A Face Recognition System Based on Local Binary Patterns and Support Vector Machine for Home Security Service Robot

Jiakailin Wang, Jinjin Zheng, Shiwu Zhang, Jijun He, Xiao Liang and Sui Feng
Department of Precision Machinery and Precision Instrumentation
University of Science and Technology of China
Hefei, China
corresponding author Jiakailin Wang,wangjiakailin@163.com

***Abstract*—In this paper, we present a real-time face recognition system for home security service robot which can be applied to recognize the person's face in front and give a warning when the identity of the person is a stranger. Considering the complexity of the actual situation, there might be some errors causing by the following factors like the angle, the size, the environment and the illumination of the human face, which can hardly be avoided. Our system is designed to reduce and eliminate the influence of the factors above. This article first used Local Binary Patterns (LBP) to detect face and cut out the face region. Then, based on the collected and pretreated face database, the target face could be recognized utilizing Support Vector Machine (SVM). However, the given face database are always irregular in size, position of the face area, scale and some other factors. The proposed face recognition system has been tested in practical application for home security service robot. The results indicate that our face recognition system is effective and real-time. It works well in multi-face recognition and stranger identification, which meet the demand of robot.**

Keywords-face recognition; stranger identification; home security service robot; local binary patterns; support vector machine

I. INTRODUCTION

Nowadays, people's security awareness and requirement for home service is gradually rising. How to combine home service and security on one robot is worth studying. On the other hand, face recognition, which is a biometric identification technology based on human face feature information, has been a hot research topic for a long time [1], [2]. The face recognition system proposed in this paper, where the camera was used to capture the video stream, can automatically detect and track the human face in the image frame, and then recognize the human's identity. However, the human faces in the video frame are not always completely front faces. Hence, the method should be able to detect the face with angle. According to some recent researches, the common methods of face detection include harr-like feature [3] and LBP feature [4]. It is found that LBP works better in incomplete frontal human faces. In face recognition, eigenfaces [5] and fisherfaces [6] have been widely used. Moreover, SVM [7],[8] with many unique advantages in solving small sample, nonlinear and high dimensional pattern recognition problems has been applied to handwriting recognition, 3D object recognition, face recognition, text image classification and other practical problems, showing a good learning ability. Compared with eigenfaces and fisherfaces, recognition using SVM has a higher accuracy at Labeled Faces in the Wild (LFW) face database [9]. The decision rules obtained from the limited training samples can still get a small error in the independent test set, which indicates the SVM face recognition is very effective in practical application.

The face recognition system introduced in this paper is applied in the home security service robot. The home security service robot is combined with motion system, remote control, network video monitoring, main control system, human-computer interaction, navigation, visual recognition, voice recognition and other modules as shown in Fig.1.

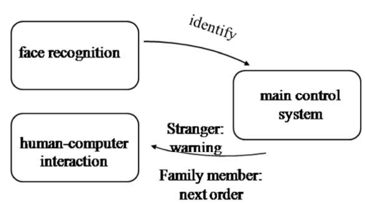

Figure 1. Face recognition module

In our family security module, the function of our face system in this part is to judge whether the person is a stranger and give a warning if not. In the design of family service function, the face recognition system is to identify the person, and then pass the identity information to some related modules for other works .

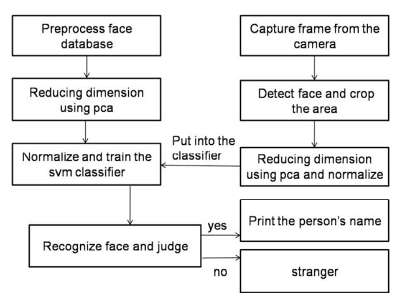

Figure 2. The total flow chart

2473-3547/16 $31.00 © 2016 IEEE
DOI 10.1109/ISCID.2016.184

Fig. 2 shows the total flow chart. This paper focuses on the face detection with LBP features, face recognition with SVM, and also the images' preprocessing. Before real-time recognition, we need to deal with the face database and train the SVM classifier including normalization, dimension reduction and preprocess explained in the next parts.

II. PREPROCESSING AND DIMENSION REDUCTION

A. Alignment

The face database used in this system is composed of dozens of each family member's wild face images. The proportion and position of the face in the collected images are not uniform. So alignment and cropping are necessary in order to facilitate the accuracy of the following works. The alignment method proposed in this paper is to detect the eye center of the people in each picture and use this position as the benchmark to align and crop.

The circular eye's region obtained by detection prints some information like eye's center coordinate and radius. Even if the proportion and position of the face in each images is different. Eye's center can be used as the benchmark to extend outward in proportion according to the proportion of eye in the whole image. The same images before alignment and after are showing in Fig.3

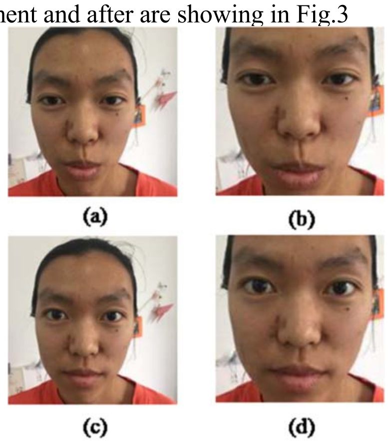

Figure 3. Before alignment and after: the images (a), (c) are before alignment and the (b), (d) are after alignment.

B. Dimension reduction

In this paper, Principal Component Analysis (PCA) [10] is used to reduce the dimension. PCA is not only to reduce the dimension of high-dimensional data, but also to remove the noise from the dimension.

PCA displays the original n features with a number of less m features. The new features are the linear combinations of the old features. The sample variance is maximized by these linear combinations, as far as possible to make the new m features not related to each other. The inherent variability in the data mapping can be captured from the old features to the new features.

III. FACE DETECTION BASED ON LBP

A. LBP (Local Binary Patterns)

LBP refers to Local binary patterns. It is an operator used to describe the local texture characteristics of images. And it has some significant advantages like rotation invariance, gray invariance and others. The original LBP operator is a powerful method of texture description. The operator labels the pixels of an image by comparing the 3×3 neighborhood of each pixel with the center value. If the neighbor pixel is bigger than the center's it is written as 1, otherwise 0.So the result is an eight-bit binary number [11]. The decimal number converted from the binary number is the LBP code of the center pixel in this window and this value can be used to reflect the texture information of the region. Then the histogram of the labels can be used as a texture descriptor. An illustration of the basic LBP operator is shown in Fig. 4. Besides, an operator named Harr-like is also widely used in face detection [12]. Compared with Harr-like feature, LBP feature is an integral feature, so the training and detection process will be faster.

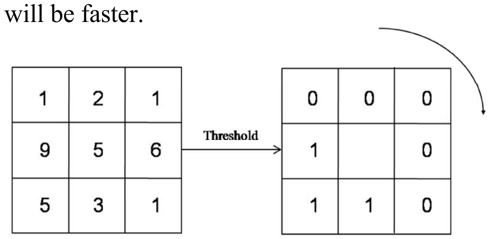

Figure 4. The basic LBP operator: the binary is 00010011 and 19 converted to decimal.

B. Face detection

In this paper, the face detection algorithm uses LBP classifier to distinguish between face and non-face. It can determine whether there is a face in the collected video frame and extract the face area.

1) Use LBP features to do detection: We use LBP operator to scan the whole image and get the LBP feature map.

2) Use AdaBoost algorithm introduced by Freund and Schapire [13], [14] to train the strong classifier from week classifiers between face and non-face:

3) Cascade the strong classifiers together to improve the accuracy.

The set of intra or extra features obtained by scanning the sub-window is over complete for the intrinsically low dimensional face appearance pattern. To reduce the dimensionality subspace techniques have been used [15], [16]. Adaboost is used to select most significant intra or extra features from a large feature set [17]. Therefore, AdaBoost works to solve the following three fundamental problems:

- Learn effective features from a large feature set.
- Train weak classifiers based on the selected features.
- Boost the weak classifiers into a stronger classifier.

IV. FACE RECOGNITION BASED ON SVM

A support vector machine (SVM) is generally a binary classifier based on minimization of structural risk. In complicated nonlinear modeling, the SVM transforms the training data into a higher dimensional feature space and there is a linear hyper-plane can separate the data [18]. In the simplest two-dimensional plane, the SVM training phase tries to find the linear function to separate the data .

Because these data are linear, so you can use a straight line to separate these two types of data, which is equivalent to a super plane this super plane can be expressed by(1).

$$f(x) = \omega^T x + b \tag{1}$$

where ω is the weight vector, x is the input data, and b is considered as bias. The SVM classifier is aiming to find possible hyper-planes which maximize the margin between two classes with support vectors at closest point from the hyper-plane.

To classify the data, the further the interval from the data points to the hyper-plane, the higher the confidence of the classification is. Therefore, in order to make the confidence as high as possible, we need to choose a hyper-plane which can maximize the interval value. This interval expressed as gap/2 is shown in the Fig. 5

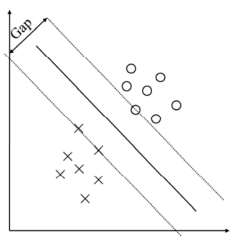

Figure 5. Interval from the data points to thehyper-plane

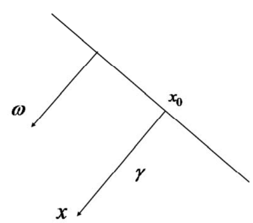

Figure 6. Distance from x to the hyper-plane

Intuitively, this hyper-plane should be the most suitable straight line to separate the two types of data. The standard is that the data on both sides has the maximum distance from the straight line. So, the maximum interval of the hyper-plane needs to be found.

The following formula can be got from Fig. 6: x is a point outside the hyper-plane and x_0 is the projection of x on the hyper-plane.

$$x = x_0 + \gamma \frac{\omega}{\|\omega\|} \tag{2}$$

where $\|\omega\|$ is the bound norm, γ is the distance from point to the hyper-plane. Because point x_0 is on the hyper-plane and $f(x_0) = 0$, so we can have:

$$\gamma = \frac{\omega^T x + b}{\|\omega\|} = \frac{f(x)}{\|\omega\|} \tag{3}$$

The geometric interval is defined as (y represents the class):

$$\tilde{\gamma} = y(\omega^T x + b) = yf(x) \tag{4}$$

$$\tilde{\gamma} = y\gamma = \frac{\gamma}{\|\omega\|} \tag{5}$$

So the objective function can be defined as the maximal interval:

$$\max \tilde{\gamma} \tag{6}$$

Constraint condition:

$$y_i(\omega^T x_i + b) = \tilde{\gamma}_i \ge \gamma, i = 1, \ldots, n \tag{7}$$

We can get the hyper-plane (classifier) by solving (6) with (7).

The problem discussed above is about two classes. But in our face recognition system, the classes are far more than two. To solve multiclass problems, there are many strategies like "one-against-one", "one-against-many" and "many-against-many". We use the "one-against-one" strategy, which constructs one SVM for each pair of classes. For a problem with 5 classes, 10 SVMs are trained to distinguish the samples of one class from the samples of another. The label which is the one with the most votes among the classifiers [19] is selected. In addition, with more classes the "one-against-one" strategy is more accurate than other methods especially when it has few training samples per class does not cause any problem [20].

V. EXPERIMENTS

A. *Experimental process*

Our program runs on the i.MX6Q e9 card computer with 4×1GHz frequency, 2GB DRR3 and 8GB eMMC flash .

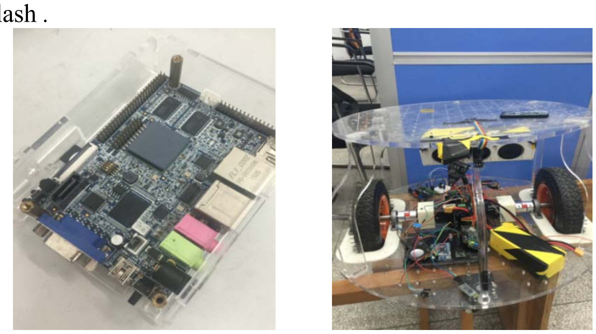

Figure 7 i.MX6Q e9 card computer and Robot chassis

a) We first preprocess the collected face images including alignment and cropping them to 200×200. The images need to be stored according to classes.

b) The second step is using PCA to reduce the dimension and normalization.

c) Put the feature vectors obtained at the second step into our SVM classifier, training data and generating model. The kernel function we used is RBF (radial basis function).

d) Use LBP to detect the face area from the video frame collected by the camera and cut out the face area.

e) Deal the collected face images with PCA and normalization.

f) Send the vectors we get from step 5 into the trained SVM classifier.

g) Recognize the face to get the identity and judge whether this person is a stranger.

In the course of the experiment, feature faces gotten after dimension reduction with PCA are shown in Fig.8.

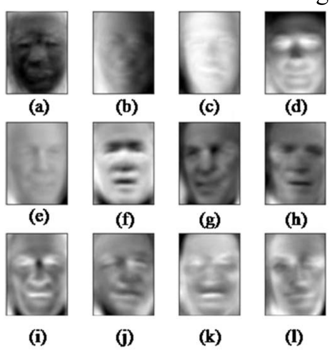

Figure 8. PCA result: Use PCA to do dimension reduction with several images from the LFW database

By the constant rotation and testing of human face in reality, The deflection angle range can be estimated in which human face can be detected. According to the results in Fig.9, we find that harr-like face detection does not work well when there is deflection angle and it cannot be avoided in reality.

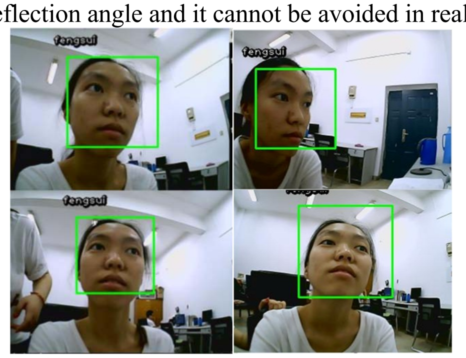

Figure 9. Harr-like results(left) and LBP results(right) of face detection with deflection angle

Compared with the Harr-like result, it can be found that LBP can detect face with bigger angle. So face detection with LBP feature is more effective in practical application and selected to service for our project.

According to some researches, Eigenfaces and Fisherfaces are chosen from some common real-time face recognition methods to compare with SVM. They are tested on a database named Labeled Faces in the Wild (LFW) with these three methods. The precision of recognizing the face correctly are shown in Fig. 10.

	precision
Eigenfaces	0.59
Fisherfaces	0.81
svm	0.86

Figure 10. Results of face recognition on LFW database

It can be seen from the experimental results that precision of SVM is better than the other two. So we choose SVM as the face recognition method in this project.

B. Actual function demonstration

Detect the faces in the camera: The faces appearing in the frame all can be detected Fig. 11.

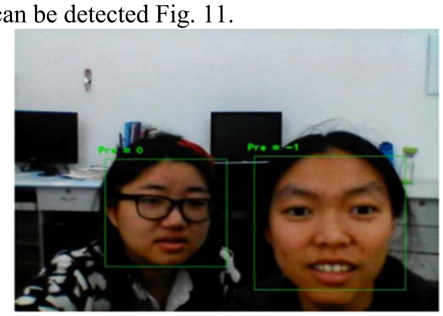

Figure 11. Multi-face recognition: Two people in front of the camera can all be recognized.

About 10 people are selected to establish the family members' face database and they can all be recognized correctly

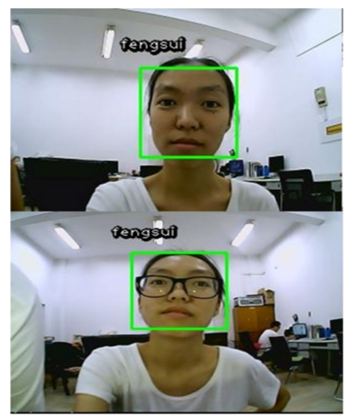

Figure 12. Face recognition (with glasses and not) and print the person's identity

According to the principle of SVM classifier, the bigger the distance of the sample to the separating hyper-plane is, the more accurate the classification result will be. By the experience gotten from experiments, a threshold is used to judge the reliability of the classification result. In other words, if the distance of the sample to the separating hyper-plane is bigger than the threshold, the identify printed is credible, otherwise the person will be recognized as a stranger. The person in Fig. 13 has no image in the face database, so she is recognized as a stranger correctly.

Figure 13. Recognize stranger

VI. CONCLUSION

The real-time face recognition system presented in this paper by combining face detection and face recognition can work well for home security service robot .The functions of identifying the identity of the person in front and giving a warning if there are strangers pass the test in reality. With comparisons of some tests, the face detection with LBP has good adaptation to the face angle and the face recognition with SVM can accurately recognize family members or strangers. Experimental results in reality show that the function mentioned can satisfy the demand of home security service robot.

ACKNOWLEDGMENT

This work was supported by NSFC-CAS Joint Fund (No.U1332130), 111 Projects (No.B07033), 973 Project (N0.2014CB931804).

REFERENCES

[1] L. Wiskott, J.M. Fellous, N. Kruger, and C. Vondermalsburg. “Face recognition by elastic bunch graph matching”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):775–779, 1997.

[2] Phillips, P. Jonathon, et al. “The FERET evaluation methodology for face-recognition algorithms.” IEEE Transactions on pattern analysis and machine intelligence 22.10 (2000): 1090-1104.

[3] Viola, Paul A., and Michael J. Jones. “Robust Real-Time Face Detection.”ICCV. Vol. 2. 2001.

[4] T. Ojala, M. Pietikainen, and D. Harwood. “A comparative study of texture measures with classification based on feature distributions”. Pattern Recognition, 29(1):51–59, January 1996.

[5] Turk, Matthew, and Alex Pentland. “Eigenfaces for recognition.” Journal of cognitive neuroscience 3.1 (1991): 71-86.

[6] Belhumeur, Peter N., João P. Hespanha, and David J. Kriegman. “Eigenfaces vs. fisherfaces: Recognition using class specific linear projection.”IEEE Transactions on pattern analysis and machine intelligence 19.7 (1997): 711-720.

[7] Phillips, P. Jonathon. Support vector machines applied to face recognition. Vol. 285. US Department of Commerce, Technology Administration, National Institute of Standards and Technology, 1998.

[8] Osuna, Edgar, Robert Freund, and Federico Girosit. “Training support vector machines: an application to face detection.” Computer vision and pattern recognition, 1997. Proceedings., 1997 IEEE computer society conference on. IEEE, 1997.

[9] Learned-Miller, Erik, et al. “Labeled faces in the wild: A survey.” Advances in Face Detection and Facial Image Analysis. Springer International Publishing, 2016. 189-248.

[10] Jolliffe, Ian. Principal component analysis. John Wiley & Sons, Ltd, 2002.

[11] T. Ojala, M. Pietikainen, and M. Maenpaa. “Multiresolution gray-scale and rotation invariant texture classification width local binary patterns”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(2002):971–987.

[12] Chen, Duan-Sheng, and Zheng-Kai Liu. “Generalized Haar-like features for fast face detection.” 2007 International Conference on Machine Learning and Cybernetics. Vol. 4. IEEE, 2007.

[13] P. Viola and M. Jones. “Robust real time object detection”. In IEEE ICCV Workshop on Statistical and Computational Theories of Vision, Vancouver, Canada, July 13 2001.

[14] Y. Freund and R.E. Schapire. “A decision-theoretic generalization of on-line learning and an application to boosting”. Journal of Computer and System Sciences, 55(1):119–139, August 1997.

[15] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. “Active shape models: Their training and application”. CVGIP: Image Understanding, 61:38–59, 1995.

[16] S. C. Yan, M.J.Li, H.J.Zhang, and Q.S.Cheng. “Ranking prior likelihood distributions of bayesian shape localization framework”. In ICCV 2003, pages 524–532, Mediterranean coast of France, October 2003.

[17] R. E. Schapire and Y. Singer. “Improved boosting algorithms using confidencerated predictions”. In Proceedings of the Eleventh Annual Conference on Computational Learning Theory, pages 80–91, 1998.

[18] Vapnik, Vladimir Naumovich, and Vlamimir Vapnik. Statistical learning theory. Vol. 1. New York: Wiley, 1998.

[19] Kreßel, Ulrich H-G. “Pairwise classification and support vector machines.”Advances in kernel methods. MIT Press, 1999.

[20] Milgram, Jonathan, Mohamed Cheriet, and Robert Sabourin. ““One Against One” or “One Against All”: Which One is Better for Handwriting Recognition with SVMs?” Tenth International Workshop on Frontiers in Handwriting Recognition. Suvisoft, 2006.