Face Recognition Using Vector Quantization Histogram and Support Vector Machine Classifier

Rong-sheng LI, Fei-fei LEE*, Yan YAN and Qiu CHEN

Department of School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, 516 Jun Gong Road, Shanghai, P. R. China

*Corresponding author

Keywords: Face recognition, Vector Quantization (VQ) histogram, Support Vector Machine (SVM), Gradient Descent Method (GDM).

Abstract. In this paper, we propose a new face recognition method combining Vector Quantization (VQ) method and Support Vector Machine (SVM) classifier. VQ method is used as a feature extractor and SVM classifier for feature classification. By applying low pass filtering and VQ processing to a facial image, a histogram including effective facial feature is generated, which is called VQ histogram. After dividing VQ histograms into training set and testing set, classifiers are trained with training examples (training histograms) by using Gradient Descent Method (GDM). Testing examples (testing histograms) can be tested with optimal classifiers for face recognition. We use the publicly available ORL face database for the evaluation of recognition accuracy, which consist of 400 images of 40 individuals. Experimental results show that the variety of filter size affects the recognition accuracy. The recognition rate increases with an increase of the ratio of training examples and testing examples, and maximum recognition rate of 98.0% is obtained.

Introduction

With the rapid development of the science and technology, people have paid much attention to the problem of information security. Personal identification systems utilizing personal biometric features naturally became more and more important due to its many applications in different fields, such as, identity authentication, access control, and surveillance. In face recognition, facial feature extraction is the key for recognizing examples accurately and creating more effective systems. There are a lot of techniques for extracting facial feature, for example, Principal Component analysis (PCA) [1], Edge Contour Feature analysis [2], Elastic Bunch Graph Matching [3], etc. Many classification and recognition techniques, such as, KNN [4], Neural networks classification [5] and Support Vector Machine (SVM) [6], have also been proposed. As a well-known image coding compression method, Vector Quantization (VQ) method [7] can be utilized for extracting facial features. For recognition problem, SVM classifier has been shown to give good generalization performance.

Before using VQ method, facial image is first divided into several small blocks. These blocks are regarded as input blocks, and then they will be matched with each codevector (template vectors) in codebook. For representing feature information of facial image, these blocks are compared with each codevector. The codevector having the maximum similarity to the block is selected, and then we count the similar index of matching codevector. After counting the matching frequency of each individual codevector, a histogram including facial feature information is obtained, which is called VQ histogram. By using a few codevectors to represent all small facial blocks, VQ method can reduce facial dimensions and reserve feature information.

SVM classifier has been proposed as a supervised learning method for pattern recognition [8,9,10]. SVM model [11] is schematically shown in Fig. 1. It is a binary classification method finding the optimal decision hyperplane with maximum margin between two classes (positive examples and negative examples). This method also bases on the concept of structural risk minimization [12]. In SVM model, some points near optimal decision hyperplane are called support vectors. These points play an important role in optimizing hyperplane. For optimizing hyperplane,
there are many approaches, such as, Sequential Minimal Optimization (SMO) [13]. Least squares, Gradient descent [14], etc. After generating optimal hyperplane, testing examples are just used to be tested for recognition rate. To many multi-class problems in real-world, one-against-all classification method [15] can solve this kind of problem by classifying a single class against all remaining classes. Therefore, we can use SVM classifier to solve most of recognition problems.

In Section II, Vector Quantization (VQ) method is introduced, and proposed face recognition method combining VQ method and Support Vector Machine (SVM) classifier is described in Section III. Then in Section IV, we present the face recognition experiments. The results and discussions are given in Section V. Finally, the conclusion is shown in Section VI.

**Vector Quantization Method**

Figure 2 shows the process steps of VQ histogram. Before VQ processing, an average low-pass filter is applied to filter input image. High-frequency noises are reduced, and most effective low frequency components are extracted by filtering processing. Next step is to divide input image into many small blocks. By using a dividing-partition that slides through as a one by one pixel, a facial image is divided into many 4x4 pixel small blocks with 15/16 overlap. For these 4x4 pixel small blocks, minimum intensity in a block is found, and then this minimum intensity is subtracted from each pixel in this block. Intensity variation in the block is extracted, and brightness variation is compensated in this step. Besides, a codebook has defined 33 4x4 pixel codevectors representing facial 33 dimensions. For these codevectors, thirty-two codevectors systematically vary in the direction and intensity, and the remaining codevector has no intensity variation. Next, Vector Quantization is used to intensity-variation blocks by using this codebook. Manhattan distance is utilized as a measure for representing the distance between an intensity-variation block and each codevector. The smaller the distance is, the more similar the codevector is to the block. The most similarly matched codevector to the block is then selected, and the index of this codevector is counted. For all of 4x4 pixel small blocks, those indexes of associated codevectors are counted, and then a VQ histogram is generated. This histogram is called VQ histogram. VQ histogram includes all indexes of most similar codevectors to small blocks. Therefore, VQ histogram includes most of facial feature information. In the registration procedure, VQ histogram is saved in a database as personal identification information. Thus far, all process steps of VQ histogram are finished, and a complete database is prepared for recognition problem in SVM processing.

In VQ method, low computational cost is a great advantage because of the needlessness of precise face detection. When the background is uniform, VQ histogram is basically irrelevant to face position, and position information can be ignored by using codebook. Thus, the computational cost can be reduced.
Proposed Face Recognition Algorithm Using VQ histogram and SVM Classifier

At the end of VQ processing, a registered database is prepared for recognition. In the next steps, SVM classifier is utilized to complete recognition processing. Fig.3 shows the process steps of SVM classifier. First of all, the registered database by VQ processing needs to be divided into training examples and testing examples. Training examples are used for optimizing hyperplane (1), and testing examples are used for getting recognition rate.

\[ w \cdot x_i - b = 0 \]  
\[ (1) \]

For this optimal hyperplane, gradient descent method is applied to find a minimum of object function (2).

\[ \frac{1}{2} \lambda \|w\|^2 + \sum_{i=1}^{n} \max[0, 1 - y_i \cdot (w \cdot x_i - b)] \]  
\[ (2) \]

According to the concept of structural risk minimization, hyperplane function (1) can be transformed into object function. This object function contains regularization term and loss function term. The regularization term is applied for avoiding over-fitting, and the loss function can reduce the number of error points in training examples. In function (2), \( x_i \) is the \( i \)th input vector, and \( y_i \) is the labels of \( x_i \). \( w \) is the normal vector to the hyperplane, and \( b \) is a bias. For initial value, both \( w \) and \( b \) are zero. Parameter \( \lambda \) can help object function to achieve a balance between regularization term and loss function term. After dividing database, both requisite classifiers and associated labels (\( y_i \)) of examples need to be initialized. For different classifiers, the associated labels of examples have different value. The value of labels (\( y_i \)) is +1 for the positive examples in a class and -1 for the negative examples. Next step is to loop through all examples for optimizing \( w \) and \( b \). Learning-rate \( \eta \) first is set up for step size of updating \( w \) and \( b \) by equation (3).

\[ \eta = \frac{0.01}{1 + \lambda \times t \times 0.01} \]  
\[ (3) \]

The initial value of parameter \( t \) is one, and \( t \) adds one when function margin is not less than one. If function margin of a current example is less than one, \( w \) and \( b \) are updated by equation (4).
\[ w = w' - \eta \cdot \left( x_i \cdot y_i \right) \]
\[ b = b' - \eta \cdot y_i \]

\( w' \) and \( b' \) represent previous \( w \) and \( b \). The condition of stopping looping is that difference value between objection function is less than a threshold. As the difference value is enough small, the optimal hyperplane with \( w \) and \( b \) is obtained, and then hyperplane can be utilized to test testing examples. For the different classifiers, the associated labels of test examples also have different value. By calculating the average of all classifiers’ recognition results, the final recognition rate is obtained.

In this SVM method, \( w \) and \( b \) are the key for generating SVM model. For optimizing hyperplane with \( w \) and \( b \), gradient descent is used as an optimization algorithm. It is a simple and ancient method for solving optimization problems. For finding a minimum of function, \( w \) and \( b \) need to be optimized constantly with appropriate direction and step size. The direction is the gradient of the function, and the step size is the learning-rate. Appropriate learning-rate is essential toward the minimum of function, and it gradually diminishes for converging when function value approximates a minimum. By using this method, the optimal hyperplane can be obtained conveniently.

In this SVM method, multi-class problem is involved. Though initial SVM classifier is only applied for a binary classification problem, one-against-all classification method can help SVM classifier to solve multi-class classification problems. By defining classifiers consecutively that separating a single class from all remaining classes, SVM can have a function of solving multi-class problem. For different classifiers, input examples must have different label value to associated classifiers. After recognition results of all classifiers are obtained, the average recognition result is just the final recognition rate.

**Experiments**

This experiment used the Cambridge ORL Face database. Some image examples of ORL database are shown in Figure 4. ORL database consists of 400 images of 40 individuals. These images contain different variations in pose, expression and facial details. Each image has a resolution of 92x112, and they are taken as input facial image for VQ processing. After registration processing, there is a database which contains 400 VQ histograms of 40 people. At the beginning of SVM processing, 400 VQ histograms are divided into training examples and testing examples. For example, the case of 200 training examples and 200 testing examples is equal to the ratio of five training examples and other five testing examples for one person. For multi-class problem of 40 people, 40 classifiers need to be initialized in SVM processing. Recognition experiment was carried out for testing examples after optimal hyperplanes are generated. By calculating the average recognition rate from 40 classifiers’ recognition results, the final recognition rate is obtained.
Results and Discussions

After VQ processing, some facial VQ histograms are shown in Figure 5. All of histograms have same horizontal ordinate. 33 scales represent 33 codevectors in codebook. It also implies that facial images are mapped into 33 codevectors. The sum of the matching frequency in each VQ histogram is also identical. It means that each facial image has the same number of pixel blocks. Though there are many similarities among VQ histograms, the histograms of different persons are obviously different in detail. Therefore, we infer that the difference in detail implies the difference of facial feature information.

After SVM processing, recognition results are obtained. Fig. 6 shows the recognition results as a function of filter size and ratio of training examples and testing examples. The different sizes of low-pass average filter are utilized as horizontal ordinate. “1:9”, “5:5” and “9:1” stand for the ratio case of training example and testing example, respectively, for example, one training example and other nine testing examples for one person. With an increase of the ratio from “1:9” to “5:5”, all recognition results increase. Though there are a few points showing recognition result descent when the ratio increases from “5:5” to “9:1”, most of recognition results increase. Consequently, it can be said that the increasing ratio can improve recognition rate. By providing a large amount of training examples, the classifiers become more robust and accurate. The highest recognition rate of 98% is obtained at the filter size of 7x7. Though another similar recognition rate appears at the filter size of 15x15, the ratio cases of “1:9” and “5:5” do not get highest recognition rate at 15x15. A low pass filter can keep out high-frequency noise and include useful feature information. Thus, more useful personal facial features can be extracted by using an appropriate filter, and VQ histogram can include more comprehensive facial feature with the appropriate filter. By combining comprehensive VQ histogram and robust classifiers, the new face recognition method can effectively work for personal identification.

In addition to increasing ratio of training examples and testing examples, there is another approach for improving recognition rate. It is to reduce threshold value. Threshold is used as a condition to limit objection function for optimal hyperplane. When the difference between objection function is less than the threshold, it indicates that w and b is optimal to hyperplane. Thus, w and b become more optimal by reducing threshold, and classifiers can become more accurate for recognition problem. Fig. 7 shows the recognition results obtained by reducing threshold. Threshold values are utilized as horizontal ordinate. By reducing threshold gradually, the recognition rate increases by about 0.2%.

Conclusions

This paper proposes a new face recognition method combining VQ method and SVM classifier. With an appropriate filter size, VQ histograms include more comprehensive facial feature information.
information. By providing a large amount of training examples, SVM model can become more robust and accurate. Experimental evidence shows the excellent face recognition performance by applying the ORL face database. VQ method can extract more useful feature information, and SVM classifier effectively solve the recognition problem. It can be said that this face recognition algorithm combining two methods as showed above, is powerful and effective for personal identification problem.

References


