A Hierarchical Face Recognition Framework Based on Integrating Template Matching and SVM Classifier

Feng Xie Guangyou Xu Linmi Tao

Key Laboratory of Pervasive Computing (Tsinghua University), Ministry of Education
Department of Computer Science and Technology, Tsinghua University, Beijing 100084, P.R.China

Abstract

This paper presented an effective face recognition algorithm by integrates a template matching classifier and a SVM classifier, since SVM classifier is good for constructing complicated optimal separating hyper-plane, while the speed of template matching classifiers is very fast. Our hierarchical framework takes both advantages of these two methods. Specifically, for each input image, a template matching classifier is applied to select a small candidate set from a large database, then the selected set is forwarded to a trained “1 vs. 1” SVM classifier to get the recognition result. Compared with our original “1 vs. 1” SVM classifier, the hierarchical framework greatly improves the speed and keeps both the accuracy and generalization power of SVM classifier. Our initial experiments demon-strated that our method can be extended to construct a high accurate real time recognition system.

Keywords: Face recognition, Template Matching, Support Vector Machine.

1. Introduction

In the recent decade, the face confirmation and recognition technology have been a hot area of computer vision. Researchers have carried out lots of studies on face recognition and verification [1].

One of the most important aspects in the face recognition is the classifier design. In general, classifiers can be categorized into the following two classes. The first class include those classifiers based on the similarity measurements, such as the template matching [4], Principal Component Analysis (PCA) [10] and some variants of PCA [7,11], etc. These algorithms estimate similarity between two images. In the recognition process, all images in database are compared with the input image and the best match is selected as the desired result. The second class include classifiers based on the discriminant classification, such as Linear Discriminant Analysis (LDA) [5] and Support Vector Machine (SVM) [9], etc. These methods utilize the machine learning theory and search the optimal separating hyperplane in the feature space. No explicit similarity measurement has been defined in these algorithms.

These two classes of classifiers have their own advantage and drawback. The speed of template matching algorithms is very fast since the measure-ment of similarity has very low computational load, while the drawback is the rank 1 recognition rate drops very fast when the size of database grows up. However, Phillips and Moon’s results [2] show that rank *M* correct rate is comparatively stable if we concern the first *M* candidates. Contrary, SVM classifiers keep good performance in case of large database by constructing very complicated optimal separating hyperplane, which maximizes the margin resultantly. Inevitably, the classifiers have heavy computational load, which is intolerable in real time systems. For instance, the recognition speed of the "1 vs. 1" SVM classifier [8] is around 6 seconds per frame on a PC (PentiumIII 933, 256M RAM) for a database with 200 individuals.

To solve this problem, we have developed a hierarchical classification framework for fast and accurate recognition on large-scale databases. In the first stage a template matching classifier is used to choose a small subset of candidates. The reason of choosing template matching method is the template matching method is less sensitive to light condition and pose disturbances concerning the first *M* rank, while the SVM classifier is difficult to have good generalization ability over large various of light conditions and pose when the number of training data is relatively small. It is true that the sequential classifier will fail if the correct result is not within this subset, but our experiments result shows that the probability of correct result falls into the first 20 candidates is greater than 99% for a 200 people database. In the second stage we have applied "1 vs. 1"

SVM classifier on this small subset to get the final output. As a result, a high recognition rate is reached with high speed since the "1 vs. 1" SVM classifier works well on a small database generated from the template matching classfier.

The rest of this paper is organized as follows. Section 2 describes the details of our new hierarchical framework. Our primary experimental results are demonstrated in Section 3. Finally, we conclude in Section 4.

2. Hierarchical Face Recognition Framework

Figure 1 demonstrates the framework of the improved recognition algorithm. First we outline our original “1 vs. 1” SVM classifier, then describe the template matching algorithm which select the candidate set.

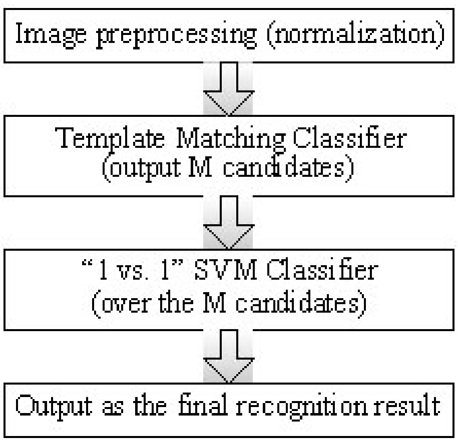

Fig. 1. The framework of improved recognition Algorithm.

2.1. “1 vs. 1” SVM classifier

Support Vector Machines show superior performance in binary-classification problems and are capable to solve face verification problem. But face recognition is a multi-classification problem so we setup two strategies to adapt SVM classifier suitable for face recognition: "1 vs. 1" and "1 vs. others".

With "1 vs. 1" strategy, a SVM classifier is trained for each two clients in the database. Totally *N*(N-1)/2* classifiers have gotten for a database with ***N*** clients. When the *(N+1)-th* client is added, the SVM classifiers between this client and other ***N*** clients are the only ones to be trained.

In recognition stage, we use competitive strategy. As shown in Fig.2, we consider a database with *N* clients. For a testing input face, the *N* clients are firstly divided into *N/2* pairs whose corresponding SVM classifiers are used to choose the winners that survive for next round competition. Doing this way repetitively, finally only one will be left as the winner to be the claim.

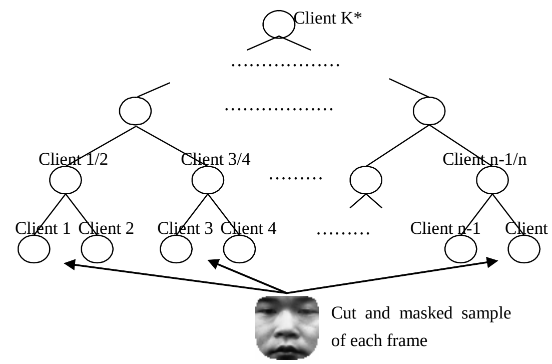

Fig. 2. “1 vs. 1” promoting procedure.

The second strategy, "1 vs. others" vote strategy, almost totally come from face verification. Consider a database with ***N*** clients. Using "1 vs. others" strategy, a SVM classifier is trained for each client in that database, which resulting totally ***N*** SVM classifiers for the database.

Although "1 vs. others" vote strategy is very simple and easy to apply, it has several defects for face recognition. First, when the database changed, SVM classifiers for all clients should be retrained, which will take a very long time. second, the number of SVs of the classifier will go larger as the number of clients in the database goes larger; this will make the recognition procedure very slow. Hsu and Lin[6] got the similar conclusion that the "1 vs. 1" strategy is more practical than "1 vs. other" strategy. Therefore we only adopt the "1 vs. 1" strategy in this paper.

2.2. Template Matching Classifier

However, even adopted "1 vs. 1" strategy there need ***N-1*** times comparisons to get the final recognition result. When the size of face database is big, the speed of computation is still slow. So we want to use other algorithm to proceed a rough recognition, getting the first *M* (*M*<<*N*) candidates, and then to use the "1 vs. 1" strategy over this *M* candidates to get the final result. The algorithm must satisfies two criteria: first, the correct result must almost always be within the subset of candidates, otherwise consecutive SVM classifier can't produce the right result; second, the speed must be faster than SVM classifier such that we can shorten the whole time cost for recognition.

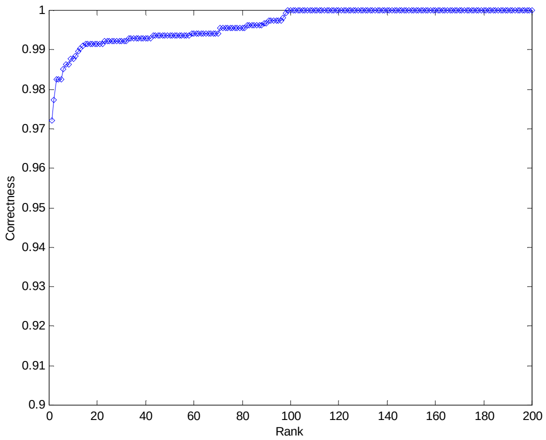

Fig. 3. Cumulative Match Characteristic (CMC).

The template matching algorithm satisfied these two criteria perfectly. In our 500 people database, we randomly select 200 people, and the cumulative match characteristic (CMC) is shown in Fig.3. The horizontal axis of the graph is rank, while the vertical axis is the recognition rate.

This means the probability of correct result falls into first 10 candidates is 98.77%, and the probability of correct result falls into first 20 candidates is 99.16%. So it can satisfy the first criteria. The contrast of computation speed is as follows (CPU PentiumIII 933, Memory 256 MB, 200 people THLIB database):

Time cost of Template Matching algorithm	Time cost of SVM classifier
0.49s	5.98s

Compared with the SVM classifier, the speed of template matching algorithm is about 12 times quick. So it can also satisfy the second criteria.

3. Experiments

We carry out the test on two databases, one is our THLIB face database, another is the ORL Database of Faces [3]. In THLIB database every people have 11 images, 9 of them are used as training set, the other two as *gallery*. We test over 50 people, 100 people and 200 people database. In ORL database every people have 10 images, we use 5 images as training set and others as *gallery*. There are $C_{10}^{5} = 252$ different kind of selection, and we test over all these selection, get the average score. The illumination and pose of all images have been normalized in preprocessing step.

First we take a look at the relation between the size of candidates subset and correct recognition rate.

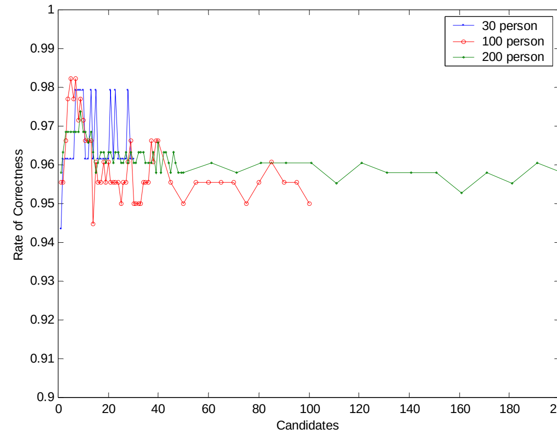

Fig. 4. Three tests over THLIB database, 30, 100, 200 person respectively, the correct recognition rate vs. the size of candidates subset.

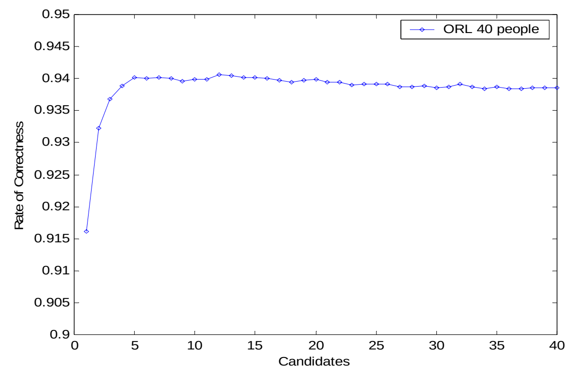

Fig. 5. Average correct rate of 252 tests over ORL database.

In Fig. 4 and Fig. 5, when the size of subset is 1, the proposed algorithm degrades to template matching algorithm and the SVM classifier has no effect; when the size of subset is the same as database size, the proposed algorithm degrades to “1 vs. 1”' SVM algorithm and the template matching algorithm has no effect. When the size of subset is between 1 and database size, they all have effect on final recognition result. From Fig. 4 and Fig. 5 we can see that when the size of subset is 10 to 15, the correct recognition rate is higher than either the single SVM classifier or template matching algorithm.

Generally, changing the kernel function of SVM classifier has little effect on correct rate, our result also confirmed it (see Fig. 6).

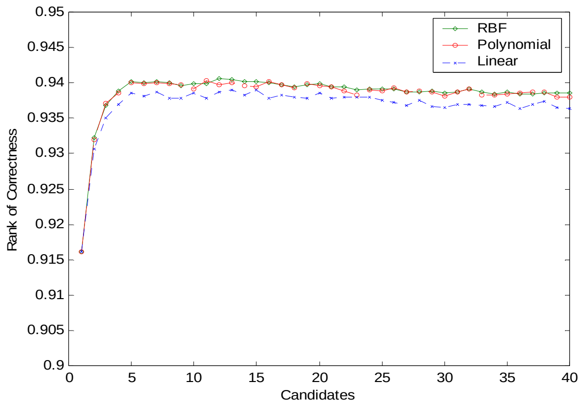

Fig. 6. Correct rate over Different type of kernel function of SVM (RBF, Polynomial and Linear).

The correct recognition rate, compared with a few common used classification algorithms:

	THLIB 200 people database	ORL 40 people database
Template Matching	95.79%	91.61%
PCA	95.26%	92.43%
LDA	96.32%	93.72%
SVM	95.79%	93.86%
Template Matching+ SVM	**96.84%**	**93.99%**

Finally, the recognition speed is also improved as expected (CPU PentiumIII 933, Memory 256 MB, 200 people THLIB database, the size of candidates subset is 10):

Time cost of our new method	Time cost of SVM classifier
0.77s	5.98s

4. Summary

This paper has presented a hierarchical face recognition framework which integrates two classical methods ie. template matching and SVM classifier. As a result, a high recognition rate is reached with high speed since the "1 vs. 1" SVM classifier works well on a small database generated from the template matching classifier. For instance, the recognition speed of the original "1 vs. 1" SVM classifier is around 6 seconds per frame on a PentiumIII 933 PC for a database with 200 individuals, while the new framework only need 0.77s. So it could be applied to implement real time face recognition system.

The relationship between the size of subset and the correct rate of recognition is still uncertain. Up till now no principle has been found to determine the size of subset. From the experimental result, we set it 10 to 15 and get good performance.

5. References

[1] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," *ACM Computing Survey*, Vol. 35, No 4, pp. 399-458, 2003.

[2] P. J. Phillips, H. Moon, P. Rauss, and S. A. Rizvi, "The FERET Evaluation Methodology for Face-Recognition Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, pp. 1090-1104, Oct 2000.

[3] F.S. Samaria and A.C. Harter, "Parameterisation of a stochastic model for human face identification," *Proceedings of the 2nd IEEE workshop on Applications of Computer Vision*, Sarasota, Florida, 1994.

[4] R. Brunelli, T. Poggio, "Face Recognition: Features versus Templates," *IEEE Transcations on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 10, pp. 1042-1052, 1993.

[5] R.A. Fisher, "The use of multiple measures in taxonomic problems," Ann. Eugenics, vol. 7, pp. 179-188, 1936.

[6] C.-W. Hsu, C.-J. Lin, "A comparison of methods for multi-class support vector machines," IEEE Transactions on Neural Networks, Vol. 13, pp. 415-425, 2002.

[7] M. Tipping and C.M. Bishop, "Probabilistic principal component analysis," *Journal of the Royal Statistical Society, Series B.*, vol. 61, pp. 611-622, 1999.

[8] L. Zhuang, H.Z. Ai, G.Y. Xu, "Training Support Vector Machines for Video Based Face Recognition," *Proceedings of SPIE-The International Society for Optical Engineering*, v 4875, n 2, pp. 737-743, 2002.

[9] B. E. Boser, I. Guyon, and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers," *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, pp. 144-152, 1992.

[10] M. Turk and A. Pentland, "Eigenfaces for Recognition," *J. Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-96, 1991.

[11] B. Scholkopf, A.J. Smola, and K.R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, Vol. 10, pp. 1299--1319, 1998.