PERSONAL AUTHENTICATION BASED ON GENERALIZED SYMMETRIC MAX MINIMAL DISTANCE IN SUBSPACE

Wende Zhang and Tsuhan Chen
Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
{wendez, tsuhan}@andrew.cmu.edu

ABSTRACT

We introduce an improved classification algorithm based on the concept of Symmetric Maximized Minimal distance in Subspace (SMMS). Given the training data of authentic samples and imposter samples in the feature space, our previous approach, SMMS, tried to identify a subspace in which all the authentic samples were projected onto the origin and all the imposter samples were far away from the origin. The optimality of the subspace was determined by maximizing the minimal distance between the origin and the imposter samples in the subspace. The Generalized SMMS relaxes the constraint of fitting all the authentic samples to the origin in the subspace to achieve the optimality and considers the optimal direction of the linear support-vector machines (SVM) as a feasible solution in our optimization procedure to guarantee that our result is no worse than the linear SVM. We present a procedure to achieve such optimality and to identify the subspace and the decision boundary. Once the subspace is trained, the verification procedure is simple since we only need to project the test sample onto the subspace and compare it against the decision boundary. Using face authentication as an example, we show that the proposed algorithm outperforms the linear classifier based on SMMS and SVM. The proposed algorithm also applies to multimodal feature spaces. The features can come from any modalities, such as face images, voices, fingerprints, etc.

1. INTRODUCTION

As the biometric applications become more popular, the biometric classification algorithms will be applied to hand-held devices. Due to the size and power constraints of these devices, the computational complexity of the verification procedure will be the main consideration for choosing the right algorithms.

The subspace approach is suitable for biometric authentication applications on power-constrained devices because of its simplicity in the verification procedure. We only need to project the test sample, which can be the features from multi-modalities, such as face images, voices, fingerprints, etc, onto the subspace. Most subspace classification algorithms determine the subspace by optimizing certain global measures such as the mean squared error (MSE) or the variance ratio. For example, in principal component analysis (PCA) [1], the subspace is determined by minimizing the reconstruction errors of samples in the subspace using the MSE criterion. In Fisher's linear discriminant analysis (LDA) [2], the subspace is defined by maximizing the ratio between the intra-class variance and the inter-class variance in the subspace. The applications of these algorithms can be found in [3][4][5] for biometric recognition and authentication. These algorithms based on global optimization provide good performance when the data distribution is close to Gaussian.

Other classification algorithms determine their decision boundaries by supporting samples [6]. Supporting samples are typically training samples near the decision boundary. Support vector machines (SVM) [7][8][9][10][11] optimize the supporting samples by maximizing the gap between the authentic samples and the imposter samples regardless of the data distribution. This "maxmin" approach outperforms the global optimization approach, when the data distribution is not Gaussian as shown by several face authentication and recognition applications [12][13][14].

A linear SVM [9] maximizes the separation gap between the authentic samples ($\mathbf{x}_{i(authentic)}$, hexagon pattern) and the imposter samples ($\mathbf{y}_{j(imposter)}$, triangle pattern) as shown in Figure 1. Alternatively, the optimal parameters (**w,** *b*) are optimized by

$$\max_{\mathbf{w},b}\{\min_{j}(\mathbf{w}^T\mathbf{y}_{j(imposter)}+b)-\max_{i}(\mathbf{w}^T\mathbf{x}_{i(authentic)}+b)\},$$

$$\mathbf{w}^T\mathbf{w}=1$$

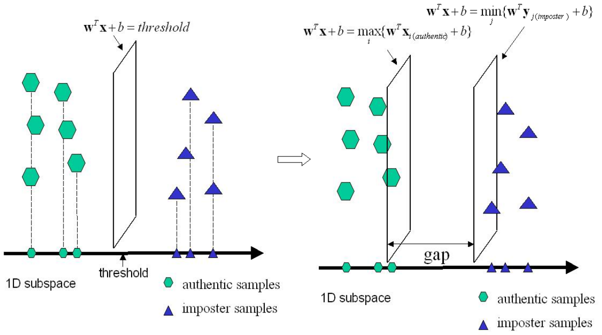

Figure 1. Linear SVM

The computation requirement for verification is low since only one inner product between the input vector **x** and the projection vector **w** is needed. We use the following criterion to decide whether the input vector **x** is authentic or not.

If $f(x)=\mathbf{w}^T\mathbf{x}+b\le threshold$, $\mathbf{x}\rightarrow$ authentic sample

This is equivalent to projecting **x** onto a one-dimensional subspace and comparing the result with the threshold as shown in Figure 1. The threshold value can be simply set to zero or chosen from the Receiver Operating Characteristic (ROC) curve to meet the false accept rate (FAR) vs. false reject rate (FRR) requirement.

A linear SVM may fail when the authentic samples (hexagon patterns) are surrounded by the imposter samples as shown in Figure 2. It is not possible to find a linear plane to separate the authentic samples from all imposter samples in this case. One way to solve this problem is to seek a separation plane by mapping the training samples to a higher dimensional space [9]. After mapping, the samples may be well separated. However, there are two potential problems in this approach. First, the advantage of simplicity is lost. When the samples are mapped to a higher dimensional space, we have to keep many support vectors to describe the separation boundary instead of only one projection vector **w**. In addition, the over fitting problem may be encountered, that is, the samples may be too few to train SVM in a high dimensional space.

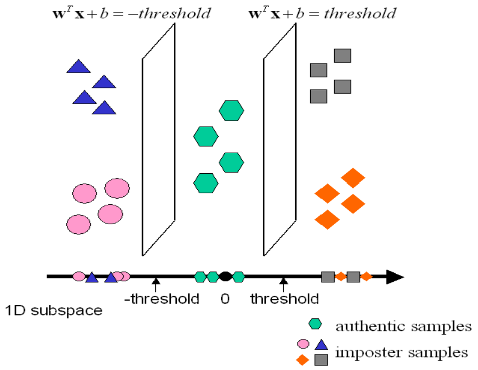

Figure 2. Samples not separable by 2 hyper-planes

In our previous work--Symmetric Maximized Minimal distance in Subspace (SMMS) [15], the authentic samples are projected onto the origin of the subspace, and the imposter samples are surround the authentic samples, hence "symmetric". The algorithm identifies the subspace and decision boundary by maximizing the minimal distance from imposter samples to the origin of the subspace. SMMS additionally requires that all the authentic samples will be projected onto the origin of the subspace. Therefore, the optimal subspace of SMMS is a sub-optimal solution to the problem of finding the max minimal distance between authentic and imposter samples in a subspace.

In this paper, we propose a new approach called Generalized Symmetric Maximized Minimal distance in Subspace (G-SMMS) to perform classification. By allowing the authentic samples to scatter around the origin in the subspace, hence "generalized", this algorithm tries to find a subspace in which the authentic samples are clustered together and the imposter samples are far away from the authentic samples. For example, in a one-dimensional subspace, the authentic samples are clustered around the origin in the subspace as shown in Figure 2. The following criterion is used to decide whether the input vector **x** is authentic or not.

$$\text{If } f(\mathbf{x})=\left|\mathbf{a}^T\mathbf{x}+b\right|\le threshold\ ,\ \mathbf{x}\rightarrow \text{authentic sample} \tag{1}$$

The samples that are not linearly separable by a linear SVM can now be well separated by G-SMMS in Figure 2 without constraint on the authentic samples.

The paper is organized as follows. In section 2, we introduce the proposed classification algorithm based on G-SMMS. In Section 3, we describe the face database used in the experiment and compare the authentication performance of G-SMMS with the linear SVM and SMMS algorithms. The conclusions are given in Section 4.

2. GENERALIZED SYMMETRIC MAXIMIZED MINIMAL DISTANCE IN SUBSPACE (G-SMMS)

We will introduce the one-dimensional G-SMMS algorithm in details and then extend the algorithm to higher dimensions.

2.1 One-dimensional Generalized Symmetric Maximized Minimal distance in Subspace (G-SMMS-1D)

To train the G-SMMS classifier, given M authentic samples, $\mathbf{x}_i$, and N imposter samples, $\mathbf{y}_j$, we try to maximize the minimal distance between the authentic samples and the imposter samples.

$$d_{G-SMMS-1D}=\max_{\mathbf{a},b}\{\min_j\left|\mathbf{a}^T\mathbf{y}_j+b\right|-\max_i\left|\mathbf{a}^T\mathbf{x}_i+b\right|\}$$
$$j=1,2,\ \dots,N;\ i=1,2,\ \dots,M;\ \mathbf{a}^T\mathbf{a}=1 \tag{2}$$

We let $b=-\mathbf{a}^T mean(\mathbf{x}_i)$, $\mathbf{p}_i=\mathbf{x}_i-mean(\mathbf{x}_i)$ and $\mathbf{q}_j=\mathbf{y}_j-mean(\mathbf{x}_i)$. The choice of b we made here is to make $\max_i\left|\mathbf{a}^T\mathbf{x}_i+b\right|$ as small as possible. It is not guaranteed for the overall $d_{G\text{-}SMMS\text{-}1D}$ to be optimal, but it is a reasonably and intuitively good choice. Then, we can optimize the minimal difference between the authentic samples and the imposter samples as shown below:

$$d_{G-SMMS-1D}=\max_{\mathbf{a}}\{V(\mathbf{a})\}=\max_{\mathbf{a}}\{\min_j\left|\mathbf{a}^T\mathbf{q}_j\right|-\max_i\left|\mathbf{a}^T\mathbf{p}_i\right|\} \tag{3}$$

To solve the "maxmin" problem (3) directly is difficult, since it is a nonlinear programming problem. One way to approximate this is to convert it into a quadratic form as follows.

$$d'_{G-SMMS-1D}=\max_{\mathbf{a}}\{\min_j\left|\mathbf{a}^T\mathbf{q}_j\right|^2-\max_i\left|\mathbf{a}^T\mathbf{p}_i\right|^2\} \tag{4}$$

The way we solve (4) is by assigning a weight to each quadratic term properly to approximate the min and max operators in (4).

We assign a weight $s_{j(q)}$ to each value $Q_j(\mathbf{a})=\left|\mathbf{a}^T\mathbf{q}_j\right|^2$ and a weight $s_{i(p)}$ to each value $P_i(\mathbf{a})=\left|\mathbf{a}^T\mathbf{p}_i\right|^2$. We constrain the sum of all the weights $s_{j(q)}$ and $s_{i(p)}$ to be 1 separately.

If we assign the non-zero weight to minimum value of $Q_j(\mathbf{a})$ and the zero weights to all the others, the minimum of $Q_j(\mathbf{a})$ will be equal to the weighted sum in equation (5).

$$\min_j Q_j(\mathbf{a})\approx\sum_j s_{j(q)}Q_j(\mathbf{a})=m_{w(q)} \tag{5}$$

Similarly, if we assign the non-zero weight to maximum value of $P_i(\mathbf{a})$ and the zero weights to all the others, the maximum of $P_i(\mathbf{a})$ will be equal to the weighted sum as follows.

$$\max_i P_i(\mathbf{a})\approx\sum_i s_{i(p)}P_i(\mathbf{a})=m_{w(p)} \tag{6}$$

To accomplish the above, the weights are adjusted iteratively to make the weighted mean $m_{w(q)}$ close to the minimum value of $Q_j(\mathbf{a})$ and the weighted mean $m_{w(p)}$ close to the maximum value of $P_i(\mathbf{a})$. To achieve this point, the weights $s_{j(q)}$ are increased for small $Q_j(\mathbf{a})$, while the other weights $s_{j(q)}$ are decreased. The weights $s_{i(p)}$ are increased for large $P_i(\mathbf{a})$, while the other weights $s_{i(p)}$ are decreased. Now given weights, we find the **a** by optimizing the following quadratic sum criterion

$$\mathbf{a}=\arg\max_{\mathbf{a}}\{\sum_j s_{j(q)}Q_j(\mathbf{a})-\sum_i s_{i(p)}P_i(\mathbf{a})\} \tag{7}$$

instead of the original criterion in (3).

Since both **a** and weights are unknown, we try to solve them iteratively. First, we initialize weights for $s_{j(q)}$ and $s_{i(p)}$ equally. Given the updated weights, we try to solve **a** by optimizing (7). Then given the current projection direction **a** along which we project $\mathbf{p}_i$ to $P'_i = |\mathbf{a}^T\mathbf{q}_j|$ or $\mathbf{q}_j$ to $Q'_j = |\mathbf{a}^T\mathbf{q}_j|$, the weights are updated by comparing the distance between the corresponding P'_i or Q'_j and the origin. After the maximum number of iterations, we choose **a** that provides the maximum value of $V(\mathbf{a})$ as indicated in (3) among all the iterations as the optimal solution for **a**. It is necessary to check with because it specifies the optimal solution while iterations for search the solution.

2.2 High-dimensional Generalized Symmetric Maximized Minimal distance in Subspace (G-SMMS-MD)

G-SMMS-1D can be extended to a higher-dimensional subspace. If we extend the subspace with orthonormal bases $\mathbf{L} = [\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_R]$ for verification, with the G-SMMS algorithm, the following criterion is used to decide whether the input vector, **x,** is authentic.

$$f(\mathbf{x}) = \sqrt{\sum_{n=1}^{R} \left|\mathbf{a}_n{}^T\mathbf{x} + b_n\right|^2} \le threshold \qquad (8)$$

Here we try to maximize the minimal distance between the authentic samples and the imposter samples in a subspace as follows.

$$d_{G-SMMS-MD} = \max_{\mathbf{L}}\{\min_j \sqrt{\sum_{n=1}^{R}\left|\mathbf{a}_n{}^T\mathbf{y}_j + b_n\right|^2} - \max_i \sqrt{\sum_{n=1}^{R}\left|\mathbf{a}_n{}^T\mathbf{x}_i + b_n\right|^2}\} \qquad (9)$$

where $b_n = -\mathbf{a}_n{}^T mean(x_i)$.

We can also solve $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_R$ by optimizing the (9) using the same procedure as G-SMMS-1D expect for redefining $Q_j = \sum_{n=1}^{R}\left|\mathbf{a}_n{}^T\mathbf{q}_j\right|^2$ and $P_i = \sum_{n=1}^{R}\left|\mathbf{a}_n{}^T\mathbf{p}_i\right|^2$ in multiple-dimensional G-SMMS.

After the iteration, we will get the optimal solution $\mathbf{L}_{optimal} = [\mathbf{a}_{1\,optimal}, \mathbf{a}_{2optimal}, ..., \mathbf{a}_{Roptimal}]$.

2.3 Using Linear SVM to improve G-SMMS-MD

For some data, the authentic and the imposter samples are already well separated by using linear SVM. For such data, the optimization we apply in G-SMMS may not provide the best solution. However, the G-SMMS formulation does incorporate the linear SVM as a special case, as we will explain.

If we perform the linear SVM on the data, we can get the max minimal distance d_{SVM} between authentic and imposter samples along optimal projection direction **w**. If we project the data to an R-dimensional subspace, which contain the projection direction **w**, we can show that the max minimal distance $d_{G\text{-}SMMS\text{-}MD}$ between authentic and imposter samples will be no smaller than the d_{SVM}.

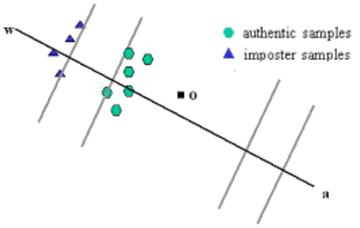

Figure 3. Linear SVM and G-SMMS-1D

For G-SMMS-1D, we let the unit projection vector $\mathbf{a} = \mathbf{w}/|\mathbf{w}|$, the origin in the original space $\mathbf{o} = mean(\mathbf{x}_i) + c * [mean(\mathbf{x}_i) - mean(\mathbf{y}_j)]$ and $b = -\mathbf{a}^T\mathbf{o}$, where the c is such a big coefficient that let the all the projected samples along the direction **a** have the negative values as shown in Figure 3. In this case, the minimal distance between authentic and imposter samples is the same for the Linear SVM and G-SMMS-1D.

For G-SMMS-MD, we let the first unit projection vector $\mathbf{a}_1 = \mathbf{w}/|\mathbf{w}|$, $\mathbf{o}_1 = mean(\mathbf{x}_i) + c * [mean(\mathbf{x}_i) - mean(\mathbf{y}_j)]$ and $b_1 = -\mathbf{a}_1{}^T\mathbf{o}_1$. The $\mathbf{a}_2, \mathbf{a}_3, \ldots, \mathbf{a}_M$ are determined by maximizing $\sum_{m=2}^{M}\left(\sum_j\left|\mathbf{a}_m{}^T\mathbf{q}_j\right|^2 - \sum_i\left|\mathbf{a}_m{}^T\mathbf{p}_i\right|^2\right)$ in the null space of $\mathbf{a}_1$. And $b_2, b_3, \ldots, b_M$ are determined by $b_n = -\mathbf{a}_n{}^T mean(\mathbf{x}_i)$. In this case, the $d_{G\text{-}SMMS\text{-}MD}$ is similar to the d_{SVM} with the big positive number c. If c goes to the positive infinitive, it will push the center **o'** of the hyper-circle in R-dimensional subspace far away from the samples along the direction $\mathbf{a}_1$, then the curves through the authentic and imposter samples in the Figure 4 will become two parallel lines and orthogonal to $\mathbf{a}_1$. Now $d_{G\text{-}SMMS\text{-}MD} = d_{SVM}$.

We will also consider the special SVM related solution here to evaluate our criterion function (9). So the final optimal $d_{G\text{-}SMMS\text{-}MD}$ will be no smaller than the d_{SVM}. The final solution will be the subspace $<\mathbf{a}_n, b_n>$ corresponding to the max minimal distance between authentic and imposter samples.

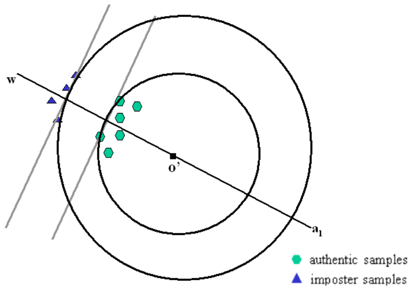

Figure 4. Linear SVM and G-SMMS-MD (M>2)

3. EXPERIMENT

We test our algorithm on a pose variation subset of the PIE face database [16] with 68 subjects. This is a very challenging database for face authentication because there are lots of pose variations and many subjects. We use 5 images for training and 4 images for testing for each subject. All the images are 64×64 grayscale cropped images. In Figure 5, the training and test face images from 3 subjects in this database are shown in different rows.

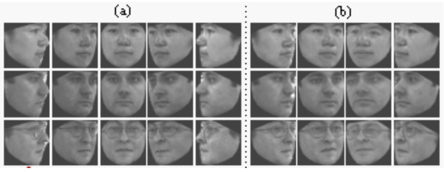

Figure 5. Sample training (a) images and test (b) images from PIE database

We use all the training images (5*68) to construct a PCA subspace. We do the experiment on first 100 PCA coefficients with 1D-7D subspaces for the proposed G-SMMS-MD as shown in Figure 6. We compare the resulting equal error rate (EER) on average with a total of 68 subjects among the SMMS, G-SMMS-

MD and the linear SVM as shown in Table 1 and Figure 6. The EER of the linear SVM is 17.54%, while any of the result from G-SMMS-MD is no worse than the linear SVM. Furthermore, the performance of G-SMMS-MD is much better than SMMS-MD as shown in Table 1.

The G-SMMS-1D performs the same as the linear SVM in this database because the data is already separable for each subject by the linear SVM and the G-SMMS-1D cannot find another subspace which provides a bigger gap between authentic and imposter samples.

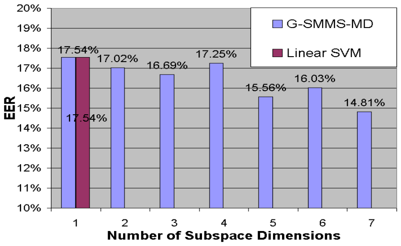

Figure 6. The comparison between the G-SMMS-MD and the Linear SVM on PIE database

	EER	Storage unit	Verification time unit
SVM-linear	17.5%	1	1
SMMS-1D	36%	1	1
G-SMMS-1D	17.5%	1	1
SMMS-3D	27%	3	3
G-SMMS-3D	16.7%	3	3

Table 1. The comparison among SMMS, G-SMMS and SVM on PIE database

When we increase the dimensions of G-SMMS, we lose the simplicity, too. However, compared with other high dimensional SVM algorithms such as the SVM-polynomial (degree=2) [9], the G-SMMS-7D is still simpler and better as shown in Figure 6. The SVM-polynomial (degree=2) requires 40-50 support vectors for each subject in this database.

	EER	Storage unit	Verification time unit
G-SMMS-7D	14.8%	7	7
SVM-polynomial (degree=2)	16.2%	40-50	40-50

Table 2. The comparison between G-SMMS-7D and SVM-polynomial (degree=2)

4. CONCLUSIONS

In this paper, we propose a new algorithm based on the concept of Generalized Symmetric Maximized Minimal distance in Subspace (G-SMMS). Using face authentication as an example, we show that the proposed algorithm is simple in the verification stage and outperforms the linear SVM and SMMS on the PIE face database. The proposed algorithm also applies to multimodal feature spaces. The features can come from any modalities, such as face images, voices, fingerprints, etc. This paper shows that the subspace classification with symmetric maximizing minimal criterion is a promising way to enhance the performance of subspace-based authentication algorithms. Along this direction, our future work is to introduce some slack variables [9] similar to SVM to improve the performance.

5. ACKNOWLEDGEMENTS

Work supported in part by General Motors Corporation.

6. REFERENCES

[1] I.T. Jolliffe, *Principal Component Analysis*, Spring-Verlag, New York, 1986.

[2] R.A. Fisher, “The use of multiple measurements in taxonomic problems”, *Annuals of Eugenics*, vol. 7, pp. 111-132, 1936.

[3] M. Turk and A. Pentland, “Eigenfaces for recognition”, *Journal of Cognitive Neuroscience*. V. 3, pp. 71-86, 1991.

[4] X. M. Liu, T. Chen and B.V.K.V. Kumar, “Face Authentication for Multiple Subjects Using Eigenflow”, *Pattern Recognition, Special issue on Biometric*, Nov 2001.

[5] P.N. Belhumeur, J.P. Hespanha and D.J. Kriegman, “Eigenfaces vs Fisherfaces: recognition using class specific linear projection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 20, pp. 711 –720, 1997.

[6] V.N. Vapnik, *Estimation of dependences based on empirical data*, Springer-Verlag, Berlin, 1982.

[7] C. Cortes and V.N. Vapnik, “Support vector networks”, *Machine Learning*, vol. 20, pp. 273-297, 1995.

[8] V.N. Vapnik, *Statistical Learning Theory*, Wiley, New York, 1998.

[9] C.J.C. Burges, “A tutorial on support vector machines for pattern recognition”, *Knowledge Discovery and Data Mining*, vol. 2, no. 2, pp. 121-167, 1998.

[10] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines*, Cambridge University Press, Cambridge, UK, 2000.

[11] K.-R. Müller, S. Mika, G. Rätsch, , K. Tsuda, B. Schölkopf, “An Introduction to kernel-based learning algorithms”, IEEE Transactions on Neural Networks, Vol. 12, pp.181 – 201, 2001.

[12] K. Jonsson, J. Kittler, Y. P. Li, and J. Matas “Support vector machines for face authentication”, in *Proceedings of British Machine Vision Conference*, T. Pridmore and D. Elliman, Ed., pp. 543–553, 1999.

[13] P.J. Phollips, “Support vector machines applied to face recognition”’ in *Proceedings of Advances in Neural Information Processing Systems 11*, M.S Keans, S.A. Solla, and D.A. Cohn, Ed., 1998.

[14] G. Guo, S. Li, and K. Chan, “Face recognition by support vector machines”, in *Proceedings of the International Conferences on Automatic Face and Gesture Recognition*, pp.196-201, 2000.

[15] W. Zhang and T. Chen, “Authentication based on Symmetric Max Minimal distance in Subspace”, accepted by *Conference of Computer Vision and Pattern Recognition*, 2003.

[16] T. Sim, S. Baker, and M. Bsat, “The CMU Pose, Illumination, and Expression (PIE) Database”, in *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, 2002.