Glasses Detection for Face Recognition Using Bayes Rules

Zhong Jing, Roberto Mariani, Jiankang Wu
RWCP Multi-modal Functions KRDL Lab
Heng Mui Keng Terrace, Kent Ridge
Singapore 119613
Email: jzhong, rmariani, jiankang@krdl.org.sg
Tel: (+65) 874-8810, Fax: (+65) 774-4990

Abstract

In this paper, we propose a method to detect, extract and to remove the glasses from a face image. The aim is to achieve a face recognition system robust to glasses presence. The extraction is realised using Bayes rules that incorporate the features around each pixel, and the prior knowledge on glasses features that were learnt and stored in a database. Glasses removal is achieved with an adaptive median filter conducted in the points classified as glasses. The experiments indicate the performance of this method is better than that of Deformable Contour, and are satisfying.

1 Introduction

Robust face recognition system requires recognising faces correctly under different environments, such as face scale, orientation, lighting, hairstyle and wearing glasses. However glasses affects the performance of face recognition system. Previous work [1] has investigated the detection of glasses existence, however the extraction of glasses in face image is rarely addressed in literatures. Originally, we used a Deformable Contour [2] based method. Since it is only based on edge information, the result would be altered by edges of nose, eyebrow, eyes, wrinkle, etc. To solve this drawback, in this paper, we propose a new glasses point extraction method, which combines some visual features to classify glasses points and non-glasses points. Based on this method, a glasses removal approach was developed to remove the glasses points from face image in order to achieve a glasses invariant face recognition system. In our approach, Bayes rule is used to incorporate the features around one pixel extracted from face image and the prior-knowledge on glasses properties stored in a database. The prior-knowledge is represented by conditional probabilistic distribution, which is obtained by training over face image database and quantized into a feature vector table. The feature vectors describe the properties around the pixels, such as texture, intensity and direction of edge etc. After the extraction of glasses points, an adaptive median filter is applied in each pixel classified as glasses point. The pixel value is substituted by result of median filter. The method is outlined in Figure 1.

The method includes two phases: the learning phase and recognition phase. The purpose of learning phase is to produce a knowledge database that describes the visual property of glasses points. The aim of the recognition phase is to identify the glasses points by combining the visual features and the prior-knowledge produced by learning. Prior to both phases, the eyes are automatically located and the faces have been normalised. All the features are extracted within the normalised images. Edge detection is performed on the normalised image to output a binary edge image. The learning and recognition operation is only realised at those pixels labelled as edge.

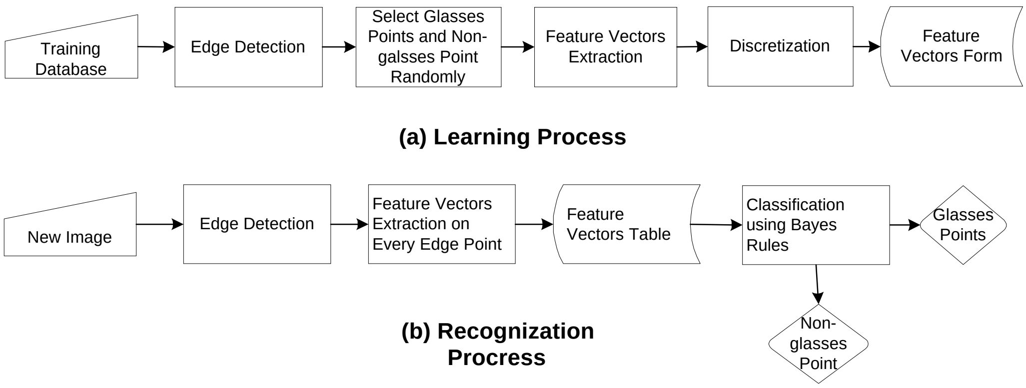
In learning phase, a set of face image wearing glasses is collected, and the glasses and non-glasses points are labelled manually on these images. The feature vectors are extracted at these points and the probability distributions of both classes are estimated over the training sets. The estimation is realised using a feature vector quantization method, which will be discussed, in section 2.2.

In recognition phase, feature vectors are extracted at each pixel points. The feature vectors are fed into a Bayes rule classifier to determine whether the pixel belongs to glasses class. The Bayes rule classifier will use the probability distribution that has been estimated in learning phase.

To remove the glasses points, adaptive median filter is applied to the image pixels classified as glasses. The filtering region is adaptively enlarged to enclose as many non-glasses pixels as the filter needs. The median value of the pixels within this region is computed and assigned to the central pixel of the region.

Our paper will be organised as follow: in the second section, we overview the Bayes rules classifier and describe the details about the learning and recognition processes. In the third section we describe the extraction of feature vectors. In the fourth section, we present the removing method for glasses points. Finally, we give the experimental results with their analysis.

Figure 1. Outline of our system



2. Glasses Point Detection Using Bayes Rule

The visual features of glasses points at the glasses frame differ drastically due to the diversity of glasses material, colours, etc. This fact indicates that it is difficult to have a unique model representing the visual features of all glasses categories. Thus, we choose a supervised learning scheme, under which the visual features of glasses are learned and stored into a feature vector table, and in the recognition phase, the learned knowledge will be recalled to determine whether the input pixel is a glasses point or not. The process can be realised by the Bayes rule [4].

2.1 Overview of Bayes Rule

The possibility of whether a pixel in face image is classified as glasses point or not can be represented as the posterior probability given the visual features at the given pixels:

$$\begin{array}{ll} P(Glasses \mid features) & \quad for\ glasses \\ P(\overline{Glasses} \mid features) & \quad for\ non-glasses \end{array}$$

We have the following decision rule to determine the pixel point:

If $P(Glasses \mid features) > P(\overline{Glasses} \mid features)$ then the pixel point is glasses point, otherwise it is not a glasses point. (1)

However, in practice, it is not applicable to directly obtain the posterior probability distribution $P(Glasses \mid features)$. Instead, we have to compute this posterior probability function from the prior probability and the conditional probability distribution (CPD) within glasses class and non-glasses class. This can be realised by Bayes formula:

$$\begin{cases} P(Glasses \mid features) = \dfrac{P(features \mid Glasses)P(Glasses)}{P(features)} \\ P(\overline{Glasses} \mid features) = \dfrac{P(features \mid \overline{Glasses})P(\overline{Glasses})}{P(features)} \end{cases} \qquad (2)$$

The conditional probability distribution $P(features \mid Glasses)$ can be obtained by learning from database of face images wearing glasses. However, the prior probability $P(Glasses)$ and $P(\overline{Glasses})$ is unknown. For solving this problem, combining the decision rule (1) and the Bayes formula (2), we have the Bayes decision rule for glasses point detection:

$$\frac{P(features \mid Glasses)}{P(features \mid \overline{Glasses})} \underset{\overline{Glasses}}{\overset{Glasses}{\gtrless}} \lambda = \frac{P(\overline{Glasses})}{P(Glasses)} \qquad (3)$$

If the left side of Equation 3 is larger than λ, we consider the pixel is glasses point, otherwise it's not. The uncertainty of prior probability is handled by the constant λ, which is used to adjust the sensitivity of glasses point detection, and its value is determined empirically.

2.2 Learning the conditional probability distribution

For the conditional probability distribution (CPD), it is necessary to define its form. In many applications of Bayes rule, it is usually assumed that the CPD has the form of Gaussian function. Unfortunately, for glasses detection, this method is not applicable, because the features of glasses points differ too much for variable types of glasses so that the Gaussian assumption is hardly accurate. Therefore, we instead use a non-parameter estimation approach.

Before training, a set of face images wearing glasses is selected from face database. Sovel edge detector is performed on the input image to produce an edge map. Only those pixles that are marked as edge would be selected for learning and recognition. Glasses points and non-glasses points are labelled manually within these images, fifty points for either left or right part of glasses frame in each image. And the neighbours of these labelled points would be added as well. This produces a 50x9x2 feature vector set for either glasses or non-glasses in one image. The feature vectors are extracted at those labelled pixels to constitute a training set. We assume the training set has *M* training samples, *M/2* for glasses and *M/2* for non-glasses. A feature vector quantization method is employed by k-mean clustering to separate the training set into *N* non-overlapped clusters. The mass centroid of each cluster is computed as the Representative Feature Vector (RFV) of the cluster. The process is illustrated as Figure 2. The result of quantization is a table, in which each entry represents a cluster with its RFV. To compute the CPD of the feature vector **f** within glasses or non-glasses, first find the nearest RFV in the table, which represent the cluster *i*:

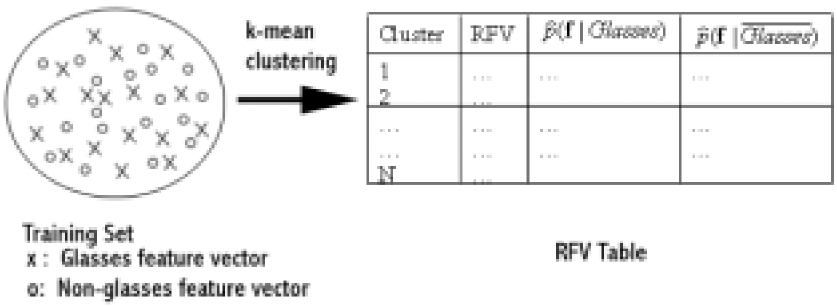
$$\|\mathbf{f}, RFV_i\| < \|\mathbf{f}, RFV_j\|, \forall i \neq j,\ i, j = 1,2,\ldots,N \qquad (4)$$

and the CPD is computed by:

$$\hat{p}(feature \mid Glasses) \approx \frac{N_{ig}}{M}$$
$$\hat{p}(features \mid \overline{Glasses}) \approx \frac{N_{i\bar{g}}}{M} \qquad (5)$$

where N_{ig} is the number of glasses points within *i*th cluster, $N_{i\bar{g}}$ is the number of non-glasses points within *i*th cluster, and *M* is the size of sample set.

Figure 2. Learning Process



2.3 Glasses Detection Process

Given a new face image, Sobel Filter is employed to obtain an edge map. For each edge pixel of the edge image, a feature vector is extracted. The nearest RFV of this feature vector is computed, and the CPD of both glasses and non-glasses of this feature vector are acquired by searching RFV Table. Finally, both CPD are fed into formula (3) to output a classification result. In our experiments, λ is set to 1.0.

3 Extraction of Feature Vectors

Feature vector is extracted at those pixels marked as edge in the edge map, which has been computed using Sobel edge detector. Three kinds of visual features are involved in glasses point detection. They are texture features, shape features using moment, and geometrical features that include the edge orientation and intensity, the distance between the point and eye, as well as the Polar Co-ordination.

Texture features are used to identify the glasses points from the eyebrows, since in some cases both of them might be very similar in shape and grey scale distribution, but different from the texture. We extract texture features based on Fast Discrete Fourier Transform [5] within an 8x8 region around the pixel. Texture features are computed based on the power spectrum after FFT. The Discrete Fourier Transform is:

$$X(u,v) = \frac{1}{NM} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} I(x,y) e^{-j2\pi(\frac{xu}{N}+\frac{yv}{M})}$$

where N,M is 8. The power spectrum is an 8x8 block:

$$P(u,v) = \sqrt{X(u,v)X(u,v)^*}$$

The texture features is derived from the power spectrum:

$$T_x = \frac{P(0,1)}{P(0,0)}, T_y = \frac{P(1,0)}{P(0,0)} \quad (6)$$

The shape features are used to identify the glasses points from the nose and other confusing points around the glasses. For this purpose, the general invariant moment [6] is employed, which is invariant to translation, rotation, scale and contrast. Given a digital image I(x,y) with the size of MxN, the (p+q)th order geometrical moments are:

$$m_{pq} = \sum_x \sum_y x^p y^q I(x,y)$$

And the central moments are:

$$\mu_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q I(x,y)$$

The first five Hu's invariant moments are:

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\phi_1 = \eta_{20} + \eta_{02}$$

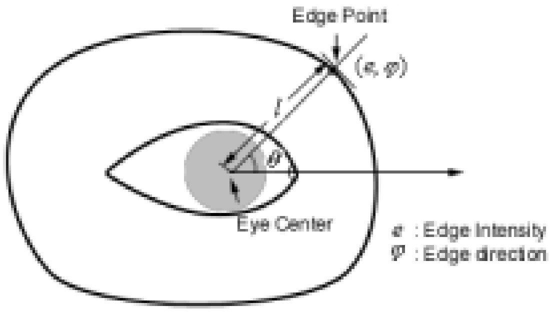
General Invariant moment is based on Hu's moments but is invariant to scale and contrast further. The first four general invariant moments are used:

$$\begin{aligned} \beta(1) &= \sqrt{\phi(2)} / \phi(1) \\ \beta(2) &= \phi(3) \cdot \mu_{00} / \phi(2) \cdot \phi(1) \\ \beta(3) &= \phi(4) / \phi(3) \\ \beta(4) &= \sqrt{\phi(5)} / \phi(4) \end{aligned} \quad (7)$$

As illustrated in Figure. 3. The geometrical features consist of the edge orientation and intensity (e,φ), the distance l between the edge pixel and the eye centre, and the direction θ of the line that

links the edge pixel and the eye centre. The geometrical features are used to distinguish the edge points belonging to the glasses frames and the edge points of the eyes and wrinkles.
Finally, we have a feature vector with the dimension of 10, where 2 features for texture, 4 features for moment and 4 features for geometrical property. In our experiments, these features are discriminative to identify the glasses points.

Figure 3. Geometrical features



4 Determination of Glasses Presence and Removal of Glasses

After glasses points classification, the number of pixels that are classified as glasses is taken as the criteria to determine whether the glasses exists or not. If the number of glasses points is greater than a threshold, we conclude that glasses exist in the face.

For every pixel in the face images wearing glasses that is classified as glasses points, the median filter is applied in a rectangular window centred at the given pixel. The rectangular window is adaptively resized to enclose as many non-glasses points as the median filter needs. The median value of the grey level of the non-glasses pixels within this region is computed and substituted to the central pixel of this region. This operation is performed for every glasses point. The result image is a face image without glasses.

5 Experiments and Results

One hundred face images wearing glasses are collected from a face image database. The image size is 200x150. The glasses varies in colour, size and shape. For training, fifty face images are randomly selected from this face database. The remaining images are used to test the performance of our method.

Before training process, we manually labelled the glasses points in glasses frames. Fifty glasses points are labelled both in left part and right part of glasses frame. These points with their neighbouring points (9 points) are used in the training step to produce the RVF table for glasses class. For non-glasses points, similar process is undertaken to produce non-glasses RVF table.

The remaining fifty face images are used to test the performance of the approach as well to compare the performance with Deformable Contour method. Figure 4(b) shows two instances of the glasses recognition. Figure 4(c) shows the resulting face image after removal of glasses from the face image. After glasses points detection based on Bayes Rules, a simple linear interpolation algorithm is employed to obtain the contour of glasses frame. The comparison results of average errors of glasses contour are shown in Table 1. The average error is computed by:

$$E = \frac{1}{M}\sum_{i=1}^{M}\left\|\mathbf{g}_i - \mathbf{t}_i\right\| \qquad (8)$$

Where, $\mathbf{t}_i$ is the ith point of glasses chain, which is obtained by manual labelling, $\mathbf{g}_i$ is the glasses point computed from glasses detection results by quantization of direction and coordinate averaging to identify corresponding points between T and G. The comparison result of glasses false detection is shown in Table 2.

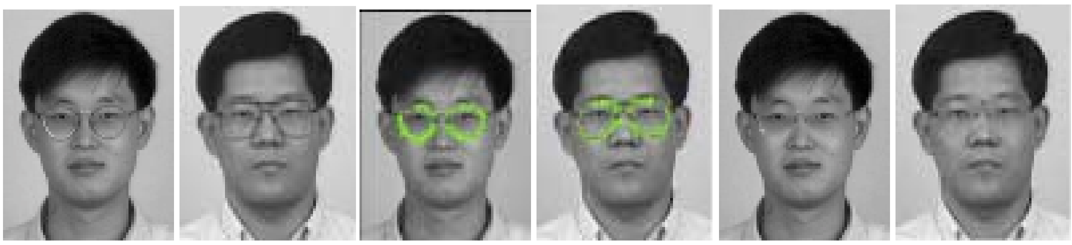
Table 1. The comparison of Deformable Contour method and Linear Interpolation based on Glasses Detection using Bayes Rule

Algorithm	Average error of Left part glasses	Average error of Right part glasses
Based on Deformable Contour	27.3297	28.9375
Linear Interpolation based on Glasses Detection using Bayes Rule	11.0963	19.627

Table 2. Comparison of False Detection of Glasses Existence

Detection Algorithm	Based on Edge Of Glasses Ridge [1]	Based on Bayes Rules
False Detection Rate	2/419	0/419

Figure 4. Results of Glasses Detection and Removal



(a) Original Images **(b) Glasses Detection based on Bayes Rules** **(c) Glasses Points Removal**

6 Conclusion

This paper describes a method for detecting glasses and an approach to remove the glasses points in face images. The experiments we conducted within a face image database showed that our method is effective and has a better performance than that of Deformable Contour based method. The visual effect of glasses removal algorithm also indicates the method can be applied in face recognition sensitive to glasses presence.

7 Refence

[1] X. Jiang, M. Binkert, B. Achermann, H. Bunke, "Towards Detection of Glasses in Facial Images", Proc. 14th ICPR, Brisbane, 1998, 1071 - 1073

[2] Kok Fung Lai, "Deformable Contours: Modelling, Extraction, Detection and Classification", Ph. D. Thesis, Electrical Engineering, University of Wisconsin-Madison, 1994.

[3] Zhong Jing and Robert Mariani, "Glasses Detection and Extraction by Deformable Contour", Proc. 15^{th}, ICPR, Spain, 2000.

[4] H. Schneiderman and T. Kanade, "Probabilistic Modeling of Local Apperance and Spatial Relationships for Object Recognition". CVPR98.

[5] Tao-I. Hsu, A.D.Calway, and R. Wilson. In "Texture Analysis Using the Multiresolution Fourier Transform", *Proc 8th Scandinavian Conference on Image Analysis*, pages 823--830. IAPR, July 1993

[6] Ming-Keui Hu, "Visual Pattern Recognition by Moment Invariants", IRE Transactions on Information Theory, pp179-187, 1962.