mignon Park**

*Korea Electronics Technology Institute
**Electrical and Electronic Eng., Yonsei University

## Abstract

This paper describes a system for tracking multiple faces in an input video sequence using facial convex hull based facial segmentation and robust hausdorff distance. The algorithm adapts skin color reference map in YCbCr color space and hair color reference map in RGB color space for classifying face region. Then, we obtain an initial face model with preprocessing and convex hull. For tracking, this algorithm computes displacement of the point set between frames using a robust hausdorff distance and the best possible displacement is selected. Finally, the initial face model is updated using the displacement. We provide an example to illustrate the proposed tracking algorithm, which efficiently tracks rotating and zooming faces as well as existing multiple faces in video sequences obtained from CCD camera.

Key words : Robust hausdorff distance, convex hull, face segmentation, tracking

## 1. Introduction

Detecting and locating human faces and facial features in an image or image sequences are important tasks for the applications like model-based coding of video sequences, recognition or identification of faces, and intelligent human-computer interaction. Up to now, much work has been done on detecting and locating faces in gray-scale images and the method like mosaic and template-based[1,2], neural network- based[3], and example-based[4] have been well studied by many researchers. These methods, however, are all computationally expensive and some of them can only deal with the frontal faces with little variation in size and viewpoint. To solve these problems, recently, color-based face detection has become a new direction and exhibited a better performance[5].

Several approaches have been published so far for the detection of facial regions and facial features using color and some other information. For example, Ref. [6] addressed the method to automatically segment out a persons face from a given image that consisted of a head-and-shoulders view of the person and a complex background scene. In Ref. [9], the face-like regions were first detected using color information and then the geometric knowledge of faces was used for further verification. The above two approaches are fast and elaborate, yet they just considered single face and the background was not complex. Ref. [7] introduced a fuzzy pattern matching method for face detection in a complex background by using skin and hair color distribution rate, but it does not work when the facial region is not ellipse-like or the image includes rotating face or turning face. In Ref. [8],

the face-like regions were detected using quantized skin color regions merging and wavelet packet analysis. This method can detect faces with various sizes and different poses; however, it is also computationally expensive due to its complicated segmentation algorithm and time-consuming wavelet package analysis.

Once faces have been segmented in an initial image, they have to be tracked through the video sequence. Tracking facial motion is useful in many applications, such as expression analysis, animation, teleconferencing or face-based biometric person authentication. A robust face tracking algorithm should be insensitive to partial occulations and shadows, to changes in face orientation, scale and lighting conditions, and it should also be computationally simple.

Besides the general purpose tracking methods, there are some approaches that are specially designed to track faces or facial features. This way, the technique presented in Ref. [10] characterizes each feature point using Gabor wavelets and could individually track each point by a phase-based displacement estimation. Other techniques introduced the constrain of real time processing. In this framework, [11] located and tracked the dominant face in the scene based on skin color information. Finally, a snake-based approach for tracking facial features was developed in Ref. [12] that included motion estimation in the energy minimization process of the snake. This allows for large displacements of the contour and for a more robust feature tracking

This paper address a new and effective multiple faces tracking algorithm, which involves face with orientation and a various of size, without much computation using convex hull and robust hausdorff distance. Our method suggests two kinds of problems how to determine the initial face model and how to do the face model update.

In this paper, we determine the initial face model from

intersection relationship (ICH) between the convex hull of skin color regions (SCH) and convex hull of hair color regions (HCH) and track the face of next frames by robust hausdorff distance measure[15,16].

The performance of the proposed algorithm is shown by an experiment. Experimental results show that the proposed algorithm successfully and efficiently tracks rotating and zooming faces as well as multiple faces in a video sequence.

# 2. Skin and Hair Color Segmentation

We will use the color analysis approach for face segmentation. For scope of this paper, we simply illustrate the skin and hair color reference map and limit our attention to the segmentation algorithm using convex hull and tracking algorithm using robust hausdorff distance.

## 2.1 Color segmentation

The head of a human consists of a face region and a hair region.

Therefore, this paper uses two kinds of color reference maps for stable and accurate face segmentation. One is a skin color reference map and another is a hair color reference map.

The first work is classifying pixels of the frame p in video sequence into skin color and non-skin color. In order to do so, we use a skin color reference map in YCbCr color space like [9].

Let's define skin color range of $C_b$ and $C_r$ in $YC_bC_r$ color space as $R_{Cb}$ and $R_{Cr}$, respectively and we call the skin color range as skin color reference map. The $R_{Cb} = [77\ 127]$ and $R_{Cr} = [133\ 173]$ is defined by [9], which propose a set of regularization processes that are based on the spatial distribution and the corresponding luminance values of the detected skin color pixels and this approach overcomes the restriction of motion analysis and avoids the extensive computation of the ellipse-fitting method.

With this skin color reference map, the skin color segmentation can now begin. Lets call the set of pixel of frame extracted from the skin color reference map as the skin likeness region. Since we are utilizing only the color information, the segmentation requires only the chrominance component of the frame.

The skin color likeness is described as

$$S(x, y) = 1, \quad If \quad [C_b(x, y) \in R_{Cb}] \cap [C_r(x, y) \in R_{Cr}] \atop 0 \qquad\qquad Otherwise \qquad (1)$$

For hair color segmentation, we make the hair color region in RGB color space. A race in this paper is limited as an oriental having black hair. Usually, the hair color has an insensitive property For hair color segmentation, we make the hair color region in RGB color space. A race in this paper is limited as an oriental having black hair. Usually, the hair color has an insensitive property to luminance. Then, we use RGB color space for hair color segmentation. Lets define hair color ranges of R, G and B in RGB color space as $R_R = [0\ 30]$, $R_G = [0\ 30]$ and $R_B = [0\ 30]$ in experience, respectively

and we call the hair color range as hair color reference map.

We extract the image points falling inside respective hair color reference map for hair color segmentation.

Lets call the set of pixel of frame extracted from the hair color reference map as the hair likeness region. We can find the set of pixels of the hair likeness regions of a frame in video sequence from equation (2)

$$H(x, y) = 1, \quad if \quad [R(x, y)] \in R_R] \cap [G(x, y) \in R_G] \atop \cap [B(x, y) \in R_B] \qquad (2) \atop 0 \qquad\qquad Otherwise$$

In the next step, we apply the opening operator to remove the noises and we find the skin and hair color region, which is final region of preprocessing. Figure.1 shows skin and hair color region of a frame p in video sequence.



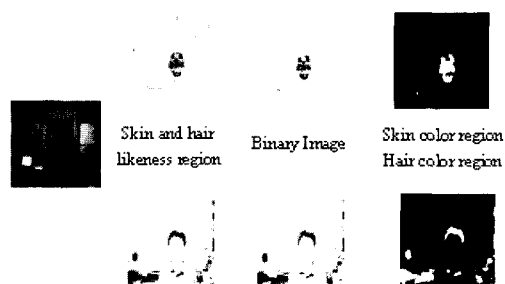Fig.1 Skin and Hair color region

## 2.2 Face Segmentation algorithm based on convex-hull

A convex polygon has the property that any line connecting any two points inside the polygon must itself lie entirely inside the polygon[14]. Then, we define set of points constructing the convex polygon as convex hull.

This paper proposed an effective face segmentation algorithm using the property of convex-hull. Usually, the skin and hair color regions with intersection of them have a very strong possibility that they may be the face and hair.
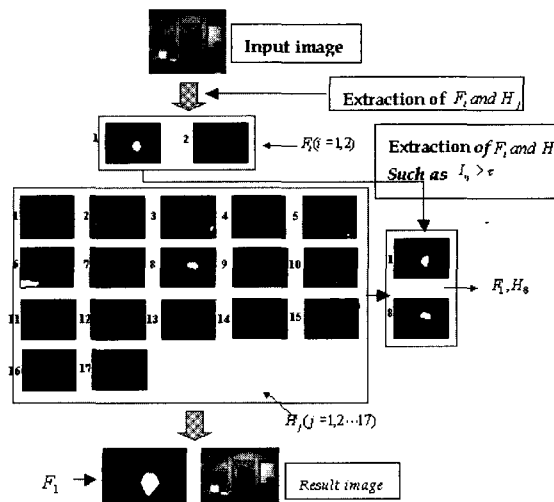


Fig. 2 The overall face segmentation

After assigning label to each color region, we make the set

of the pixels in the convex-hull of hair color region as Hj and that of skin color region and intersection region as $F_i$, $I_{ij}$, respectively. (i = 1~n, j = 1~m)

$$F_{ij} = F_i \cap H_j$$

set the value of pixels in Fi     if $n[I_{ij}] > \tau$
       0                 otherwise       (3)

where, n[.] is the number of element in the set, n and m denote the number of the set of the pixels in the convex-hull surrounding skin likeness regions and that of hair likeness regions.

Figure.2 shows the face segmentation procedure via convex-hull in a frame. Using equation (3), we can select the $F_1$ and $H_8$, where we choose $\tau$ as 10 from experimental result. Consequently, F1 can be determined to be a face location.

# 3. Face Tracking Algorithm

Motion estimation is still very challenging task because the apparent motion flow is highly sensitive to noise. In this section, we explain the robust haudorff distance and propose a new automatic tracking algorithm having good performance in case of existing multiple faces or changing the number of faces in the frame.

## 3.1 Robust Hausdorff Distance

The use of the hausdorff distance for binary image comparison and computer vision was originally proposed by Huttenlocher and colleagues[13]. In this paper, we can track using face model constructing set of outline points surrounding a face region in frame. Let's call the set of outline points surrounding a face region as convex hull's information.

The hausdorff distance measure computes the distance between the points obtained from the convex hull's information in a frame p and the convex hull's information in frame p+1. The conventional hausdorff distance between two point sets, the convex hull's information of initial model in the frame p, $M = \{m_1, ..., m_{N_M}\}$ and that of update model in the frame p+1, $C = \{c_1, ..., c_{N_C}\}$ of sizes NM and Nc respectively, is defined as

$$H(M,C) = \max(h(M,C), h(C,M)) \quad (4)$$
where $h(M,C) = \max_m \min_c \| m\text{-}c \| \quad (5)$
and $h(C,M) = \max_c \min_m \| c\text{-}m \| \quad (6)$

For every model point m, the distance to the nearest convex hull's information c is computed, then the maximum value is assigned to h(M,C). Also, for convex hull's information of current frame c, the maximum distance to the nearest model point m is calculated, then it is assigned to h(C,M). Hausdorff distance is the larger value of the two maxima. That is, h(M,C)=d means that every point of the model is within the distance d from some convex hull's information. But, there are some problems in equation (5) and
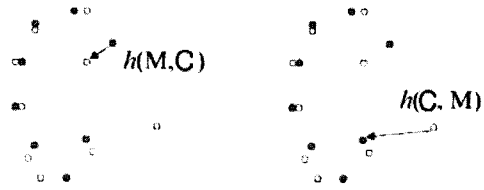
equation (6).



Fig. 3 Hausdorff distance

If the model or object is outlying as can be seen in Figure.3, the hausdorff distance will be very large, even if most of the points are matched well

Thus, we use a robust hausdorff distance measure in [15] proposing a robust hausdorff distance based on the robust statistics such as M-estimation to reduce the problems. The direct distance hR(M,C) is defined as

$$h_R(M, C) = \frac{1}{N_M} \sum_m \rho(d_c(m)) \quad (7)$$

where dc(m) represents the minimum distance at point m to the convex hull's information of current frame C, and the cost function $\rho$ is defined by

$$\rho(x) = \begin{array}{ll} |x|, & |x| \leq \tau \\ \tau, & |x| > \tau \end{array} \quad (8)$$

where $\tau$ is a threshold to eliminate outliers. It is useful to reduce the effect of outliers and to compare portion of images. In this paper, we implement above robust hausdorff distance to track a face in video sequences. The algorithm is as follows.

It is assumed that the hausdorff distance is smaller than a threshold T. A small shape change in all directions, the displacement $t \in T$ of a face model is obtained by minimizing the following formula :

$$\sigma = \min_t H_{M,t} (M, C) \quad (9)$$

That is, $h_M(M+t, C)$ and $h_M(C-t, M)$ and are computed, then the larger of the two is selected as $h_{M,\tau}(M, C)$. When it is made the smallest value by t, t is selected as the best possible displacement.

## 3.2 The initializing of a face model

First of all, the face position has to be determined. In section 2, we presented a face model in an input image frame using face segmentation algorithm. Let's call the convex
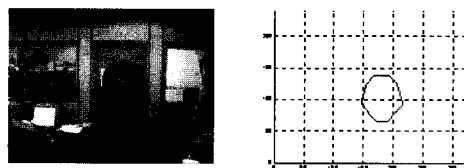


Fig. 4 The initial model

hull's information of the frame p as Mx (x=1~m') and m' denotes the number of the convex hull' s information in the frame p.

The right figure on the figure.4 shows convex hull' s information in the frame p.

### 3.3 Model update

Since the tracked object can rotate or change its shape, it is necessary to update the corresponding model every frame. We invest the face regions using face segmentation algorithm to the new model of frame p+1 for finding the best possible displacement t. Let' s call the convex hull' s information of the frame p+1 as $C_y$(y=1~n') and n' denotes the number of the convex hull' s information in the frame p+1.

If the x is equal to the y, we shift the face model of the frame p to the new position of best match to update these parts, which was located by the robust hausdorff distance or else we consider the face model of frame p as initial model for tracking the new face again.

Model update          if x=y
Initialization of a face model     otherwise      (10)

When the model of the frame p is shifted to the new position with the displacement t. the part of convex hull' s information with t that belongs to the updated model is selected by the following equation.

$$M_x' = [c \in C_y \mid h(C_y - t, M_x)]$$

where, $M_x'$ denotes update model       (11)

The selection of $\sigma$ is very critical. If $\sigma$ is too small, we have little tolerance to the shape change. If it is too large, some nearby noises are determined as new model. In our experiment, it is 3 pixels is used.

## 4. Experimental results

We have applied the proposed tracking algorithm to track multiple faces in video sequence containing faces, which is obtained by CCD camera. The obtained images are 160 120 RGB model images of laboratory environment. The host computer administering the hole system uses IBM PC Pentium III 700 MHz.

First, we segment skin and hair likeness regions from the background using skin and hair color reference map. An initial face model is obtained from SCH with ICH. Finally, the model is updated or intialized according to relationship between the number of convex hull' s information of the frame p and that of the frame p+1. Figure.5 and Figure.6 show skin and hair color regions and the initial model in the frame p, respectively. Figure. 7 represents the tracking results. Also, this paper can track multiple faces from video frame with robust hausdorff distance and the algorithm has the robust property when the frame includes rotating or zooming face. Figure 8 and 9 illustrate their properties, respectively.

## 5. Conclusion

A new tracking algorithm based on robust hausdorff distance was presented. A model of the face was automatically derived and matched against subsequent frames using the robust hausdorff distance. To accommodate for rotation and changes in shape, the model was updated every frame by a novel update technique that consists of two components for slowly and rapidly changing or moving parts. Matching the model based on robust hausdorff distance using convex hull' s information is remarkably robust from the noise. The new position is accurately detected even when the face undergo large changes in shape or the background is moving because of analyzing the color of successive frames using the proposed face segmentation algorithm. Also, the system can track the multiple moving faces from sequence or the rotating face.



(a) The frame p
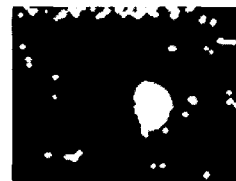


(b) Skin color regions       (c) Hair color regions
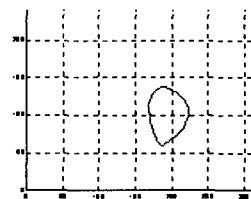Fig. 5 Skin and Hair color region in 1st frame



Fig. 6 Initial face model



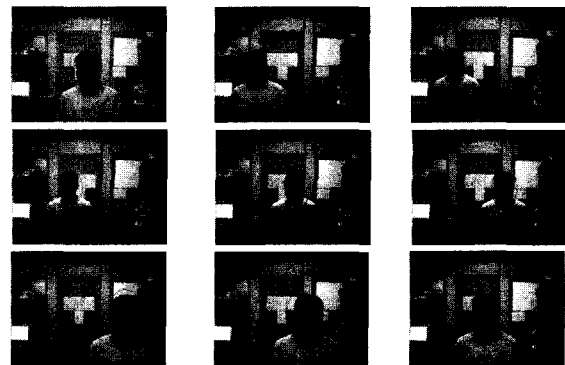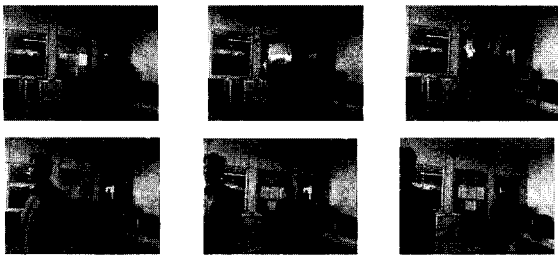Fig. 7 Tracking results

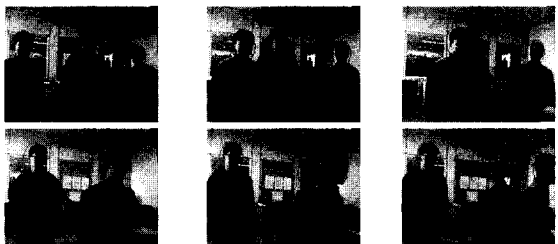Fig. 8 Two people tracking



Fig. 9 Zooming and rotation tracking



Fig. 10 Tracking in case of closing two faces

## Reference

[1] Yang, G.Z., Huang, T.S.: 'Human face detection in a complex background' , Patten Recognition., 1994, 27, (1), pp. 43-63

[2] Miao, J., Yin, B., Wang, K., Shen, L., Chen, X.: 'A hierarchical multiscale and multiangle system for human face detection in a complex background using gravity-center template' , Pattern Recognition, 1999, 32, (7), pp. 1237-1248

[3] Rowley, H.A., Baluja, S., Kanade, T.: 'Neural network-based face detection' , IEEE Trans. Pattern Anal. Mach. Intell. 1998, 20, (1), pp. 23-38

[4] Sung, K. K., Poggio, T.: 'Example-based learning for view-based human face detection' , IEEE Trans. Pattern Anal. Mach. Intell. 1998, 20, (1), pp. 39-51

[5] Lee, C. H., Kim, J.S., Park, K.H.: 'Automatic human face location in a complex background using motion and color information' , Pattern Recognition, 1996, 29, (11), pp. 129-140

[6] Sobottka, K., Pitas, I.: 'A novel method for automatic face segmentation, facial feature extraction and tracking' , Signal Process. Image Commun., 1998, 12, (3), pp. 263-281

[7] Wu, H., Chen, Q., Yachida, M.: Face detection from color images using a fuzzy pattern matching method, IEEE Trans. Pattern Anal. Mach. Intell., 1999, 21, (6), pp. 557-563

[8] Garcia, C., Tziritas, G.: 'Face detection using quantized skin color regions merging and wavelet packet analysis' , IEEE Trans. Multimedia, 1999, 1, (3), pp. 264-277

[9] Chi, D., Ngan, K.N.: 'Face Segmentaion Using Skin-Color map in Videophone Applications, IEEE Trans. Circuits and systems for video technology' , June,1999, 9, (4), pp. 551-564.

[10] McKenna, S.J., Gong, S., Wurtz, R., Banin, D., Tanner,, J.: 'Tracking facial feature points with gabor wavelets and shape models' , Proceedings of the first International Conference on Audio and Video Based Biometric Person Authentication, Crans-Montana, Swotzerland, March 1997, Springer, Computer Science, Vol, 1206, pp.35-42

[11] Qian, R., Sezan, M.I., Matthews, K.E.: 'A robust real-time face tracking algorithm' , Processing of the IEEE International Conference on Image Processing, ICIP-98, Chicago, USA, 1998, pp. 131-135

[12] Pardas, M., Sayrol, E.: 'Tracking of contours combining snakes with motion estimation' , Proceding of the IEEE International Conference on Image Processing, ICIP-00, Vancouver, Canada, September 2000, pp. TA07.14.

[13] Huttenlocher, D. P., Klanderman, G. A., Rucklidge W. J.: 'Comparing Images Using the Haudorff distance' , IEEE. Trans. Pattern Anal. Machine. Intelligence 1999, 15, (9), pp. 850-863

[14] Hussain, Z.: ' Digital Image Processing' , ELLIS HORWOOD LIMITED in New York in London.

[15] Sim, D., Kwon, O.: 'Rae-Hong Park, Object Matching Algorithms Using Robust Hausdorff Distance Measure s' . , IEEE Trans. Image Processing, March, 1999, 8, (3), pp. 425-429.

[16] Huttenlocher, D.P., Klanderman, G. A., Rucklidge, W.J.: 'Comparing Images Using the Hausdorff Distance' , IEEE Trans. Pattern Analysis and Machine intelligence, Sep, 1993, 12, (9), pp. 850-863

**Chang-Woo Park**
received the B.S. degree in electronics from Korea University in 1997 and M.S. and Ph.D. degrees in electronics from Yonsei University, Seoul, Korea, in 1999 and 2003, respectively. He has been a researcher in Korea Electronics Technology Institute under the Ministry of Commerce, Industry and Energy. His current research interests include intelligent system, fuzzy system and robotics.

Phone : +82-32-621-2854
Fax : +82-32-621-2855
E-mail : drcwpark@keti.re.kr

**Young-Ouk Kim**

received B.S., M.S degrees in electrical engineering from Hong-Ik University in 1995, 1997. Since 2003, he has been under Ph. D. course at Department of the Graduate School of Advanced Imaging Science, Multimedia and Film in Chung-Ang University. Since 1997, he has been a senior researcher in Korea Electronics Technology Institute under the Ministry of Commerce, Industry and Energy. His research interests are video tracking, automatic zoom, face recognition, intelligent space, HCI theory and applications.

Phone : +82-32-621-2847
Fax : +82-32-621-2855
E-mail : kimyo@keti.re.kr

**Ha-Gyeong Sung**

received the B.S. and M.S. degrees in mechanical engineering from Hanyang University in 1986 and 1995 respectively and received Ph.D. degree in mechanical engineering from Ajou University, Korea, in 2003. He has been a chair of center in Korea Electronics Technology Institute under the Ministry of Commerce, Industry and Energy. His current research interests include human computer interface, motor design and tele-operating system.
received

Phone : +82-32-621-2841
Fax : +82-32-621-2855
E-mail : sunghg@keti.re.kr

**Mignon Park**

received the B.S. and M.S. degrees in electronics from Yonsei University, Seoul, Korea, in 1973 and 1977, respectively, and the Ph. D. degree from the University of Tokyo, Tokyo, Japan, in 1982. Since 1982, he has been a Professor in the Electrical and Computer Department of Yonsei University. His research interests include fuzzy control and application engineering, robotics, and fuzzy biomedical system, etc.

Phone : +82-2-2123-2868
Fax : +82-2-312-2333
E-mail : mignpark@yonsei.ac.kr