Face Tracking Using Skin-Color and Robust Hausdorff Distance in Video Sequences

박 정 호, 박 창 우, 박 민 용
연세대학교 전기 · 컴퓨터 공학과

Jungho Park, Changwoo Park, and Mignon Park
Dept. of Electrical · Computer Engineering, Yonsei University
Tel: (02) 361-2868, Fax: (02) 312-2333, jhpark@yeics.yonsei.ac.kr

ABSTRACT

We propose a face tracking algorithm using skin-color based segmentation and a robust Hausdorff distance. First, we present $L^*a^*b^*$ color model and face segmentation algorithm. A face is segmented from the first frame of input video sequences using skin-color map. Then, we obtain an initial face model with Laplacian operator. For tracking, a robust Hausdorff distance is computed and the best possible displacement t is selected. Finally, the previous face model is updated using the displacement t. It is robust to some noises and outliers. We provide an example to illustrate the proposed tracking algorithm in video sequences obtained from CCD camera.

I. INTRODUCTION

Face tracking is an challenging and important problem in computer vision. This subject has been attracting researchers for many years. It has a wide request for many applications such as, interactive human-computer communications, virtual reality interfaces, and automated surveillance systems and video conferencing. The complexity of these issues depends on environment under which the algorithm is applied. In this paper, we applied the tracking algorithm to face tracking in video sequences obtained from CCD camera in laboratory environment.

So far, many different approaches for face segmentation or detection and tracking have been proposed in literature [6]. Eigenfunction and neural networks(NNW) are the commonly used techniques to locate and detect human faces. Eigenfunctions are specific for a particular pattern. So, it has limitation of the variation on viewpoint. NNWs perform well to detect faces in still images but difficult to train on scenes with no-faces in the scenes. Skin-color based approach is simple and non-parametric forms of density estimation, but if there is not sufficiently accurate model for apparent color, good parametric models for density estimation cannot be obtained. Feature-based approach has the flexibility to be extended to a difficult scale, orientation and viewpoints of faces.

In this paper, we propose an approach using skin-color based segmentation[1] and model matching by Hausdorff distance measure[2, 4, 3]. In the first frame, we segment a face and set a face model. Then, we track the face using Hausdorff distance measure.

The remainder of this paper is organized as follows. Section 2 describes color models and a face segmentation algorithm, which is improved on the basis of skin color map. In section 3, we describe robust Hausdorff distance, which is used to determine the best possible displacement. Also initialization and update the face model are given. Experimental results are reported in section 4. Finally, we conclude in section 5.

II. SKIN-COLOR BASED SEGMENTATION

Segmentation of a face in video sequences is a difficult job because of occlusion, cluttered backgrounds, and variation of viewpoint and illumination. The importance of face segmentation is illustrated by its many applications, such as, object-based image encoding, modeling, face recognition and face tracking. Recently there have been many approaches on face segmentation, but no single method assures robustness. Some proposed approaches involved the combination of shape, motion, and statistical analysis for robustness[6]. In these days, a new approach using color information is proposed. In this paper, we will use the color analysis approach for face segmentation. It is fast, and invariant to orientation and viewpoints. But its principal limitation is high sensitivity to illumination and occlusion.

A. Color Model

An image can be described by several different color space models. RGB model is a most commonly used hardware-oriented model. It is simple but very sensitive to illumination. HIS model decouples the intensity component I from the chrominance information. Hue and saturation component represent pure color and a measure of the degree to which a pure color is diluted by white light. They are intimately related to the way in which human beings perceive color. $L^*a^*b^*$ model is another hardware-oriented model. Unlike RGB model, it separates luminance from chrominance information. Here L^* represents luminance, a^* and b^* represent chrominance components. HIS and $L^*a^*b^*$ are have been reported to be suitable for discrimination color information for modeling skin-color. In this paper, we use the $L^*a^*b^*$ model because of simple analysis. If there is a good contrast between skin and background, a simple segmentation can provide reliable results. Otherwise, it is difficult to segment exact facial region.

B. Face Segmentation Algorithm

In this paper, a face is segmented from the first frame of input image sequences to track a face. We assume that there is always a face in the input image sequences. The first frame of RGB image sequences obtained by CCD camera is transformed into $L^*a^*b^*$

▣ The first step of the algorithm is to classify pixels of the input image into skin region and non-skin region. Here we use the skin-color map in $L^*a^*b^*$ color model. The result of this step is described as

$$S(x,y)=\begin{cases}1, & \text{if } [a^*(x,y)\in R_{a^*}]\cap[b^*(x,y)\in R_{b^*}]\\ 0, & \text{otherwise}\end{cases}$$

where R_{a^*} and R_{b^*} are respective skin-color ranges of a^* and b^* in CIE chromaticity diagram.

▣ The second step of the algorithm is to remove the small area of background image. Thus we investigate the density distribution of the previous result. With 3×3 window we obtain density map, $D(x,y)$, which counts the skin-color pixels in each window. If $D(x,y)$ is 9, it is classified as full-density point. Once the density map is obtained, erode any full-density point(i.e., set to zero) surrounded by less than four full-density points.

Ⅲ. FACE TRACKING ALGORITHM

Motion estimation is still a very challenging task. Because the apparent motion flow is highly sensitive to noise. Thus, a robust matching algorithm must be able to detect moving faces which are rotating and translating in shape. If a face is segmented and facial edges are reliably detected, the face can be tracked over time in an image sequences. In this Section, we propose a new automatic tracking algorithm using robust Hausforff distance.

A. Robust Hausdorff Distance

Color or gray scale imges are not suitable for template or object matching because they are too sensitive to changes in illumination. Thus, low-level matching algorithms using distance transform and Hausforff distance have been investigated. They are simple and insensitive to the changes of image characteristics.

The Hausdorff distance measure computes the distance between two sets of edge points obtained from the model of previous frame and an Edge of current frame.

In our algorithm, the edge images are obtained by Laplacian operator. The conventional Hausdorff distance between two point sets, model, $M = \{m_1, \cdots, m_{N_M}\}$ and Edge, $E = \{e_1, \cdots, e_{N_E}\}$ of sizes N_M and N_E, is defined as

$$H(M,E) = \max\ (h(M,E),\ h(E,M)) \qquad (1)$$

where

$$h(M,E) = \max_{m \in M} \min_{e \in E} \|m - e\| \qquad (2)$$

and

$$h(E,M) = \max_{e \in E} \min_{m \in M} \|e - m\| \qquad (3)$$

For every model point m, the distance to the nearest Edge point e is computed, then the maximum value is assigned to $h(M,E)$. Also, for every Edge point e, the maximum distance to the nearest model point m is calculated, then it is assigned to $h(E,M)$. Hausdorff distance is the larger of the two maxima. That is, $h(M,E) = d$ means that every point of the model is within the distance d from some Edge points. But There are some problems in equation (2) and (3). If the model or object is outlying as can be seen in Fig. 1., the Hausdorff distance will be very large, even if most of the points are matched well. Thus, we use a robust Hausdorff distance measure to reduce the problems. It is described as

$$h_R(M,E) = \frac{1}{N_M} \sum_{m \in M} \rho(d_E(m)) \qquad (4)$$

where $d_E(m)$ represents the minimum distance at point m to the Edge set E, and the cost function ρ is defined by

$$\rho(x) = \begin{cases} |x|, & |x| \le \tau \\ \tau, & |x| > \tau \end{cases}$$

where τ 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 :

$$\sigma = \min_{t \in T} H_t(M,E) \qquad (5)$$

That is, $h_R(M+t,\ E)$ and $h_R(E-t,\ M)$ are computed, then the larger of the two is selected as $H_t(M,E)$. When the Hausdorff distance, $H_t(M,E)$ is smallest, t is selected as the best possible displacement.

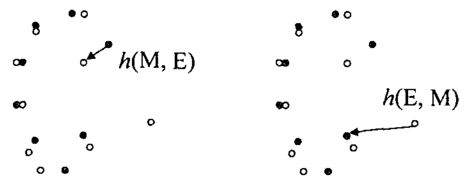


Fig. 1 Hausdorff distance. (a) $h(M,E)$, (b) $h(E,M)$, In this figure, Hsusdorff distance is equal to $h(E,M)$.

B. Initialization of Face Model

First of all, the face position has to be determined. In section 2, we presented the face segmentation algorithm in the first input image frame using skin-color information. The initial model can be extracted by simple computation of the difference between consecutive frames, but it is too noisy. Thus, we propose a different edge based approach. Face edges are obtained by Laplacian operator. Then, it is set to the initial model for Hausdorff distance measure. Using the model and Edge set of current frame, we calculate the best possible displacement.

C. Update Model

After the best possible displacement t is obtained, the face model is updated. Here, we determine which part of the Edge set must be included in the updated model. The model of the previous frame is shifted to the new position with the displacement t. With t, the part of Edge set which belongs to the updated model is selected by the following equation.

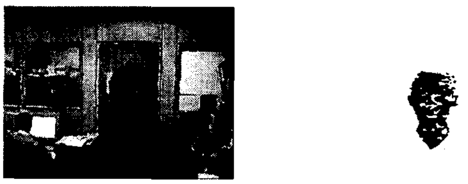
$$M_n = \{ e \in E_n \mid h_R(E_n - t, M_{n-1}) \le \sigma \} \qquad (6)$$

The selection of σ is very critical. If σ is too small, we have little tolerance to the shape change.

If σ is too large, some nearby noises are determined as new model. In our experiment, σ of 3 pixels is used.

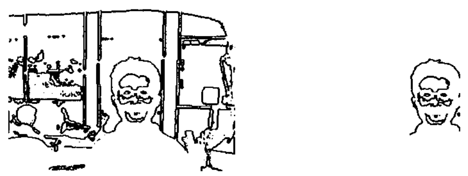
IV. EXPERIMENTAL RESULTS

We have applied the proposed tracking algorithm to track a face in image sequences containing a face, which is obtained by CCD camera. The obtained images are 320×240 RGB model images of laboratory environment. First, we transform the RGB model images into $L^*a^*b^*$. We then segment the face from the background using skin-color map. A face model is obtained by Laplacian operator. Finally, the displacement is computed by a robust Hausdorff distance and the model is updated. Fig. 5. shows the tracking results



(a)First frame image (b)Segmented face

Fig. 2. The result of face Segmentation.



(a) Edges of the first frame (b) Obtained Face model

Fig. 3. Face model

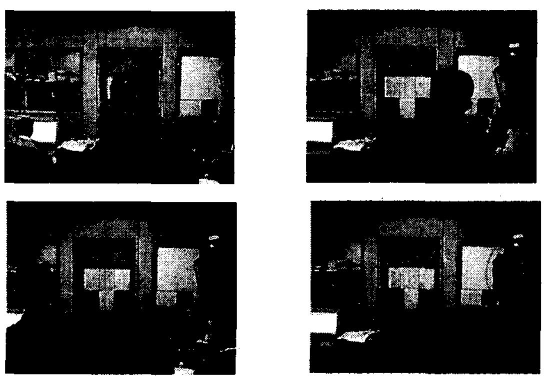


Fig. 4. Tracking Results

V. CONCLUSION

In this framework, we have presented an automatic approach for face segmentation and tracking over time. Face segmentation is done by $L^*a^*b^*$ color model and skin-color map in a first image frame. Once faces are segmented, we extract a face model, which can be tracked over time. Tracking the face is performed by computing displacement using a robust Hausdorff distance. If a best possible displacement is found, the face model is updated. The robustness of our approach was tested on different color image sequences containing faces. More research are required for the cases including occlusion, severe variation of illumination, and rotation of face.

VI. REFERENCES

[1] Douglas Chai and King N. Ngan, "Face Segmentation Using Skin-Color Map in Videophone Applications", IEEE Trans. PAMI, vol. 9, No.4, JUNE 1999, pp.551-564.

[2] Thomas Meier and King N. Ngan. "Automatic Segmentation of Moving Objects for Video Object Plane Generation", IEEE Trans. Circuits and Systems for Video Technology vol. 8, No. 5, Sept. 1998, pp. 525-538.

[3] Dong-Gyu Sim, Oh-Kyu Kwon, and Rae-Hong Park, "Object Matching Algorithms Using Robust Hausdorff Distance Measures", IEEE Trans. Image Processing, vol. 8, No. 3, March 1999, pp. 425-429.

[4] Hertog Nugroho and Shinji Ozawa, "Tracking Multiple Moving Objects from Monocular Image Sequences with Hausdorff Distance and Context Information", T. IEE Japan, vol.118-C, No 7/8, 1998, pp. 1138-1147.

[5] Liang Zhang, "Tracking a face for knowledge-based coding of videophone sequences", Signal Processing: Image Communication, V.10 N.1-3, Oct. 1997, pp. 93-114.

[6] V. Ronda, M. H. Er., W. Ser, "Face Detection, tracking and recognition-A Study", Proceedings of ICARCV '98, Singapore, Dec. 1998, pp. 50-55