Face Detection by Orientation Map Matching

Li Bai and LinLin Shen
School of Computer Science & IT, University of Nottingham, Nottingham, UK
E-mail: {bai, lls}@cs.nott.ac.uk

Abstract

This paper describes a fast and accurate face detection method based on edge orientation map matching of colour images. Face detection speed and accuracy is crucial to realtime face recognition applications but still remains a problem even though much research effort has been devoted to face detection algorithms. What make our method different from existing works are that only simple computation is involved so that faces can be detected in less than 1 second, and both face shape and skin colour information is taken into account whereby false detection is reduced. In addition our algorithm is robust to noise as edge orientation maps are computed for each image block, rather than for each individual pixel. Experimental results demonstrate the advantages of our approach.

1 Introduction

The face detection problem has been addressed in the literature over the last ten years or so and is still an unsolved problem in terms of speed and accuracy. A typical face detection system breaks down an image into small subwindows and uses a trained classifier to determine whether a subwindow is a face image. It is difficult to know prior the size of the subwindows as not all the faces in an image are of the same size. For this reason face detection process has to be repeated for several resolution levels. A typical such system is Rowley's neural network-based face detector [1]. The detector consists of a set of neural networks each applied to the subwindows of an image. Inputs to the neural networks are pixel values of 20x20 subwindows, and outputs from these networks are put through an arbitrating process to arrive at the final decision as to whether a subwindow is a face image. Feraud's face detection methods [2] are also based on neural networks and the size of the subwindows is 15x20.

Many face detection works rely on colour models [3][4][5]. Image pixels can be labeled as skin or non-skin pixels according to the values of the pixels with respect to the colour model. Though distribution of colours provides a valuable cue for the presence of faces in images, it should not be used as the sole method of locating faces. This is not only because people have different skin colours but also because skin colour varies under different lighting conditions and there are often other skin regions other than faces in the image such as neck and hands. In [4], colour information is first used to segment the skin region and ellipse fitting is then applied to identify face regions. However, it is not always appropriate to describe the shape of a face as an ellipse due to different poses.

Other popular approaches are based on template and facial feature matching. The essence of feature-based approaches to face detection is to find the arrangement of certain features such as the eyes, nose, and the mouth in the image, which forms a face pattern. A number of papers describe some promising applications of wavelet analysis to extract facial features [7]. In template matching a face model is defined and used to look for characteristic patterns in an image associated with the face model. Fröba and Küblbeck [6] extract edge orientation information from a face model and match this against a 32x40 subwindow at every pixel location in the image. To find faces larger or smaller than the face model, a pyramid of edge orientation fields is built by varying the resolution of the input image. Since this method uses only edge orientation information, false detection often occurs where edge frequency or texture is high. It is observed during our experiments that this algorithm frequently classifies regions with leaves and grass as faces.

This paper describes a fast face detection algorithm that combines skin colour information and face shape information in the form of face orientation maps. A colour image is converted to a skin probability image first based on a Gaussian skin colour model, from which face orientation map is extracted. Details of the two steps are described in section 2 and 3 respectively. Finally the orientation map is matched against a pre-defined model as described in section 4. A database consisted of colour images downloaded from the Internet is built to evaluate the performance of the algorithm. Though the images in the database contain complex background and the faces are in different orientations, our algorithm is fast and has produced accurate results.

2 Colour Analysis

It is widely accepted that the colour of human skin is distinctive from the colour of many other natural objects in an image. There are different ways of representing colours (colour models) and some of them are less sensitive to variations in luminance, such as the YCbCr colour model which separates luminance (Y) from chrominance (CbCr) or lightness is separated from colour. By analyzing the skin colour statistics, it is concluded that skin colours are distributed over a small area in the chrominance plane and the major difference between skin tones is intensity [4]. So to make the best use of colour features, an image is first converted into luminance and chrominance channels in the YCbCr colour space. This is done as follows.

Let $\mathbf{w}_{ij} = [Cb_{ij}\ Cr_{ij}]^T$ denote a vector composed of the chrominance components Cb and Cr for a pixel (i, j). The class-conditional pdf of $\mathbf{w}_{ij}$ belonging to the skin class x is modeled by a two-dimensional Gaussian [3],

$$p(\mathbf{w}_{ij} \mid x) = (2\pi)^{-1} \mid \Sigma \mid^{-1/2} \exp\left(-\frac{1}{2}\left[\mathbf{w}_{ij} - \boldsymbol{\mu}\right]^T \Sigma^{-1}\left[\mathbf{w}_{ij} - \boldsymbol{\mu}\right]\right) \qquad (1)$$

where the mean vector $\boldsymbol{\mu}$ and the covariance matrix Σ are estimated from the training set. The contour of the pdf shown above defines an ellipse in the CbCr domain, whose center and principal axis is determined by $\boldsymbol{\mu}$ and Σ respectively. Figure 1 shows the distribution of skin colours in the CbCr domain.

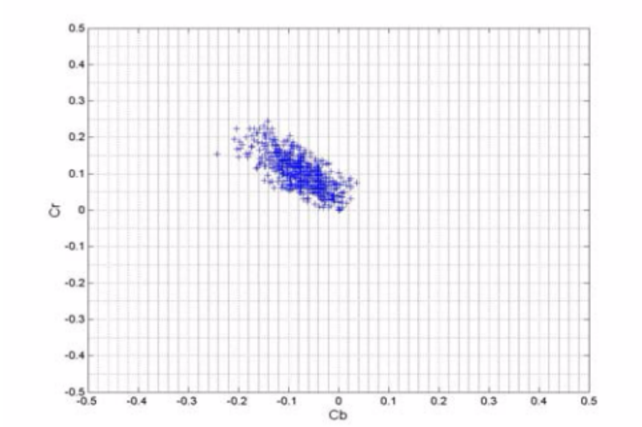
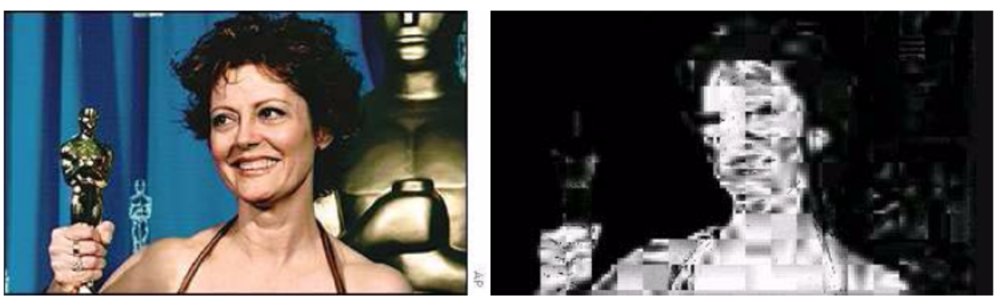


Figure 1 Distribution of skin colours in the CbCr domain

The skin probability image P is then created and used to extract face orientation information. Each pixel in P indicates the probability of an image pixel belong to the skin class x:

$$P(i, j) \sim p(\mathbf{w}_{ij} \mid x) \tag{2}$$

Figure 2 (a) and (b) shows an input colour image I and its corresponding skin probability image P. The example indicates that skin probability images isolate skin regions and suppress the backgrounds. This reduces the search space for the presence of faces.



(a) (b)

Figure 2 An input colour image (a) and its skin probability image (b)

3 Orientation Map Extraction

Since pixel-based edge orientation is susceptible to noise, a block-based orientation map extraction approach [8] is adopted in our algorithm. Two 3×3 Sobel operators, S_x for horizontal filtering and S_y for vertical filtering, were convolved with the image $I(x, y)$ to generate two gradient images $G_x(x, y)$ and $G_y(x, y)$.

$$G_x(x, y) = S_x * I(x, y) \tag{3}$$
$$G_y(x, y) = S_y * I(x, y) \tag{4}$$

Similar to the algorithm used in [8], the gradient images $G_x(x, y)$ and $G_y(x, y)$ are divided into a series of non-overlapping windows of size *w×w*, and each pixel *(x, y)* in the window centered at pixel *(i, j)* is assigned the same orientation value *O(x, y)* as shown below:

$$V_y(x,y) = \sum_{u=i-\frac{w}{2}}^{u=i+\frac{w}{2}} \sum_{v=j-\frac{w}{2}}^{v=j+\frac{w}{2}} 2G_x(u,v)G_y(u,v)$$

$$V_x(x,y) = \sum_{u=i-\frac{w}{2}}^{u=i+\frac{w}{2}} \sum_{v=j-\frac{w}{2}}^{v=j+\frac{w}{2}} (G_x^2(u,v) - G_y^2(u,v)) \qquad (5)$$

$$O(x,y) = \frac{1}{2}\tan^{-1}\left(\frac{V_y}{V_x}\right)$$

After the orientation field of an input image is calculated, the certainty level of edge orientation at pixel *(x, y)* in the window centered at pixel *(i, j)* is defined as below: [8]

$$C(x,y) = \sqrt{\frac{1}{w \times w} \frac{(V_x^2(x,y) + V_y^2(x,y))}{V_e(x,y)}} \qquad (6)$$

where

$$V_e(x,y) = \sum_{u=i-\frac{w}{2}}^{u=i+\frac{w}{2}} \sum_{v=j-\frac{w}{2}}^{v=j+\frac{w}{2}} (G_x^2(u,v) + G_y^2(u,v)) \qquad (7)$$

A *3 x 3* box filter (or an averaging mask) is used to smooth the orientations obtained previously so as to remove any abrupt changes in orientation that are caused by noise in the low quality image regions. The edge information on homogenous parts of the image where no grey value changes occur is often noisy and provides no useful information for the detection [6]. A threshold T_c is applied to the certainty level *C(x, y)* to generate an edge certainty level field $C_t(x,y)$.

$$C_t(x,y) = \begin{cases} C(x,y) & \text{if} \quad C(x,y) > T_s \\ 0 & \text{else} \end{cases} \qquad (8)$$

The edge orientation information can then be rewritten as a vector field as shown below:

$$\mathbf{V}(x,y) = C_t(x,y)e^{jO(x,y)} \qquad (9)$$

Figure 3 (a) and (b) show the orientation maps extracted from Figure 2 (a) and (b) respectively. Compared with that in (a), background edges have been reduced in (b) and thus the chances of false detection are reduced. The orientation map matching based on skin probability image produces better results than could have been on original images.

(a) (b)

Figure 3 Orientation maps of the original image (a) and of skin probability image (b)

4 Orientation Map Matching

A face model is built from a sample of hand-labeled face images. Ten color face images are cropped, aligned and scaled to the size 33×42. Ten orientation maps were extracted from the face images and a model orientation map was calculated by averaging the ten orientation maps. For face detection, the orientation model $\mathbf{V}_m(x,y)$ is slid over the input image and the similarity between the model and the underlying image block centered at pixel *(i,j)* is calculated as below:

$$S(i,j) = \sum_{u=i-\frac{w}{2}}^{u=i+\frac{w}{2}} \sum_{v=j-\frac{h}{2}}^{v=j+\frac{h}{2}} sim(V_m(u,v), V_I(i+u, j+v)) \tag{10}$$

where $\mathbf{V}_I(x,y)$ is the orientation map of the input image, w is the width of the model orientation map, h is the height of the model orientation map and

$$sim(\mathbf{V}_1, \mathbf{V}_2) = \begin{cases} \cos(|\arg(\mathbf{V}_1) - \arg(\mathbf{V}_2)|) & if \quad |\mathbf{V}_1|, |\mathbf{V}_1| > 0 \\ 0 & else \end{cases} \tag{11}$$

A pyramid of orientation maps at different resolutions is used to detect faces of different sizes. The size ratio between two resolution levels is set to be 1.25. At each resolution level, the input image is matched against the model template and a similarity map is generated according to equation (10) and (11). From these similarity maps, the pixel with the maximum similarity score is found and a face is detected at the location of this pixel.

5 Experimental Results

An image database consisting of 240 colour images downloaded from the Web is used in the experiments to test the algorithm. All images contain faces in complex backgrounds and are in different orientations and sizes, and the largest image is of size 360×420. A face is confirmed if the position and the size of a detected region is reasonable to contain a face. Since the model template used in our experiments is of size 33×42, it produces 154 orientation vectors when a window of size 3×3 is used to calculate the orientation maps. The average detection time is less than 0.6 seconds on a Pentium IIII, 1800 MHz. Figure 4 shows the detection results for several images from our database.

Figure 4 Face detection results

6 Conclusion

In this paper, we have described an orientation map based face detection algorithm. A skin probability image is extracted first from a colour image and edge orientation map is matched against that of a face model template. Experiments have shown that the proposed algorithm is both fast and accurate. We also compared our algorithm with those of Fröba and Küblbeck [6]. They extract edge orientation maps from gray level images and match them against the orientation map of a model template. When we tested their algorithm against our image database, false detection often occurs in the image where edge frequency and texture is high. They also extract edge orientation maps using two 3×3 Sobel operators for each pixel and this has shown to be susceptible to image noise. Instead of using the gray level images we use the skin probability images and we compute edge orientations for image blocks rather than for pixels. As is the case with most of skin colour based detection methods, strong lighting conditions may decrease the algorithm performance as some background pixels might be classified as skin pixels. Other factors that may affect the results are image block size, threshold chosen, the way to scale orientation maps, and the efficiency of the matching algorithm. We are currently investigating these issues to improve the algorithm.

References

[1] Rowley, Shumeet Baluja, and Takeo Kanade, Neural Network-Based Face Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 1, January 1998.

[2] Rapha el Feraud, Olivier J. Bernier, Jean-Emmanuel Viallet, and Michel Collobert, A Fast and Accurate Face Detector Based on Neural Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 1, January 2001.

[3] BMenser and F.Muller, Face Detection in Colour Images Using Principal Components Analysis. *Proceedings Seventh International Conference on Image Processing and its Applications*, vol. 2, pp.620–624 , July 1999.

[4] Eli Saber, a.Murat Tekalp, Frontal-view Face Detection and Facial Feature Extraction Using Colour, Shape and Symmetry Based Cost Function, *Pattern Recognition Letters*, 19(8):669-680, June, 1998.

[5]] Saber, E. Takalp, A.M., Eschbach, etc., Automatic Image Annotation Using Adaptive Colour Classification. *Graphical Models and Image Process*. 58, 115-126, 1996.

[6] Bernhard Froba and Christian Kublbeck. Real-Time Face Detection Using Edge-Orientation Matching. *3rd International Conference on Audio and Video Based Biometric Person Authentication*, p78-83, Sweden, June, 2001.

[7] F.Smeraldi, O.Carmona, and J.Bigun. Saccadic Search with Gabor Features Applied to Eye Detection and Real-time Head Tracking. *Image and Vision Computing*, 18:323-329, 2000.

[8] Lin Hong, Anil Jain, Sharath Pankanti, Ruud Bolle, An Identity Authentication System Using Fingerprints. *Proceedings of the IEEE*, Vol. 85, No. 9, Sept. 1997, pp. 1365-1388.D. Goldberg, *Genetic Algorithms in Search, Optimisation and Machine Learning*, Addison Wesley, Massachusetts, USA, 1989.