

# Multi-view Face Recognition with Min-Max Modular SVMs

Zhi-Gang Fan and Bao-Liang Lu\*

Department of Computer Science and Engineering,  
Shanghai Jiao Tong University, 1954 Hua Shan Road,  
Shanghai 200030, China  
zgfan@sjtu.edu.cn, blu@cs.sjtu.edu.cn

**Abstract.** Through task decomposition and module combination, min-max modular support vector machines ( $M^3$ -SVMs) can be successfully used for difficult pattern classification task.  $M^3$ -SVMs divide the training data set of the original problem to several sub-sets, and combine them to a series of sub-problems which can be trained more effectively. In this paper, we explore the use of  $M^3$ -SVMs in multi-view face recognition. Using  $M^3$ -SVMs, we can decompose the whole complicated problem of multi-view face recognition into several simple sub-problems. The experimental results show that  $M^3$ -SVMs can be successfully used for multi-view face recognition and make the classification more accurate.

## 1 Introduction

Support vector machines (SVMs)[1] have been successfully applied to various pattern classification problems. However, SVMs require to solve a quadratic optimization problem and need training time that are at least quadratic to the number of training samples. Therefore, many large-scale problems are too hard to be solved by using traditional SVMs. To solve large-scale multi-class problems, we have proposed a min-max modular support vector machines ( $M^3$ -SVMs) in our previous work[3].

In this paper, we explore the use of  $M^3$ -SVMs in multi-view face recognition. Multi-view face recognition is a more challenging task than frontal view face recognition. Face recognition techniques have been developed over the past few decades. But many of those existing face recognition techniques, such as Eigenfaces and Fisherfaces [4,5], are only effective for frontal view faces. The difficulties of multi-view face recognition is obvious because of the complicated nonlinear manifolds existing in the data space. Using  $M^3$ -SVMs, we can decompose the whole complicated problem of multi-view face recognition into several simple sub-problems. Every individual sub-problem becomes less complicated than the original problem and it can be solved effectively.

---

\* To whom correspondence should be addressed. This work was supported in part by the National Natural Science Foundation of China via the grants NSFC 60375022 and NSFC 60473040.

## 2 Min-Max Modular Support Vector Machines

Before using M<sup>3</sup>-SVMs, for a  $K$ -class problem, we should divide the  $K$ -class problem into  $K(K - 1)/2$  two-class sub-problems according to one-against-one strategy[2,8]. The work procedure of M<sup>3</sup>-SVMs consists of three steps: task decomposition, SVMs training and module combination. First, every two-class problem is decomposed into smaller two-class problems. Then, every small two-class SVMs is trained. At last, all trained individual modules are integrated to get a solution to the original problem.

Let  $\mathcal{X}^+$  and  $\mathcal{X}^-$  be the given positive and negative training data set for a two-class problem  $\mathcal{T}$ ,

$$\mathcal{X}^+ = \{(x_i^+, +1)\}_{i=1}^{l^+}, \quad \mathcal{X}^- = \{(x_i^-, -1)\}_{i=1}^{l^-} \tag{1}$$

where  $x_i \in \mathbf{R}^n$  is the input vector, and  $l^+$  and  $l^-$  are the total number of positive training data and negative training data of the two-class problem, respectively.

According to [3],  $\mathcal{X}^+$  and  $\mathcal{X}^-$  can be partitioned into  $N^+$  and  $N^-$  subsets respectively,

$$\mathcal{X}_j^+ = \{(x_i^{+j}, +1)\}_{i=1}^{l_j^+}, \quad \text{for } j = 1, \dots, N^+ \tag{2}$$

$$\mathcal{X}_j^- = \{(x_i^{-j}, -1)\}_{i=1}^{l_j^-}, \quad \text{for } j = 1, \dots, N^- \tag{3}$$

where  $\cup_{j=1}^{N^+} \mathcal{X}_j^+ = \mathcal{X}^+$ ,  $1 \leq N^+ \leq l^+$ , and  $\cup_{j=1}^{N^-} \mathcal{X}_j^- = \mathcal{X}^-$ ,  $1 \leq N^- \leq l^-$ .

After decomposing the training data sets  $\mathcal{X}^+$  and  $\mathcal{X}^-$ , the original two-class problem  $\mathcal{T}$  is divided into  $N^+ \times N^-$  relatively smaller and more balanced two-class sub-problems  $\mathcal{T}^{(i,j)}$  as follows:

$$(\mathcal{T}^{(i,j)})^+ = \mathcal{X}_i^+, \quad (\mathcal{T}^{(i,j)})^- = \mathcal{X}_j^- \tag{4}$$

where  $(\mathcal{T}^{(i,j)})^+$  and  $(\mathcal{T}^{(i,j)})^-$  denote the positive training data set and the negative training data set for subproblem  $\mathcal{T}^{(i,j)}$ , respectively.

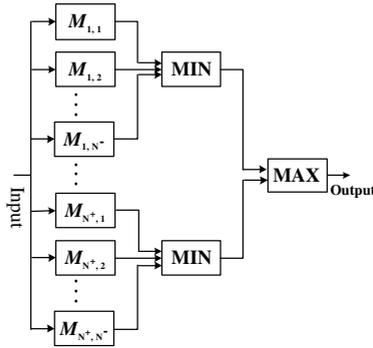
In the learning phase, all the two-class sub-problems are independent from each other and can be efficiently learned in a massively parallel way. After training, all the individual SVMs are integrated into a M<sup>3</sup>-SVM with MIN and MAX units according to the two combination principles, namely the minimization principle and the maximization principle [2,3].

According to the minimization and maximization principles, the  $N^+ \times N^-$  smaller SVMs are integrated into a M<sup>3</sup>-SVM with  $N^+$  MIN units and one MAX unit as following equations (5, 6). Fig.1 shows the structure of M<sup>3</sup>-SVMs.

$$M^i(x) = \min_{j=1}^{N^-} M^{(i,j)}(x) \quad \text{for } i = 1, \dots, N^+ \tag{5}$$

and

$$M(x) = \max_{i=1}^{N^+} M^i(x) \tag{6}$$



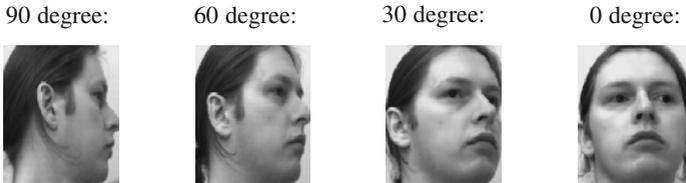
**Fig. 1.** A  $M^3$ -SVM consisting of  $N^+ \times N^-$  individual SVMs,  $N^+$  MIN units, and one MAX unit

where  $M^{i,j}(x)$  denotes the transfer function of the trained SVM corresponding to the two-class subproblem  $\mathcal{T}^{(i,j)}$ , and  $M^i(x)$  denotes the transfer function of a combination of  $N^-$  SVMs integrated by the MIN unit.

### 3 Experiments

We use the UMIST database [6], a multi-view face database consisting of 575 gray-scale images of 20 subjects, each covering a wide range of poses from profile to frontal views. Figure.2 depicts some sample images of one subject in the UMIST database. The overall database is partitioned into two subsets: the training set and test set. The training set is composed of 240 images: 12 images per person are carefully chosen according to face poses. The remaining 335 images are used to form the test set. All input images are of size  $112 \times 92$ . We have used feature selection method to reduce the dimensionality of feature space and the feature selection method we have used is described in detail in [7]. All of the experiments were performed on a 3.0GHz Pentium 4 PC with 1.0 GB RAM.

As shown in Fig.2, using task decomposition principle of  $M^3$ -SVMs, we have divided all face images in each face class into 4 subsets according to the face poses. An RBF kernel for SVMs is used. The parameter  $C$  is set to 10000 and an optimal  $\sigma$  is used. The experimental results are shown in Table 1 and we can see that  $M^3$ -SVMs is more accurate than traditional SVMs.



**Fig. 2.** All images in each class are divided into 4 subsets according to the face poses

**Table 1.** Test results on UMIST face database

Methods	No. features	$\sigma$	Training time (s)		Test time (s)	Correct rate (%)
			Parallel	Serial		
SVMs (rbf kernel)	300	30	0.862	13.588	1.522	92.8358
	200	25	0.748	12.654	0.976	92.2388
	150	25	0.703	11.865	0.757	90.1493
	100	20	0.685	11.269	0.478	82.3881
M <sup>3</sup> -SVMs (rbf kernel)	300	20	0.531	15.273	1.647	93.1343
	200	15	0.447	13.413	1.215	92.5373
	150	10	0.386	12.587	0.873	91.3433
	100	10	0.359	12.165	0.526	83.8806

## 4 Conclusion and Future Work

Through task decomposition and module combination, we have applied M<sup>3</sup>-SVMs to multi-view face recognition. Comparing to traditional SVMs, our experiments show that M<sup>3</sup>-SVMs can improve the accuracy of multi-view face recognition. In the future work, we will enhance the task decomposition method.

## References

1. Vapnik, V.N.: The Nature of Statistical Learning Theory. Springer-Verlag, New York (2000)
2. Lu, B.L., Ito, M.: Task decomposition and module combination based on class relations: a modular neural network for pattern classification. *IEEE Transactions on Neural Networks*, vol.10, (1999) 1244 -1256
3. Lu, B.L., Wang, K.A., Utiyama, M., Isahara, H.: A part-versus-part method for massively parallel training of support vector machines. In: *Proceedings of IJCNN'04, Budapest, July 25-29(2004)*
4. Turk, M., Pentland, A.: Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, vol. 3, no. 1, (1991) 71-86
5. Belhumeur, P., Hespanha, J., Kiregeman, D.: Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, (1997) 711-720.
6. Graham, D.B., Allinson, N.M.: Characterizing virtual eigensignatures for general purpose face recognition. In: *Face Recognition: From Theory to Applications, NATO ASI Series F, Computer and Systems Sciences*, vol. 163, (1998) 446-456.
7. Fan, Z.G., Lu, B.L.: Fast Recognition of Multi-view Faces with Feature Selection. submitted to *ICCV05, Beijing (2005)*
8. Hsu, C., Lin, C.: A Comparison of Methods for Multiclass Support Vector Machines. *IEEE Trans. Neural Networks*, vol. 13, no.2, (2002) 415-425.