METHOD FOR LOCALIZATION OF HUMAN FACES IN COLOR-BASED FACE DETECTORS AND TRACKERS

Vladimir Vezhnevets

Department of Computational Mathematics & Cybernetics.
Moscow State University, Vorobjovy Gory, Moscow, 119899, Russia
vvp@graphics.cs.msu.su

Abstract. This paper is devoted to problem of face localization in color images, more particularly, to the task of grouping pixels with high skin probability to form possible face candidate regions. Using connected components analysis for pixels grouping has a significant drawback, for the noise of skin color detection and non-face skin-colored objects can form connected regions of skin pixels with vividly not face-like shape. Two methods for extracting ellipse-shaped skin regions are described, both capable of successfully finding faces on images with considerable noise and in presence of other skin-colored objects.

Keywords: face detection, face tracking, color-based segmentation.

Introduction

Locating of human face in an image is a necessary background for automatic face recognition, facial expression analysis and other high-level tasks of human face perception. That is why automatic detection and localization of human faces in two-dimensional natural, complex scene images and tracking along image sequences receive much attention nowadays.

Color is a distinctive feature of a human face. Using color as a feature for face detection and tracking has several advantages: first, color is nearly invariant to orientation of the surface and invariant to lighting to some extent, second, processing of color information is much faster than processing other facial features. Most of color-based face detection and tracking algorithms use two steps of processing. First, skin color detector creates an image, which indicates the likelihood of each pixel to represent skin (called skin probability or skin likelihood image). Second, pixels with high skin likelihood are grouped into face candidate regions, which are than tested to satisfy definite criteria (usually some shape restrictions). The straightforward way of grouping skin pixels to form face candidate regions is connected components analysis [1, 2, 3, 4]. This method has a significant drawback, for the noise of skin color detection and non-face skin-colored objects can form connected regions of skin color, that fail to pass the restrictions applied to face candidates (see figure 1).

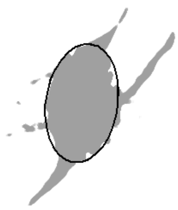
Skin detection examples
(darker pixels represent skin color)
Fig. 1

This problem has inspired design of an ellipse-specific pixel-grouping method, which can successfully cope with noisy skin detection results and images with multitude of skin-colored objects. This method is capable of finding groups of skin-colored pixels with approximately elliptic shape, not necessarily well separated from other skin-colored pixels. Also, it allows different restrictions to be specified on the possible shape of the extracted ellipses (size, orientation, relationship between axes). The following sections describe the method in detail and give the results of application to different images.

1. Ellipse-specific face detection

The task statement is: find and mark groups of skin pixels forming "almost elliptic" shape. Often the results of skin color detection exhibit noise, or many skin colored objects create considerable clutter in skin probability image and make face region not so clearly distinguishable (see figure 1).

The idea is to fit an ellipse to the image region in the way, that the ellipse covers as many points of high skin likelihood as possible, while retaining little skin-colored pixels at its border and little non-skin colored pixels inside. This means that the most desirable pixels pattern is an ellipse with high skin likelihood inside, and very low likelihood at the border. Figure 2 shows an example of desired detection result on a noisy image.



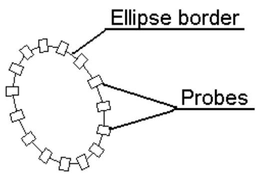
Extraction of elliptic region from a noisy image
Fig. 2

Two methods of achievement of the stated goal are presented here; both are using successive approximations to find the ellipse best fitted to skin pixels patterns.

1.1. First method of ellipse detection

First method is similar to "radial spanning", described in [6] and to snake used for face tracking in [2]. Both mentioned methods are used to find approximate size and shape of a single cluster of skin-colored pixels. Neither of them takes the advantage of expected elliptic shape of the face, allowing almost arbitrary deformations of the extracted cluster, which may lead to inaccurate detection of the face region. Also, it is not easy to incorporate face region shape and position restrictions in those methods.

The suggested method uses an elliptic model, shown here in the Fig. 4. It is initialized at some location in the image near the expected position of face. Then it deforms itself step-by-step to find the most desirable position and configuration. A number of rectangular probes, positioned on the ellipse border, evaluate the distribution of skin-colored pixels in the border area of the ellipse and decide, whether each of them should move outwards or inwards.



Elliptic model
Fig. 4

The desirable displacement for each probe is calculated using the following rules:

$$P_{in} = 2 \cdot \frac{N_{inside}}{S}; P_{out} = 2 \cdot \frac{N_{outside}}{S};$$

N_{inside} - number of skin pixels that lie both inside probe and ellipse;

$N_{outside}$ - number of skin pixels that lie inside probe and outside ellipse;

S - probe area;

If $P_{in} < T_{in}$, where T_{in} is a threshold value, the probe moves inwards with speed V_{in}. Else, if $P_{out} > T_{out}$, where T_{out} is a threshold value, the probe moves outwards with speed V_{out}.

One step of model updating includes:

1. Calculation of the displacement of all the probes;
2. Fitting an ellipse to the centers of the repositioned probes.

Ellipse fitting can be performed, for example, by the procedure of fast ellipse fitting described in [5] and implemented in Intel OpenCV library [7].

This simple scheme shows good robustness and quickly finds the ellipse containing face even when it is initialized not inside the face region. Fig. 5 shows some results of face region detection.

Example of face region detection, showing
ellipse initialization, and the final configuration
Figure 5

The model fitting procedure behavior is controlled by these parameters:

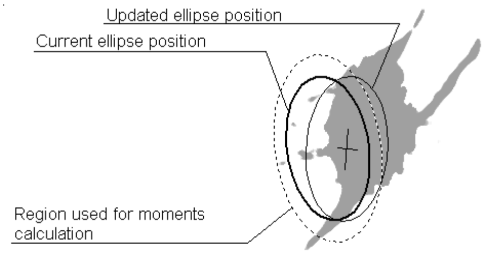
1. The number and sizes of the probe rectangles;
2. The positioning of the probes (the probes can be fixed, or they may be shifted on each step along the ellipse border, to avoid lock-ups in definite pixel configurations);
3. The threshold values controlling the movement of the probes (T_{in}, T_{out});
4. The speed of the probes moving outwards and inwards (V_{in}, V_{out}).

In case there are know n some restrictions on the possible orientation, size or proportions of face to be detected in the image, such conditions can be easily incorporated into ellipse fitting procedure, increasing robustness of the whole method.

1.2. Second method of ellipse detection

Second method of extraction ellipse -like groups of skin colored pixels differs from the already described one by considering the skin pixels distribution not only on the border of the ellipse, but also inside it. This method is particularly well suited to the task of tracking a skin -colored ellipse (face) along the sequence of frames, when the proportions of the ellipse are fixed and the ellipse size does not change dramatically between successive frames.

Elliptic model is initialized near the expected position of the face and then is adapted step -by-step to fit the image data. The step of the adaptation process consists of the considering the skin pixels lying inside the ellipse of a slightly larger size, and calculating the centroid and second order moments of the pixels formation inside this larger ellipse (see Fig. 6).



Second method of elliptic model updating
Fig. 6

Updated ellipse position and orientation is evaluated using the calculated moments:

$$\mu_x = \frac{\sum_{(x,y)\in S} x \cdot I(x,y)}{\sum_{(x,y)\in S} I(x,y)}; \qquad \mu_y = \frac{\sum_{(x,y)\in S} y \cdot I(x,y)}{\sum_{(x,y)\in S} I(x,y)};$$

$$\mu_{02} = \frac{\sum_{(x,y)\in S} (y-\mu_y)^2 \cdot I(x,y)}{\sum_{(x,y)\in S} I(x,y)}; \qquad \mu_{20} = \frac{\sum_{(x,y)\in S} (x-\mu_x)^2 \cdot I(x,y)}{\sum_{(x,y)\in S} I(x,y)};$$

$$\mu_{11} = \frac{\sum_{(x,y)\in S} (x-\mu_x)(y-\mu_y) \cdot I(x,y)}{\sum_{(x,y)\in S} I(x,y)};$$

Where:

S - ellipse with slightly larger size, than the elliptic model;
$I(x,y)$ - pixel intensity at (x, y) position of likelihood image;

New ellipse center and axis are computed as follows: (μ_x, μ_y) - is new ellipse center p oint, $(\mu_{11}, \mu_{02} - \mu_{20} + \sqrt{(\mu_{02} - \mu_{20})^2 + 4\mu_{11}^{\ 2}})$ - unnormalized ellipse major axis vector. Ellipse size can be recalculated by evaluating new axes length (for example see [2]) and determining new ellipse size with fixed proportions that suits best to given axes. Estimation of t he face region with the method described here works well even on images with deformed skin pixels region (see Fig. 7).

Examples of face localization using second method
Fig. 7

Conclusion

Two methods for detection of ellipse-shaped face candidate regions in the skin likelihood images were described in this paper. Both are based on a simple rule - finding ellipse configuration, which maximizes number of skin-colored pixels inside and minimizes number of skin colored-pixels at the border. Both methods show robust behavior on skin probability images with much noise and many skin-colored objects and can be successfully used in color-based face detection and tracking systems. Of course, erroneous detection occurs sometimes, but the fact is that usually there exist very few stable configurations for the elliptic model in the image. So by varying the initialization position we can get a small number of possible face candidates. The described methods augmented with some additional checks (like, for example, presence of the facial features inside the ellipse) can easily reject the false candidates.

References

1. S. Tsekeridou and I. Pitas: “Facial features extraction in frontal views using biometric analogies”, in *Proceedings IX European Signal Processing Conference*, 1998, vol. 1

2. K. Sobottka and I. Pitas: “Segmentation and Tracking of Faces in Color Images”, *Second International Conference on Automatic Face and Gesture Recognition 1996*, Killington, Vermont, USA, pp. 236-241, 14-16 October 1996.

3. J.-C. Terrillon, M. David, S. Akamatsu: "Automatic Detection of Human Faces in Natural Scene Images by Use of a Skin Color Model and of Invariant Moments", In *Proceedings of the Third International Conference on Automatic Face and Gesture Recognition,* Nara, Japan, 1998, pp. 112-117, 1998

4. Bernd Menser and Mathias Wien: “Segmentation and tracking of facial regions in color image sequences. In *Proc. SPIE Visual Communications and Image Processing 2000, volume 4067, pages 731-740*, Perth, Australia, June 2000

5. Andrew Fitzgibbon, Maurizio Pilu, and Robert B. Fisher: Direct Least Square Fitting of Ellipses, in *IEEE Transactions on Pattern Analysis And Machine Intelligence, Vol. 21, No. 5, May 1999, pp.476-480*.

6. K. Toyama: Radial spanning for fast blob detection. In *Proc. Int’l. Conference on Computer Vision, Patt. Recognition, and Image Processing*, 1998.

7. Bradski, G: Open Source Computer Vision Library, 2001 *http://www.intel.com/research/mrl/research/opencv.*