LOCALIZATION OF FACIAL REGIONS AND FEATURES IN COLOR IMAGES

Karin Sobottka Ioannis Pitas

Department of Informatics, University of Thessaloniki 540 06, Greece

e-mail:{sobottka, pitas}@zeus.csd.auth.gr

Index Terms: Face recognition, color image processing, watersheds.

1 Introduction

Recognition of human faces out of still images or image sequences is a research field of fast increasing interest. There are many different applications for systems coping with the problem of face localization and recognition, e.g. model-based video coding, security systems, mug shot matching. Due to variations in illumination, background, visual angle and facial expressions, the problem is complex. A critical survey about existing literature on human and machine recognition of faces is given in [1].

In a first step of face recognition, the localization of facial regions and the detection of facial features as eyes, nose and mouth is necessary. In this framework we present an approach that copes with problems of this first step.

For the detection of facial regions and facial features, approaches have been published using texture, depth, shape and color information or combinations of them ([2], [4], [3], [6]).

Our approach is based on the observation that human faces are characterized by their oval shape and their skin-color, also in the case of varying light conditions. Therefore we locate face-like regions on the base of shape and color (HSV) information. The hypotheses for faces are verified by searching for facial features inside of the face-like regions. For that we first enhance facial features by applying morphological operations and then we extract candidates for facial features by combining thresholding and a modified watershed method.

2 Localization of face-like regions

Faces are significantly characterized by their oval shape and specific skin-color. For that reason, we have decided for an approach based on shape and color information. The effectiveness of using color information was also shown in [2].

As discriminating color information we consider the attributes hue and saturation. These attributes are very similar to the human impressions of color, which are different from the more often used primary components of Red, Green and Blue [9].

(a) (b) (c)

Figure 1: Detection of facial regions

To obtain robustness against changes in light conditions, we perform region segmentation by considering appropriately defined domains in the hue and saturation domain. An example of such a segmentation is shown in Figure 1a,b. The skin of the hands and of the face are well localized. There are some falsely detected pixels at the shirt and in the background.

The oval shape of a face can be approximated by an ellipse. Therefore looking for faces in an image means to detect objects with nearly elliptical shape. This can be done based on edges [3] or, as we will show here, based on regions. The advantage of considering regions is that they are more robust against noise and changes in illumination. In a first step connected components are determined and then we check for each connected component, if its shape is nearly elliptical or not.

We determine connected components by applying a region growing algorithm at a coarse resolution of the segmented image. Then for each connected component C with a given minimum size, the best-fit ellipse E [5] is computed on the base of moments.

An ellipse is exactly defined by its center $(\overline{x}, \overline{y})$, its orientation θ and the length a and b of its minor and major axis.

The center $(\overline{x}, \overline{y})$ of the ellipse is given by the center of gravity of the connected component. The orientation θ of the ellipse can be computed by determining the least moment of inertia. As result

$$\theta = \frac{1}{2} \cdot arctan\left(\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}}\right) \tag{1}$$

is obtained, with $\mu_{i,j}$ denoting the central moments of the connected component.

The length of major and minor axis of the best-fit ellipse can also be computed by evaluating the moments of inertia. With I_{min} the least and I_{max} the greatest moment of inertia of an ellipse with orientation θ,

$$I_{min} = \sum_{(x,y)\in C} [(x-\overline{x})cos\theta - (y-\overline{y})sin\theta]^2 \tag{2}$$

$$I_{max} = \sum_{(x,y)\in C} [(x-\overline{x})sin\theta - (y-\overline{y})cos\theta]^2 \tag{3}$$

the length a of the major axis and the length b of the minor axis results in:

$$a = \left(\frac{4}{\pi}\right)^{1/4} \left[\frac{(I_{max})^3}{I_{min}}\right]^{1/8} \qquad b = \left(\frac{4}{\pi}\right)^{1/4} \left[\frac{(I_{min})^3}{I_{max}}\right]^{1/8}. \tag{4}$$

On the base of the already computed elliptical parameters, a reduction of the number of potential face candidates is possible. This is done by applying to each ellipse decision criteria concerning its orientation and the relationship between major and minor axis.

For the remaining candidates we assess, how well the connected component C is approximated by its best-fit ellipse E. For that purpose the "holes" inside of the ellipse and the points of the connected component that are outside of the ellipse are counted. Then the ratio of the number of false points to the number of points of the interior of the ellipse is calculated. Based on a threshold on this ratio, the ellipses that are good approximations of connected components are selected and considered as face candidates.

For the example, we obtain the results shown in Figure 1c. The ellipses are drawn with white color into the image of the detected connected components.

Subsequently the determined face candidates are verified by searching for facial features inside of the ellipses.

3 Extraction of facial features

Our approach to the extraction of facial features is based on the observation that, in intensity images, eyes and mouth differ from the rest of the face because of their lower intensity. The reason for that is the color of the pupils, the sunken eye-socket and the light red color of the lips. Even if the eyes are closed, the darkness of the eye sockets is sufficient to extract the eye regions. Therefore, in the following we consider the intensity information in the interior of the ellipses (Fig. 2a).

In a preprocessing step, we enhance the dark regions inside of the ellipses by using morphological operations. First we apply a grayscale erosion [8] and then an extremum sharpening operation [7] to improve the contrast of the interior of the ellipse. Results are illustrated in Fig. 2b.

The extraction of candidate regions for eyes and mouth is based on the idea to immerse the graylevel relief into water up to a predefined depth and then flood the cutted minima by a modified watershed method. The first step, the immersion of the graylevel relief up to a predefined depth, is realized by thresholding. Then connected components are determined and different labels are assigned to the pixels of different connected components. This result is considered as initial segmentation and we use it as input for our modified watershed method. A very efficient realization of conventional watershed algorithm is given in [10]. We use it partially to obtain also a very efficient modified version for our application.

First the pixels of the interior of the ellipse are sorted according their graylevels. This ensures random access to the pixels and direct access to the neighbours. Beginning with the initial regions, determined by the thresholding step, the regions are slowly flooded. This is done by applying an iterative method. Considering the local contrast ensures that the merging process stops, also if there is no other region as direct neighbour.

(a) (b) (c)

Figure 2: Extraction of candidates for facial features

The resulting image is shown in Figure 2c. Besides of regions for eyes and mouth, also regions for beard and hair are extracted. They are removed by postprocessing methods.

4 Conclusion

We have presented an approach that detects facial regions in color images and calculates facial features.

Facial parts are determined on the base of color and shape information. Therefore, we first extract regions with skin-like hue and saturation values and then compute the best-fit ellipse for each of the regions. Based on a function that assesses how well the regions are approximated by their best-fit ellipse, we choose potential face candidates. These hypotheses are verified by searching for facial features inside of the ellipses.

On the base of the observation that eyes and mouth differ from the rest of the face because of their lower intensity, we first enhance the dark regions inside of the ellipse by applying morphological operations and then extract candidates for them by using thresholding and a modified watershed method. The success of this method was verified on a large number of color images containing faces.

References

[1] R. Chelleppa, C. L. Wilson, S. Sirohey, "Human and Machine Recognition of Faces: A Survay," *Proc. of IEEE,* vol. 83, no. 5, pp. 705-740, May 1995.

[2] Y. Dai, Y. Nakano, "Extraction of Facial Images from Complex Background Using Color Information and SGLD Matrices," *International Workshop on Automatic Face- and Gesture Recognition*, pp. 238-242, ed. Martin Bichsel, Zurich, Switzerland, June 1995.

[3] A. Eleftheriadis, A. Jacquin, "Automatic face location and tracking for model-assisted coding of video teleconferencing sequences at low bit rates," *Signal Processing: Image Communication*, vol. 7, no. 3, pp. 231-248, July 1995.

[4] G. Galicia, A. Zakhor, "Depth Based Recovery of Human Facial Features from Video Sequences," *IEEE Conference on Image Processing*, vol. 2, pp. 603-606, Washington, USA, October 1995.

[5] A. K. Jain, "Fundamentals of Digital Image Processing," Prentice Hall, 1989.

[6] Y. Li, H. Kobatake, "Extraction of Facial Sketch Images and Expression Transformation Based on Faces,", *IEEE Conference on Image Processing*, vol. 3, pp. 520-523, Washington, USA, October 1995.

[7] H. Niemann, "Pattern Analysis and Understanding", Springer-Verlag, 1990.

[8] I. Pitas, A. N. Venetsanopoulos, "Nonlinear Digital Filters: Principles and Applications", Kluwer Academic Publishers, 1990.

[9] S. Tominaga, "Color Image Segmentation Using Three Perceptual Attributes," *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 628-630, 1986.

[10] L. Vincent, P. Soille, "Watersheds in Digital Space: An Efficient Algorithm Based on Immersion Simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583-598, June 1991.

[11] H. Wu, Q. Chen, M. Yachida, "An Application of Fuzzy Theory: Face Detection," *International Workshop on Automatic Face- and Gesture Recognition*, pp. 314-319, ed. Martin Bichsel, Zurich, Switzerland, June 1995.