PCA와 SVM에 기반하는 빠른 얼굴 탐지 방법

하춘뢰* · 신현갑* · 박명철* · 하석운**

A Fast Method for Face Detection Based on PCA and SVM

Chun-Lei Xia* · Hyeon-Gab Shin* · Myeong-Chul Park* · Seok-Wun Ha**

요 약

얼굴인식기술은 컴퓨터비전 분야에서 중요한 역할을 담당하고 있다. 본 논문에서는, PCA와 SVM 기술을 사용하는 빠른 얼굴인식기술을 제안한다. 제안한 시스템에서는, 먼저 지역 히스토그램 분포를 분석하여 생성한 통계적 특성을 사용함으로써 얼굴가능영역을 필터링한다. 이 과정에서 대부분의 비얼굴 영역이 제거되기 때문에 탐지 과정의 처리속도가 향상된다. 다음으로는 PCA 특징 벡터가 생성되고, SVM 분류기를 사용하여 테스트 영상 내에 얼굴이 존재하는지를 탐지한다. 본 논문에서의 테스트 영상은 CMU 얼굴 데이터베이스를 사용하였으며, SVM의 학습을 위한 얼굴과 비얼굴 샘플들은 MIT 데이터 세트로부터 선택하였다. 얼굴탐지 실험결과, 제안한 방법에서 좋은 성능을 나타내었다.

ABSTRACT

Human face detection technique plays an important role in computer vision area. It has lots of applications such as face recognition, video surveillance, human computer interface, face image database management, and querying image databases.

In this paper, a fast face detection approach using Principal Component Analysis (PCA) and Support Vector Machines (SVM) is proposed based on the previous study on face detection technique. In the proposed detection system, firstly it filter the face potential area using statistical feature which is generated by analyzing the local histogram distribution, the detection process is speeded up by eliminating most of the non-face area in this step. In the next step, PCA feature vectors are generated, and then detect whether there are faces present in the test image using SVM classifier. Finally, store the detection results and output the results on the test image.

The test images in this paper are from CMU face database. The face and non-face samples are selected from the MIT data set. The experimental results indicate the proposed method has good performance for face detection.

키워드

SVM, PCA, histogram distribution, face potential area

Ⅰ. INTRODUCTION

Face detection is a very useful and helpful technique and it plays an important role in the real applications. It is the first step of any face processing system is detecting the locations in images where faces are present. Face detection is the first step for any automatic face recognition system, first step in many surveillance systems, first step for object recognition or object tracking. And it is also could be a part of identification system.

* 경상대학교 컴퓨터과학과

** 경상대학교 컴퓨터과학과 교수, RICIC 연구원, 교신저자

접수일자 : 2007. 6. 1

Face detection has been researched for more the 20 years. Nowadays, face detection is still concerned by amount of researchers in computer vision fields. There are several difficulties, such as pose, facial expression, occlusion, illumination are the factors that can affect the accuracy of face detection.

People used lots of methods or algorithms to solve the face detection problem [1]. In template matching, the correlation values with the standard patterns are computed for the face contour, eyes, nose, and mouth independently. The existence of a face is determined based on the correlation values. This approach has the advantage of being simple to implement.

Neural networks have been applied successfully in many pattern recognition problems. However, one drawback is that the network architecture has to be extensively tuned to get exceptional performance.

Support vector machine operates on structural risk minimization, which aims to minimize an upper bound on the expected generalization error. SVM have a good performance in pattern recognition problem too, it is become popular in recent years. SVM is selected as the classifier in the proposed face detection system, because the experiments show SVM has a better performance than neural network. More details about SVM are described in later sections.

In this paper, we proposed a fast method for face detection, it is called face potential area selection method. This face potential area selection method could eliminate most of non-face area according the character of histogram distribution of face samples and non-face samples. By combining PCA transform and SVM algorithm, a face detection system is presented.

This paper is organized as follow: The detail of PCA transform and PCA feature generation are given in section II. In section III, how does the Support Vector Machine work is explained. SVM for face detection is presented in section IV. The proposed method, face potential area selection is presented in section V. In section VI, it shows the detection result of the proposed method. Finally, Conclusion and future works are given in the last section.

II. PCA FEATURE EXTRACTION

Principal Component Analysis (PCA) is a useful statistical tool that has found application in fields such as face detection and image compression and it is a powerful tool for analyzing data. PCA is a common technique for finding patterns in data of high dimension. Usually, PCA is used for reducing dimension of feature vectors. Applying PCA transform, the largest features of samples could be obtained.

The central idea of PCA is to find a low dimensional subspace (the feature space) which represents most of the variation within the sample data. Sample data can be approximately reconstructed with only part of their projection onto the PCA subspace. After projection onto the principal component axis, the feature vectors are found on the principal component. And the feature vectors are sorted in descent.

Let $T_X = \{x_1, \cdots, x_l\}$ be a set of training vectors from the n-dimensional input space R^n. The set of vectors $T_Z = \{z_1, \cdots, z_l\}$ is a lower dimensional representation of the input training vectors T_X in the m-dimensional space R^m. The vectors T_Z are obtained by the linear orthonormal projection

$$z = W^T x + b, \tag{1}$$

where the matrix $W[n \times m]$ and the vector $b[m \times 1]$ are parameters of the projection. The reconstructed vectors $T_{\tilde{X}} = \{\tilde{x}_1, \cdots, \tilde{x}_l\}$ are computed by the linear back projection

$$\tilde{x} = W(z - b), \tag{2}$$

obtained by inverting (1). The mean square reconstruction error

$$\epsilon MS(W,b) = \frac{1}{l}\sum_{i=1}^{l} \| x_i - \tilde{x}_i \|^2, \tag{3}$$

is a function of the parameters of the linear projections (1) and (2). The Principal Component Analysis (PCA) is the linear orthonormal projection (1) which allows for the minimal mean square reconstruction error (3) of the training data T_X. The parameters (W, b) of the linear projection are the solution of the optimization task

$$(\mathbf{W}, b) = \underset{\mathbf{W}', b'}{\operatorname{argmin}}\ \varepsilon_{MS}(\mathbf{W}', b'), \qquad (4)$$

subject to

$$\langle w_i \bullet w_j \rangle = \delta(i,j), \forall\, i,j$$

where $w_i, i = 1, \ldots, m$ are column vectors of the matrix $W = [w_1, \ldots, w_m]$ and $\delta(i,j)$ is the Kronecker delta function. The solution of the task (4) is the matrix $W = [w_1, \ldots, w_m]$ containing the m eigen vectors of the sample covariance matrix which have the largest eigen values. The vector b equals to $W^T \mu$, where μ is the sample mean of the training data.

Figure 1 shows the PCA feature vectors and subspace. The small circles on the PCA subspace are the vectors which are projected onto the principal component axis, they are the feature vectors generated by PCA transform.

In general, we don't have to put all the feature vectors to the application, because just a part of largest PCA feature vectors can contain most features of the sample data, so just a part of largest vectors are needed as feature vectors in the application. This is called dimension reduction of the feature vectors. PCA is an effective and frequently algorithm used for feature generation.

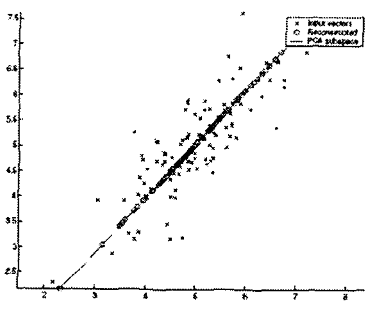
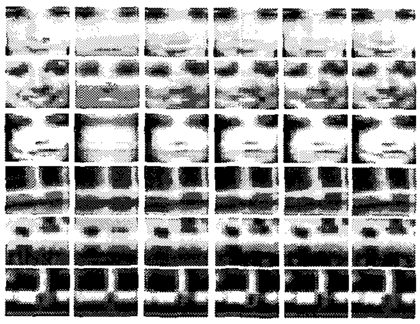


Fig. 1. Principal Component Analysis subspace

PCA transform has an inverse transform, the original image could be reconstructed by the inverse PCA transform using the feature vectors which are generated from the PCA transform. Here given some face and non-face sample images, after PCA transform and inverse PCA transform, the sample images are reconstructed. The image reconstruction results are shown in Figure 2. Images are reconstructed by *different numbers of PCA feature vectors.*

The first column in Figure 2 is original image, the other columns are the reconstructed images, and we can compare the difference between original image and the reconstructed image. From the comparison, we found the image reconstructed by 5 feature vectors has the most similarity with the original image, that means only 5 feature vectors contain most information of the original image. And also, 6 or 7 feature vectors could contain more information of the original image. But too many feature vectors will make the training procedure and detection procedure taking more time and the detection ratio won't be improved too much. Using 5 feature vectors is proven by practice tests, the results indicate it has better performance. 5 feature vectors for each sample is the optimal choice. Consider the feature representation and system performance, 5 feature vectors are selected as the input feature vectors in the face detection system.



(a) (b) (c) (d) (e) (f)

Fig. 2. PCA reconstruction results. (a) Original image (b)-(f) Reconstructed images by 1-5 feature vectors.

Ⅲ. SUPPORT VECTOR MACHINES

Support Vector Machines (SVM) was proposed by Vapnik V.N in 1995. After more then 10 years practice and development, it has become an increasingly popular tool for machine learning task involving classification, regression or other pattern recognition problems. They exhibit good generalization performance on many real-life data sets and the approach is well-motivated theoretically. For the object recognition, classification problem, lots of experiments indicate SVM a more accuracy and generalization performance then neural network or other algorithms, so in this paper SVM instead of neural network is selected as the classifier.

In SVM algorithm, there is a hyperplane used for separate different classes. The goal of SVM is to find out an optimal separable hyperplane to solve the classification task. The optimal hyperplane has maximal distance from each class to the hyperplane. In this section, first, discuss linear separable classification. And then, nonlinear case and kernel function are described.

In linear separable case, all the data of the same class are on the same side of the hyperplane. We illustrate an example to explain this problem in figure 3.

This is basic binary classification example in figure 3. It can be seen from figure 3, the cross points on the left side of hyperplane are pattern one, and the circle points on another side are pattern two. In linear case, the classification hyperplane is a straight line.

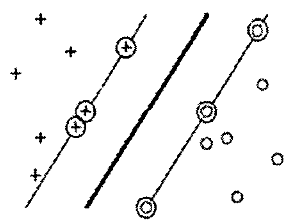


Fig. 3. Hyperplane and support vectors in linear separable case

We suppose a set of given data is that

$\{x_i, y_i\}, i = 1 \ldots l, y_i \in \{-1, 1\}, x_i \in R^d$.

The separating hyperplane is $W \cdot X + b = 0$.

The margin is $\frac{1}{\| W \|} + \frac{1}{\| W \|} = \frac{2}{\| W \|}$.

Then we compute the minimal margin, we use the Lagrange multiplier figure out the follow equations

$$W = \sum_{i=1}^{n} a_i y_i X_i, \tag{5}$$

$$L(a) = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i,j=1}^{n} a_i a_j y_i y_j x_i \cdot x_j. \tag{6}$$

The α in equation (5) is the Lagrange multiplier, the vectors which are associated with $\alpha i \neq 0$ called support vectors. Those vectors are on the boundary of the separating margin, that are the points with circles presented in Figure 3.

We maximize equation (6) subject to constraint

$$\sum_{i=1}^{n} \alpha_i y_i = 0 \text{ with } \alpha \geq 0.$$

The optimal Lagrange multipliers have been computed, and the optimal separating hyperplane is found.

But unfortunately, not in all the cases there isa hyperplane that can solve the classification problem. In nonlinear case, we have to map the original feature space into a high-dimensional space that we can figure out an optimal hyperplane, this is the solution for nonlinear classification.

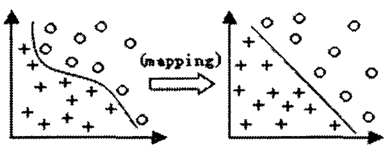


Fig. 4. The mapping from input space to a high dimension feature space

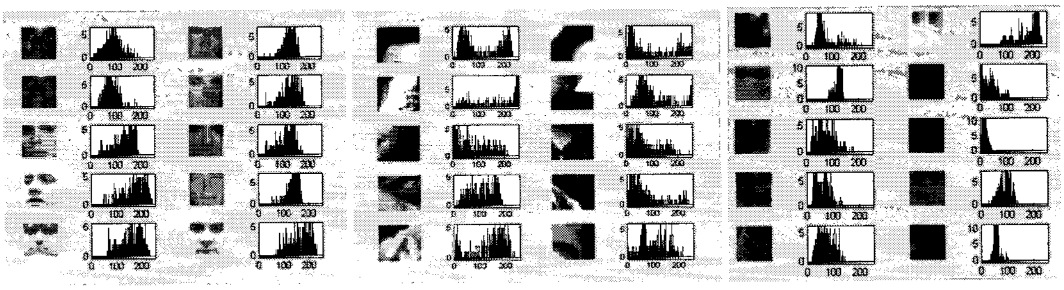
A kernel function is used here, it response to the mapping from input space to feature space.

Three kernel functions are usually used, polynomial function, RBF function and sigmoid function.

Polynomial: $K(X_i, X_j) = (\gamma X_i^T xj + r)^d, \gamma > 0.$

RBF:

$K(X_i, X_j) = \exp(-\gamma \|X_i - X_j\|^2), \gamma > 0.$



(a) (b) (c)

Fig. 6 Histogram distribution of face and non-face samples

Sigmoid: $K(X_i, X_j) = \tanh(\gamma X_i^T X_j + \tau)$.

Here, γ, τ and d are kernel parameters.

The kernel function nonlinearly maps samples into a higher dimensional space. The RBF kernel has less numerical difficulties; its computation is not as complex as the other kernel's. So the RBF kernel is chosen in the experiments.

Ⅳ. SVM FOR FACE DETECTION

Face detection is thought as a binary classification problem in SVM. There are only two patterns are involved in the classification. One pattern is defined as face pattern, another pattern is non-face pattern. In the training phase, it tries to find out a hyperplane which can separate face patterns and non-face patterns.

The face detection results depend on the accuracy of the SVM classifier. In the detection procedure, a given input data, SVM classifier identify it is a face pattern or non-face pattern by the hyperplane. In another word, there need an effective and clear separating hyperplane. Figure 5 illustrates the face classification using SVM. The blue bold curve is the hyperplane. On the left side of the hyperplane in the feature space is defined as face pattern. On the contrary, if the input data falls in the right side of hyperplane it is classified as non-face pattern. As mentioned in previous section, RBF kernel function is used in SVM face detection. The optimal parameters could be found for SVM classifier by lots of training tests.

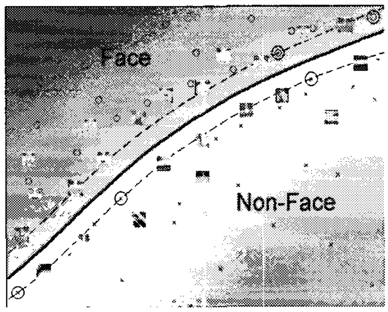


Fig. 5. Face detection in SVM feature space

Ⅴ. FACE POTENTIAL AREA SELECTION

In the detection procedure, it scans each position on the test image, and identify whether there is a face present or not. Usually, in an image, non-face areas are far more than face areas. It always wastes us much more time to detect face on large number of non-face area. It also might cause more misclassification regions in the non-face area.

In this paper, we propose a method for eliminating most of the non-face area in gray images, that it can save more detecting time and reduce the misclassification ratio. Face area has different statistical character with most of the non-face area. By analyzing histogram distributions, it shows face and non-face area have different histogram distribution features. It can be seen from figure 6. The histogram of face area has Gaussian-like distribution, the pixels are always distributed in normal distribution. But non-face area histogram has irregular distribution, it could be distributed in a short range or distributed in a wide range, and also it could be distributed in normal distribution. So we can find some rules for distinguish most of the non-face areas.

According to the histogram distribution feature, the face potential area could be selected. The histogram which is distributed in a small range, its mean value must be a high value. If the histogram distribution is widely distributed, it has a low mean value. The histogram of face image is a Gaussian-like distribution; the mean value is an intermediate value. This intermediate value is in a fixed range. By a number of tests, we figure out the range of histogram mean value of the face potential area. The fixed range is the mean value range of face potential area. So if the mean value of a sample area is in that fixed range, this sample area could bea face potential area. Otherwise, it is filtered as non-face area.

Ⅵ. EXPERIMENTS AND RESULTS

The experiments are taken based on MIT data sets and CMU face database. MIT and CMU data sets are usually used for testing face detection algorithms. The MIT face data sets provide face and non-face training samples. The SVM classifier is trained using these samples. In the detection procedure, all the test images are from the CMU face databases.

In order to get the best performance, optimal parameters are chosen for training the SVM classifier by a number of practices.

In the detection part, the face potential areas are selected first, and then identify each possible position in the selected face potential area. The feature vectors are generated by PCA transform and delivered to SVM classifier. Finally, we mark all of the detected face regions in the test images.

Figure 7 shows face potential area selection and face detection results. More detection results are shown figure 8.

Ⅶ. CONCLUSIONS

This paper proposed a fast face detection method based on PCA and SVM. Combining PCA and SVM has high performance in face detection task. In addition, the proposed method filters the face potential areas to save the detecting time and reduce the wrong detection results. It can be seen from the results, most of the non-face areas have been eliminated. The detection time is saved by eliminating the non-face areas. Obviously, the proposed method works faster than the face detection that only combined PCA and SVM. The experimental results indicated a high detection ratio and low misclassified ratio. In the future works, the performance of classifier and the face potential area selection method are also could be improved.

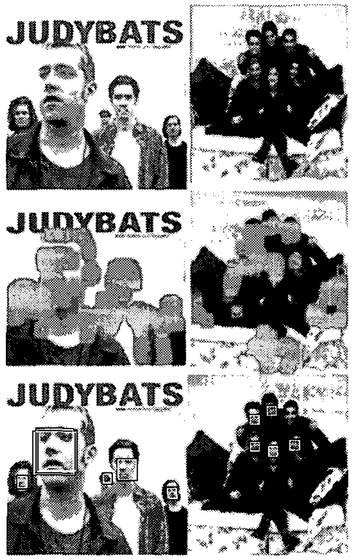


Fig. 7. Selected face potential area and detection results

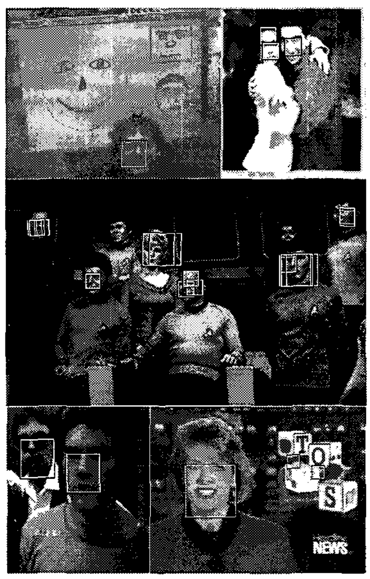


Fig. 8. Detection results

REFERENCES

[1] Ming-Hsuan Yang, David J. Kriegman, Narendra Ahuja "Detecting Faces in Images: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, Jan. 2002.

[2] Clyde Shavers, Robert Li, Gary Lebby "An SVM-based Approach to Face Detection", Proceedings of the 38th Southeastern Symposium on System Theory Tennessee Technological University, Mar. 2006.

[3] Lindsay I Smith "A tutorial on Principal Components Analysis", http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf, Feb. 2002.

[4] E. Osuna, R. Freund, F. Girosi "Training Support Vector Machines: An Application to Face Detection," Proceedings IEEE Conference Computer Vision and Pattern Recognition, 1997.

[5] Christopher J.C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, 1998

[6] B. Heisele, T. Poggio, Massimiliano Pontil "Face Detection in Still Gray Images", AI Memo 1687, Massachusetts Institute of Technology, 2000.

[7] Rafael C. Gonzalez, Richard E. Woods "Digital Image Processing (Second Edition)", Prentice Hall, 2002.

[8] C. Papageorgiou, T. Poggio "A Trainable System for Object Detection", International Journal of Computer Vision, Volume 38, Issue 1, Jun. 2000.

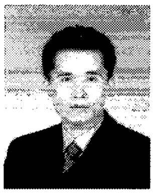
저자소개

하 춘 뢰(Chun-Lei Xia)

2005년 [중국]안산과학기술대학교 컴퓨터과학과(공학사)
2007년 경상대학교 컴퓨터과학과 석사수료
※관심분야 : 컴퓨터비전, 얼굴인식, 객체추적

신 현 갑(Hyeon-Gab Shin)

2000년 진주산업대학교 (농학사)
2003년 경상대학교 소프트웨어학과 (공학석사)
2006년 경상대학교 컴퓨터과학과 박사수료
※관심분야 : 컴퓨터비전, HCI, 객체인식, 임베디드 시스템



박 명 철(Myeong-Chul Park)

1999년 한국방송대학교 컴퓨터과학과 (이학사)
2002년 경상대학교 소프트웨어학과 (공학석사)
2007년 경상대학교 컴퓨터과학과(공학박사)
※관심분야 : 컴퓨터비전, 시각화, 임베디드소프트웨어

하 석 운(Seok-Wun Ha)

1979년 부산대학교 전자공학과 (공학사)
1986년 부산대학교 전자공학과 (공학석사)
1995년 부산대학교 전자공학과(공학박사)
2002년-2003년 캘리포니아대학교 연구교수
1993년 3월 ~ 경상대학교 컴퓨터과학과 교수
※관심분야 : 신경회로망, 영상처리, 컴퓨터비전