MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11-13, 2002, Nara-ken New Public Hall, Nara, Japan

8—18

Face Detection Using Scan-Line Based Hough Transform and MLP

Nam-Kie Kim, Il Choi and Sung-Il Chien[1]
School of Electronic and Electrical Engineering, Kyungpook National University

Abstract

A face boundary can be approximated by an ellipse with five-dimensional parameters. This property allows an ellipse detection algorithm to be adapted to detecting faces. However, the construction of a huge five-dimensional parameter space for a Hough transform is quite unpractical. Accordingly, we propose a scan-line-based Hough transform (SLHT) method for detecting faces from a symmetric contour in an image. The idea is based on the use of a constant aspect ratio for a face, gradient information, and scan-line-based orientation decomposition, thereby allowing a 5-dimensional problem to be decomposed into a two-dimensional one to compute a center with a specific orientation and a one-dimensional one to estimate a short axis. In addition, a two-point selection constraint using geometric and gradient information is also employed to increase the speed and cope with a cluttered background. After detecting candidate face regions using an SLHT, a multi-layer perceptron verifier is adopted to reject false positives. The proposed method was found to be relatively fast and promising.

1 Introduction

Detecting human faces from background in a scene is a very important task that constitutes the first step for a large number of applications: face recognition, face tracking and surveillance, remote conferencing, etc. [1,2]. Although significant progress has been made in the last two decades, there is still much to be done, and a robust face detection system is needed that can fully cover variations in scale, orientation (upright, rotated), pose (frontal, profile), occlusion, cluttered background, and lighting conditions. Moreover, when detection methods are used in a real-time system, it is important to consider the requirements of time and memory. As such, accuracy may need to be sacrificed for speed. Recently, several methods [3-6] based on elliptical approximation have been proposed using the fact that the overall shape of a head is very similar to an ellipse. However, these methods mainly focus on the localization of a single upright face in a simple background. Moreover, these methods often detect many false positive faces that might have the shapes of ellipse. Thus, these methods could have difficulties of detecting faces correctly in images with a cluttered background, and various face orientations and lighting conditions.

The proposed method detects faces in an image using a scan-line-based Hough-like ellipse fitting method. A Hough transform has long been recognized as a robust technique for detecting line features, yet applying it to detect ellipses requires a 5-dimensional array for the accumulator. Moreover, a great deal of computing time is needed to transform a feature point in the input image into many points in the parameter space. Therefore, to solve these problems, we use the fact that the ratio of the long to the short axis of a face is constant, along with gradient information extracted using a Sobel operation [7]. In addition, the parameter space is effectively reduced using scan-line-based orientation decomposition. Finally, since the gradient direction satisfies reflectional symmetry on a face contour on a scan-line, the computing time for detecting a face can be reduced.

The remainder of this paper is organized as follows: Section 2 gives a detailed description of the proposed face detection method using SLHT, section 3 presents experimental results using two database sets, and section 4 provides a summary and conclusions.

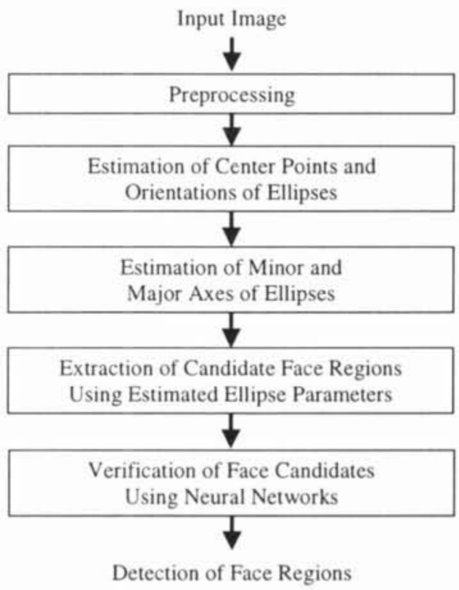

Fig. 1. Basic flowchart of proposed face detection method.

2 Face Detection using Hough-like Ellipse Fitting

The overall shape of the frontal view of a face (or the whole head) is very similar to an ellipse, not only in relation to facial expressions or the presence or absence of structural components, such as a beard, mustache, and glasses, but also with variations in the lighting conditions, orientation, or size of the face. Based on this fact, a face region can be detected using a Hough-like ellipse detection algorithm.

[1] Address: 1370, Sangyeok-dong, Buk-gu, Daegu 702-701, Korea. E-mail: **sichien@ee.knu.ac.kr**

The proposed method for detecting faces in an image, as shown in Fig. 1, is as follows. First, gradient magnitudes and directions are extracted from an input image using Gaussian smoothing and Sobel operators in a preprocessing step. Second, a search for ellipse regions is conducted by the SLHT technique. Finally, the determination of whether or not such regions are facial regions is performed by an MLP verifier.

2.1 Preprocessing

The proposed face detection method requires an edge map that includes the gradient magnitudes and orientations of the input image as a Hough transform is used to find any ellipses. To extract the edge map, Gaussian smoothing is applied to conspicuous edges of a face contour and to reduce the influence of dense edges represented in non-interest regions, such as books on a bookshelf, or clothes with a check pattern, before using a Sobel operator.

2.2 Ellipse detection using SLHT

This section describes the proposed ellipse detection method for detecting candidate face regions. A face boundary can be approximated by an ellipse represented by five parameters, given by

$$\frac{((x-x_c)\cos\alpha+(y-y_c)\sin\alpha)^2}{a^2}+\frac{(-(x-x_c)\sin\alpha+(y-y_c)\cos\alpha)^2}{b^2}=1 \quad (1)$$

where a and b are the short and long axis, respectively, (x_c, y_c) is the center point, and α is the orientation of the long axis. However, the construction of a huge five-dimensional parameter space for directly extending a Hough transform to determine a solution is very difficult [8]. Therefore, the proposed strategies to reduce the complexity of the parameter space are as follows. First, the ratio $k=b/a$ is made into a fixed constant, as the aspect ratios of faces are approximately equal to each other in an image. Second, we use the fact that the slope of a point can be calculated using the corresponding gradient direction extracted from a Sobel operation. Third, a scan-line decomposition method is used to determine the orientation of an ellipse. These distinctive strategies can decompose the 5-dimensional problem into a two-dimensional one for computing the center with a specific orientation and a one-dimensional one for estimating the short axis. As such, the ellipse detection can be divided into two stages as follows.

In the first stage, the centers and orientations are estimated using the above strategy. The equation (1) for its derivative is given by

$$(k^2A\cos\alpha-B\sin\alpha)(x-x_c)+(k^2A\sin\alpha+B\cos\alpha)(y-y_c)=0 \quad (2)$$

where $A=\cos\alpha+m\sin\alpha$ and $B=-\sin\alpha+m\cos\alpha$. Plus, $m\ (=dy/dx)$ is the slope of a point on an ellipse and can be computed using the gradient direction (φ) at the point using $m=-\cot\varphi$ [9]. Equation (2) has three parameters, the orientation (α) and center point (x_c, y_c) of an ellipse, because k and m are known.

To solve the orientation problem, a scan-line-based orientation decomposition method for computing a Hough transform is introduced to achieve a considerable reduction in the computational load, as shown in Fig. 2. Here, a Hough transform is divided into several steps, each corresponding to a selectively attentional decomposition with a specific scan direction. An SLHT uses a set of lines that have the same slope to scan an image line by line in an oblique raster scan fashion. Decomposed Hough maps are constructed for each scan direction. The direction of a scan-line θ_l is defined as $\theta_l=l\pi/L-\pi/2$ where $l=0,1,\cdots,L-1$ and L represents the maximum number of attentional decompositions. In this case, the orientation of an ellipse (α_l) can be easily estimated as it is perpendicular to the scan direction (θ_l).

To calculate the center point of an ellipse, equation (2) is rewritten as follows

$$Wx_c+Zy_c=K \quad (3)$$

where $W=k^2A\cos\alpha-B\sin\alpha$, $Z=k^2A\sin\alpha+B\cos\alpha$, and $K=Wx+Zy$.

Using two points $p_i(x_i, y_i)$ and $p_j(x_j, y_j)$ on a scan-line, as shown in Fig. 3, the corresponding φ_i and φ_j, and estimated orientation $\hat{\alpha}_l$, the center point on a Hough map can be estimated from equation (3)

$$\begin{bmatrix}\hat{x}_c(\hat{\alpha}_l)\\ \hat{y}_c(\hat{\alpha}_l)\end{bmatrix}=\frac{1}{D}\begin{bmatrix}Z_jK_i-Z_iK_j\\ -W_jK_i+W_iK_j\end{bmatrix},\ i\neq j \quad (4)$$

where $D=W_iZ_j-W_jZ_i\neq 0$, $(\cdot)_i$ and $(\cdot)_j$ represent the value of $(\cdot)$ at p_i and p_j, respectively. Here, the condition $D\neq 0$ produces a constraint given by

$$|\varphi_i-\varphi_j|\neq 0,\ \pi\ . \quad (5)$$

As such, only two points need to be considered for the Hough map and straight lines arising from clutter in an image can be remove. To further reduce the Hough calculation, the above constraint is combined with a reflectional symmetry condition as follows

$$|\varphi_i+\varphi_j-2\theta_l|=\pi \quad (6)$$

and a distance condition that considers the minimum and maximum width of a face. These constraints are called the two-point selection constraint. Thus, for each scan line with a specific direction, two points that satisfy the combined constraints are selected. Then a decomposed Hough map with a specific direction is constructed using equation (4). The center points of ellipses with an orientation can be easily detected using local peak detection in the map. An SLHT can only detect an ellipse with a specific orientation in an image by selecting a specific value or group of specific values l. This is a feature of selective attentional decomposition.

In the second stage, after estimating the center and orientation of an ellipse, the remaining parameter, the short axis, is calculated using equation (1) as follows

$$\hat{a}=\sqrt{((x-\hat{x}_c)\cos\hat{\alpha}_l+(y-\hat{y}_c)\sin\hat{\alpha}_l)^2+\frac{(-(x-\hat{x}_c)\sin\hat{\alpha}_l+(y-\hat{y}_c)\cos\hat{\alpha}_l)^2}{k^2}} \tag{7}$$

and the long axis is computed using the ratio k. Thus, a candidate face region modeled by an ellipse can be detected efficiently.

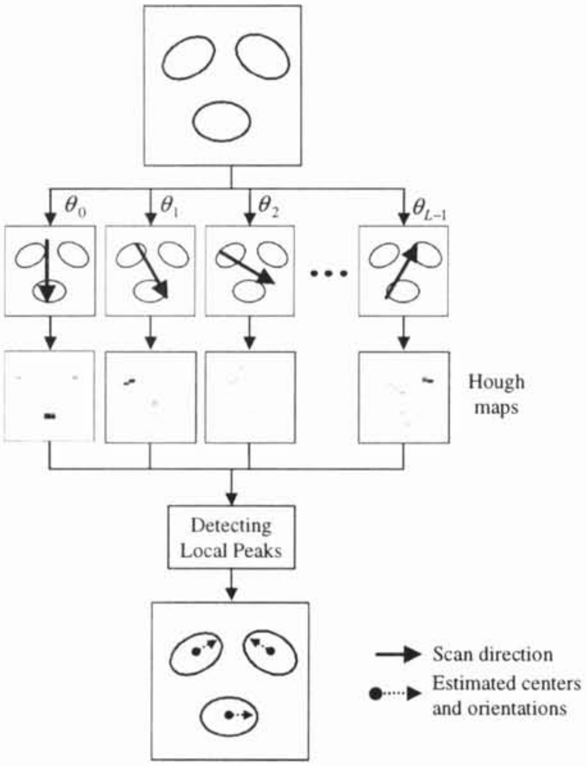

Fig. 2. Conceptual diagram of proposed scan-line-based Hough transform.

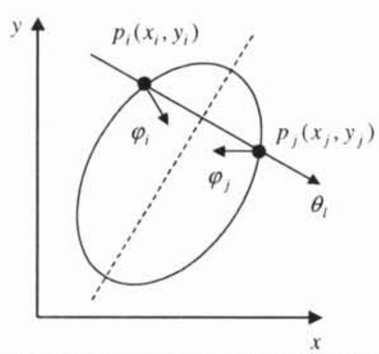

Fig. 3. Selection of two points satisfying reflection symmetry on scan line with specific orientation θ_l.

2.3 Face verifier using MLP

An MLP verifier is adopted for rejecting false positives. First, the size of a candidate face region is normalized to 12×18 pixels. To cope with variations in the illumination conditions, the region is histogram-equalized, then input into an MLP neural network trained by backpropagation. The classifier consists of 216 input layer nodes, 10 hidden layer nodes, and 2 output layer nodes. In section 2.2, when non-face images are gathered that look like ellipses, these negative examples incorrectly identified as a face are used in retraining the neural network to further enhance the MLP verifier, sometimes called a bootstrap method.

3 Experimental Results

The proposed method was evaluated with various images, including the MIT face database and CMU face database. Currently, the proposed method was evaluated using two sets of images. Set A contained 243 gray-level images (total 315 faces) acquired from a CCD camera with 320×240 pixels, while set B contained 64 images (total 137 faces) obtained from the CMU face database, TV scenes, magazines, and the Internet.

Figure 4 shows the edges of an image extracted using a Sobel operation and the edges of an image extracted based on satisfying the two-point selection constraint. Up to 595,778 edge-pairs could be selected on the scan-line, as shown in Fig. 4 (a). Whereas, the number of edge pairs satisfying the two-point selection constraint was no more than 4000. Practically speaking, the number of edge pairs was reduced by more than 99% when experimenting on set A. As such, the two-point selection constraint contributed to a decrease in the computing time without any drop in efficiency. Generally, the edges represented on a face contour were not excluded by the two-point selection constraint, as shown in Fig. 4 (b).

Some test images with detected faces are shown in Fig. 5. The proposed method was able to detect faces with glasses, rotated faces, and colored faces with complex and cluttered backgrounds and lighting variations. The faces missed did not have an ellipse shape as the head was profiled (third row, second column) or the background color was similar to that of the face (third row, third column). The average computing time for an image with 320×240 pixels with multiple faces was less than one second with a Pentium III (500MHz). After verification, the face detection results are given in Table 1.

Table 1. Face detection results.

	Missed face rate	Detection rate	False positive rate
Set A	31/315	90.2%	18/315
Set B	25/137	81.8%	13/137
Total	56/452	87.6%	31/452

(a) (b)

Fig. 4. (a) Edge image. (b) Extracted edge points satisfying two-point selection constraint.

4 Summary and Conclusions

We proposed a novel face detection algorithm using a SLHT method for detecting faces from a symmetric contour in an image. The idea is based on the use of a constant aspect ratio for a face, gradient information, and scan-line-based orientation decomposition to decompose a 5-dimensional problem into a two-dimensional one to compute a center with a specific orientation and a one-dimensional one to estimate the short axis. In addition, a two- point selection constraint using geometric and gradient information is also added to the above idea to increase the speed and cope with a cluttered background. After detecting candidate face regions, an MLP verifier is adopted to reject false positives.

Conventional face detection methods based on elliptical approximation mainly focus on the localization of a single upright face in a simple background. However, these methods often detect many false positive faces due to the shapes of the ellipses, cluttered backgrounds, face orientations, and lighting variations. The proposed method can provide a solution to cope with these problems based on adopting a two-point selection constraint for an SLHT and an MLP verifier trained in a bootstrap manner. One of the most important features of an SLHT is that the combination of directional decompositions means that only ellipses with specific orientations are responded to in an image.

References

[1] M.H. Yang, D.J. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34-58, 2002.

[2] H.A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 23-38, 1998.

[3] Y.H. Kwon and N.V. Lobo, "Face detection using templates," *Proc. of the 12th IAPR*, vol. 1, pp. 764-767, 1994.

[4] Y. Yokoo and M. Hagiwara, "Human faces detection method using genetic algorithm," *Proc. of IEEE Conf. Evolutionary Computation*, pp. 113-118, 1996.

[5] J. Wang and T. Tan, "A new face detection method based on shape information," *Pattern Recognition Letters*, vol. 21, pp. 463-471, 2000.

[6] D. Maio and D. Maltoni, "Real-time face location on gray-scale static images," *Pattern Recognition*, vol. 33, pp. 1525-1539, 2000.

[7] D.H. Ballard, "Generalizing the hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, no. 2, pp. 111-122, 1981.

[8] Y. Lei and K. Wong, "Ellipse detection based on symmetry," *Pattern Recognition Letters*, vol. 20, pp. 41-47, 1999.

[9] C. Kimme, D. Ballard, and J. Sklansky, "Finding circles by an array of accumulators," *Commun. ACM*, vol. 18, no. 2, pp. 120-122, 1975.

Fig. 5. Sampled results of face detection for various images.