International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:3, No:8, 2009

Fusion Classifier for Open-Set Face Recognition with Pose Variations

Gee-Sern Jison Hsu

Abstract**—**A fusion classifier composed of two modules, one made by a hidden Markov model (HMM) and the other by a support vector machine (SVM), is proposed to recognize faces with pose variations in open-set recognition settings. The HMM module captures the evolution of facial features across a subject's face using the subject's facial images only, without referencing to the faces of others. Because of the captured evolutionary process of facial features, the HMM module retains certain robustness against pose variations, yielding low false rejection rates (FRR) for recognizing faces across poses. This is, however, on the price of poor false acceptance rates (FAR) when recognizing other faces because it is built upon within-class samples only. The SVM module in the proposed model is developed following a special design able to substantially diminish the FAR and further lower down the FRR. The proposed fusion classifier has been evaluated in performance using the CMU PIE database, and proven effective for open-set face recognition with pose variations. Experiments have also shown that it outperforms the face classifier made by HMM or SVM alone.

Keywords**—**Face recognition, open-set identification, hidden Markov model, support vector machines.

I. Introduction

THIS paper aims at *open-set face recognition* with pose variations. Open-set face recognition can be better interpreted using a gallery set and a probe set. A gallery set contains the subjects enrolled to the system with one or a few facial images per subject, and a probe set refers to the facial images unseen to the system and presented to the system for recognition. The images in both sets are disjoint. When both sets have the same individuals, it is known as *closed-set identification*, and each probe face has one and only one matched subject in the gallery. Many former algorithms were evaluated using this scenario. The closed-set identification is often not the case in real life, but the open-set recognition is. In open-set recognition, the probe set is larger than the gallery set, and those in the probe set but not in the gallery act as imposters trying to break in the gallery. In such a scenario, one must first determine whether a probe face exists in the gallery, and if it exists, one will have to identify what the matched subject is from the gallery set. Open set face recognition is considered more general, and thus more difficult, than closed-set identification because it actually adds in a detection task on top of an identification task. It is reported in [1] that in an open-set scenario a small size of gallery is easier to recognize than a large size.

Gee-Sern Jison Hsu is with the National Taiwan University of Science and Technology (phone: +886-2-2730-3234; e-mail: jison@mail.ntust.edu.tw).

Poses, illumination conditions, and expressions are generally acknowledged as three challenging parameters in face recognition. Quite a number of methods for recognizing faces across poses use 3D approaches [1]-[7], and among them the 3D morphable model [1, 2] may be the most well-known and considered an effective tool for handling poses. However, the 3D approaches suffer from intensive computation, possible imprecise alignments, and undesirable artifacts generated on the model-based virtual views. Therefore, many researchers have been working on approaches with perspectives different from 3D ones.

A geometry assisted probabilistic approach is reported in [8], which approximates a head with a 3D ellipsoid model, so that any face image is a 2D projection of such a 3D ellipsoid at a certain pose. Linear object classes (LOC) are introduced in [9, 10], which are formed by the prototypical views to a specific class of objects, as faces for example. LOC has the properties that the virtual views of any object of the same class under uniform affine 3D transformations can be generated if the corresponding transformed views are known for the set of prototypes. If a training set consists of frontal and rotated views of a set of prototypical faces, any rotated view of a new face can be generated from a single frontal view. Two issues have limited the application of LOC in practice, one is the finding of correspondence between the model and an image, and the other is the completeness of available examples for building the prototypes. The virtual view generation problem is reformulated as a prediction problem in [11], and solved by linear regression. This method is inspired by the idea that the linear mapping between non-frontal patches and frontal patches maintains better than that of the global case in the case of coarse alignment.

This paper reports a model that fuses a Hidden Markov Model (HMM) with a Support Vector Machine (SVM), each of which has been applied to face identification and recognition, and each has some specific advantages and weakness [12]-[19]. The proposed fusion model can keep their advantages and compensate for their weakness. It has been reported and also observed in our experiments that HMM is good for closed-set face identification with pose variations, but poor for open-set face recognition [20] with unacceptable false acceptance rates (FAR). The poor FAR is primarily caused by the HMM built upon within-class samples only. To exploit the HMM's strength in recognizing faces across poses and effectively diminish the FAR it induces, the proposed fusion model connects a HMM module with a SVM module following a special architecture. When enrolling a subject to the gallery, the HMM module, built from the subject's facial images, searches for those in the gallery whose faces look similar to the subject's face, and uses these similar faces as part of the training sample for building the SVM module. Because of the HMM-selected training samples, the SVM module effectively builds in a hyperplane using the between-class samples, substantially suppressing the FAR while maintaining a good level of the false rejection rate (FRR).

In the following, we will first review the face recognition algorithms using HMM or SVM in Section II, and then present the proposed fusion model in Section III. The fusion model has been evaluated in performance using the CMU PIE database, and the results will be presented in Section IV. The conclusion of this study with highlights on the future work will be given in Section V.

II. A Review on Face Recognition using HMM and SVM

A. Face Recognition by Hidden Markov Models

This approach considers the left-to-right and top-to-bottom variation in a face image an evolutionary process that can be described by a Hidden Markov Model (HMM). A HMM consists of states, the observation symbols of each state, the state transition probability that governs the transition from one state to the others, the observation generation probability that governs the observations generated by the states, and the initial state transition and observation generation distributions [21]. Given a sequence of observation symbols $\{O_j\}_{j=1}^{s}$, a general problem in HMM is to determine the model parameters $\lambda \equiv (A, B, \pi)$ so that the resultant HMM can best describe $\{O_j\}_{j=1}^{s}$, where $A = \{a_{j,k}\}$ is the set of state transition probabilities, $a_{j,k}$ is the transition probability from state i_k to state i_j; $B = \{b_{r,k}\}$ is the set of state observation probabilities, $b_{r,k}$ is the probability of observing O_r at state i_k; $\pi = \{\pi_l\}$ is the set of initial state probabilities, π_l is the probability of the initial state being at state i_l. Note that the number of states is assumed known in the above settings.

For face recognition applications [12]-[15], the observations are extracted from a face image and must preserve the evolutionary geometric variation across the face. The work in [12] uses 1-D HMM that takes in the observations formed by the DCT coefficients of a series of partially overlapping rectangle windows running from top to bottom across the face.

The 1-D HMM model is extended in [13, 14], which use a 2-D HMM, or an embedded HMM, which takes in as observations the DCT coefficients of a series of small square windows partially overlapped right to left and then top to bottom. The embedded HMM segments a face into a series of rows that represent some invariant features from top to bottom, such as the forehead, eyes, nose, and mouth [13], as shown in Fig. 1. These rows constitute the *super-states*. Embedded in each super-state, a number of *regular states* are assumed to capture the variation left to right in that super-state. The case shown in Fig. 1 shows that a face is segmented into 4 rows of super-states for forehead, eyes, nose, and mouth. The forehead area is further segmented into 3 states, the eyes area into 5, and the nose and mouth areas each into 4. Assuming the generation of the observations governed by a mixture of Gaussians, one can apply the Viterbi algorithm to obtain the best HMM for describing the observations.

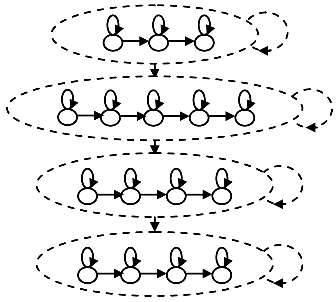

Fig. 1 Segmentation of a face image into rows of super-states and each embeds a number of states left to right

Because the HMM model is built upon a subject's facial images, we believe that it can characterize the evolutionary dynamics of the local facial features captured by the running windows across the subject's face. The evolutionary dynamics of local features may remain the same for pose variations to some extent. It is reported in [20] and observed in our experiments that an HMM facial model trained on one pose can well identify the same face in some neighboring poses in closed-set identification scenarios. However, because the HMM facial model is built upon the within-class samples only, which refer to the subject's own facial images, it lacks the measure of the dissimilarity between two different faces, resulting in poor performance when rejecting imposters. In summary, the HMM facial model can work well in closed-set identification, but fails to do a good job in open-set face recognition.

B. Support Vector Machines for Face Recognition

Support Vector Machines (SVMs) establish the optimal separating hyperplane for solving binary classification problems. Given a training dataset, SVM can reach the right balance between the accuracy attained on this training dataset and the ability to classify other disjoint datasets, making it a classifier with the best generalization performance [22]. For nonlinear classification applications, SVM often takes a kernel $K(\cdot,\cdot)$ that satisfies Mercer's condition to build the best separating hyperplane in the following form:

$$f(\mathbf{x}) = \sum_{i=1}^{n_s} \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b \qquad (1)$$

where $\{\mathbf{x}_i\}_{i=1}^{n_s}$ are the support vectors, the most represent-tative ones from the training samples, and each is associated with a class labeled by $y_i \in \{-1,1\}$; $\{\alpha_i\}_{i=1}^{n_s}$ and b are the coe-fficients given by solving an associated complex quadratic programming problem. The kernel $K(\cdot,\cdot)$ maps the training samples to a high-dimensional space in which the nonlinear classification becomes linear. The most popular kernels are linear, polynomial, and radial basis function (RBF) [22]. Given a probe image $\mathbf{x}$, represented in some specific feature vector form, one can determine its class using the following distance function,

$$d(\mathbf{x}) = \frac{1}{\left\| \sum_{i=1}^{n_S} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}_i) \right\|} \sum_{i=1}^{n_S} \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b \tag{2}$$

The sign of $d(\mathbf{x})$ shows on which side of the separating hyperplane $\mathbf{x}$ is located, and its magnitude gives the distance of $\mathbf{x}$ away from the hyperplane. The larger the distance, the more reliable the classification result.

The feature vectors in [16, 17] are generated by Eigenface and Fisherface decompositions, and those in [18] are normalized gray-valued pixel vectors formed by the face images after histogram equalization and lighting intensity subtraction. SVM is good at extracting the information for the discrimination of unclassified features, e.g., the features from PCA; however, it is prone to be over-trained with well-classified features, e.g., the features from LDA [17].

SVM aims at binary classification, but can be extended to multiple classification with two schemes. One is the one-to-the-rest scheme. If there are n classes in the training set, the negative samples of one specific class are the conglomerate of all the rest n-1 classes. The other is pairwise approach, each classifier only involves two out of the n classes, so there will be $n(n-1)/2$ classifiers in total.

The major disadvantage of the two schemes is the expensive computation due to the huge amount of training data. In the one-to-the-rest scheme, the support vectors of each class require the training upon the whole data set; and in the pairwise approach, there are too many classes to train. Training upon a large number of data poses a serious threat to the application scope of SVMs [22].

III. Fusion of HMM and SVM

Aiming at open set face recognition, we propose a fusion model, which keeps the advantages of both HMM and SVM methods, and effectively overcomes the weakness of each. The fused model is composed of a HMM module and a SVM module. In the enrollment stage, a subject's HMM module is built from the subject's facial images, and acts as a filter to select those in the gallery whose faces look similar to the subject's face. These similar faces are then used as part of the training sample for building the SVM module, substantially reducing the amount of training samples for the SVM, and also effectively suppressing the HMM-induced FAR in the recognition stage. Furthermore, in the recognition stage the similarity between a probe face and each subject in the gallery can be readily measured by the subject's HMM module, and the rank-n pool can be quickly determined which includes n subjects whose faces are considered similar to the probe face. The way of developing the HMM modules is similar to those reported in [12]-[15], the originality of this work is on (1) the fusion of the HMM module and the SVM module so that both can be integrated, and (2) the training of SVM using a generic negative sample set and a subject-oriented negative sample set.

A. Features for the Fusion Model

The features for the HMM module must preserve the variations of local facial features captured by the window moving from left to right and top to bottom. Similar to the work in [10, 11, 12], the DCT coefficients taken from each patch captured by a moving square window are used in this work. Only the low frequency parts, i.e., those in the upper triangular of each patch's DCT coefficients, are extracted. The DCT coefficients from the overlapping patches are good for building the HMM module. But they are inappropriate for building the SVM module, because of the high dimensionality of the feature space formed by these coefficients. Consider a case with a 64x64 face and a 8x8 patch overlapped with its neighbors for 3 pixels, if the largest 15 DCT coefficients are taken from each patch, it will result in a feature vector of 5415 in dimension. This size of dimension can paralyze a computer when it is running a SVM session with a large number of training samples.

A solution to the above is to downsample the features and use some subset of the features. Assuming that each patch in the above example overlaps its neighbors for 2 pixels only and just the largest 9 DCT coefficients are taken, a feature dimension of 900 is attained. In our experimental study, we tried 5 DCT coefficients from each patch, and the accuracy degraded at a negligible degree but came with faster training due to the reduced feature dimension of 500.

B. Development of the SVM module in the Fusion Model

We propose a special design to the generation of the SVM module for each subject when enrolling to the gallery set. This design consists of the following steps:

1. For each subject, the positive sample set is formed by the subject's own face images, but the negative sample set is composed of a *generic* (or *white*) set and a *subject-oriented* set of face images.
2. The generic negative set aims at carrying a wide spectrum of face variations across individuals, poses, illuminations and other factors. The generic negative set can be made, or *approximated*, by selecting the *representative* samples from a large face database using self-organizing maps (SOM's) and principal component analysis (PCA). Each subject in the gallery set shares this same generic negative set, but has a subject-oriented (or tailor-made) negative sample set.
3. Using the easy-to-be-misclassified samples to strategically carve the SVM hyperplane, the subject-oriented negative sample set is meant to reduce the false acceptance rate (FAR) and improve the recognition rate. This sample set is formed by the face images of those in the gallery who look similar to the enrolling subject, and the similarity is measured by the subject's HMM module. The subject's HMM module can select a rank-m pool of m similar faces to form the subject-oriented negative samples. This subject-oriented negative set and the generic negative set constitute the complete negative training set for building the subject's SVM module. We have tested a few kernels, and decided to use the radial basis function (RBF) because it gives better performance than others.

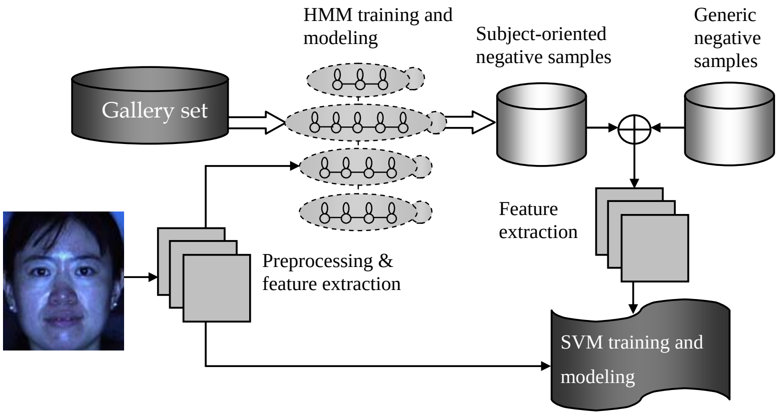

Fig. 2 Flowchart of the training of a fusion model

4. When a subject is enrolled to the gallery, the SVM modules of those in the rank-*m* pool who are considered similar in faces to the subject's must also be reassessed. Each subject in the rank-*m* pool will use his HMM module to measure the similarity of the new enrolled face to his own, if the new enrolled one enters his rank-*m* pool list, his SVM module will be updated.

The subject-oriented negative set sub-optimizes the separating hyperplane in the sense that it increases the gap between the subject's samples and those who can be easily misclassified into the subject's class. The SVM module built in this way can offer some similar classification performance to that built using the 1-to-the-rest training scheme, but the training can be much faster and the model size can be smaller because of the relatively small amount of samples considered in training.

Given a face to be enrolled to the gallery set, the aforementioned training procedure for making its fusion model is flowcharted in Fig. 2.

C. Face Recognition using the Fusion Model

When a probe face image is given to the fusion model with a gallery set, the HMM module of each subject in the gallery will give a likelihood score showing the similarity between the probe face and the subject's face. We can select *n* subjects with *n* highest likelihood scores from the gallery and form a rank-*n* candidate pool, and then compare the distances measured by the SVM modules of the *n* candidates and determine the one with the shortest distance from the SVM hyperplane. This twofold scheme allows some manipulation to the FAR and FRR of the system, as discussed below:

1. When we enhance the security level, this means that we ask for a lower FAR on the price of a higher FRR. We can take in less candidates given by the HMM module, i.e., a smaller *n*, so that each probe face only has a few chances to be matched. On the contrary, if we lower down the security level, a larger *n* can be chosen, and a probe face can have more chances to find a match.
2. The SVM module can also be designed for different security requirements using different sizes of negative training sets. The smaller the negative set, the higher the FAR and the lower the FRR. On the contrary, the larger the negative set, the stricter the security level will be. It is, however, by no means trivial to collect a sufficiently large set of *good* negative samples by approaches, such as bootstrapping. The proposed rank-*n* candidate pool pre-selected by the HMM module can easily adjust the size of the negative set according to the similarity between a subject and those who have better chances of being falsely recognized as the subject.

IV. Experiments for Performance Evaluation

The performance of the proposed fusion model was evaluated using the CMU PIE database [23]. Although PIE offers 68 subjects and each with 13 different poses, 43 different illumination conditions, and 4 different expressions, we selected 5 poses from each subject, as an example shown in Fig. 3, and each pose with 7 different illumination conditions, as shown in Fig. 4. Because this study focuses on poses, it can be seen that the variation across the different illumination conditions in Fig. 4 is not substantial.

Fig. 3 Five poses from PIE database are considered in our experiments

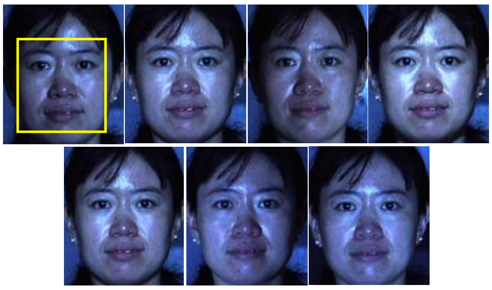

Fig. 4 Seven similar illumination conditions were selected, and the yellow box shows the facial area considered in our experiments

A. Test Protocols for Performance Evaluation

Two different test protocols were considered. One studied the impact of different number of images available for training each subject's fusion model in the gallery set, and the other studied the performance variation with different gallery sizes. Details are as follows:

Protocol-1. Different number of facial poses available for each subject to be enrolled to the gallery: for the fusion model we ran two different test scenarios. The **Test-1** scenario started with 2 poses, frontal (F) and left-sided (L); and then 3 poses, frontal (F), left-sided (L), and right-sided (R); then 4 poses, F, L, R, and upward (U); and finally with all 5 poses, F, L, R, U, and downward (D). The **Test-2** scenario again starts at 2 poses, but with F and U; then 3 poses, F, U, and D; and then 4 poses, F, U, D, and L. In each scenario, we randomly selected 34 subjects out of the 68 available from the PIE database for enrolling to the gallery. For each subject in the gallery, we randomly selected the facial image samples from one out of the seven selected illumination conditions for building the fusion model. The samples in the rest six illumination conditions were used to compute the FRR (false rejection rate). The samples of the other 34 individuals were used to compute the FAR (false acceptance rate). The randomized selection scheme was repeated for 12 times, and the average rate was reported. The samples used in the **Test-1** scenario were also used to build a SVM model and a HMM model for performance comparison. To attain a fair comparison, the thresholds in all algorithms were adjusted to make the FAR at 0.005, except for the HMM model. If the FAR of the HMM was set at 0.005, the FRR would have been over 0.9; therefore, the FAR was set to 0.15.

Protocol-2. Different galley sizes and probe sizes: the gallery size and probe size refer to the number of individuals in the gallery and probe sets, respectively. Given an upper bound of 68 subjects in the PIE database, we tested the gallery sizes of 10, 20, and 34 with the probe sizes of 20, 40, and 68, respectively. Those in the gallery set were also in the probe set, but with different sets of images. The data partition was same as in the Protocol-1, but aims at open-set settings: for each gallery size, an equal size of *imposters* are there trying to break in. For each gallery size, we repeated the test with a random selection on the subjects for enrollment, and then a random selection of one out of the seven illumination conditions to provide training samples. As this protocol aims at the performance of the proposed fusion model handling galleries of different sizes in open-set settings, we used all 5 poses available from the selected illumination condition for enrollment.

B. Sample Preprocessing and Test Results

All faces were aligned by the eyes, and normalized to 64x64 pixels in size according to the distance between the eyes. Each facial image was converted from the original color image into an 8-bit gray-scale image. We subtracted the best-fit linear plane from each image to reduce possible illumination impacts, and then equalized its intensity histogram.

To make the generic negative set for the SVM module, we collected a large number of face images from other benchmark databases, including FRGC [24, 25], AR [26], and XM2VTS [27], and some from the internet. Our collection had 8,156 facial images with different poses, expressions, ethnic backgrounds, and under various illumination conditions. 626 representative ones were extracted using a self-organizing map (SOM) with facial features extracted by PCA (Principal Component Analysis).

Rank-5 candidate pool was used in all experiments, for either training or testing, i.e., *n*, the number of similar faces selected to train the HMM module is 5, and the number of candidates selected to validate a probe face is also 5. For the proposed fusion model, 8x8 squares overlapped for 4 pixels have been chosen along with the major 15 coefficients from each square's DCT map served as the features for the HMM module. As mentioned in Section III-*A*, part of these features were used by the SVM module to form the feature vector of dimension 500. The HMM-only for comparison used the same features as those used in the fusion model, but the SVM-only used the 666 low-frequency DCT coefficients extracted from each image.

All test results are the average of 10 randomized selections of the gallery sets with the associated probe set. The performance for Protocol-1 is shown in Fig. 5 where the two ways of pose variations, Test-1 and Test-2, are compared with its SVM-only and HMM-only counterparts. With a pre-selected FAR at 0.005, the FRR of the fusion model with both Test-1

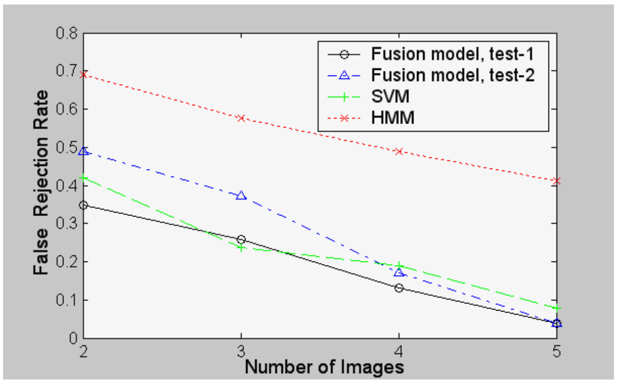

Fig. 5 Performance of the fusion model varies with the number of images available for enrollment to the galley set, compared with the HMM and the SVM algorithms (FAR=0.005)

and Test-2, and the SVM-only case are shown, together with a pre-selected FAR at 0.18 for the HMM-only case. Choosing FAR 0.18 for the HMM is because that its FRR has been found to reach a level over 0.85 when the corresponding FAR reduces to below 0.05. Fig. 5 gives the following observations:

1. The fusion model outperforms the SVM-only and the HMM-only algorithms, especially for the case of Test-1. Enrolled with two poses, frontal (F) and left-sided (L), the fusion model gives a FRR at 0.34, better than the SVM-only with FRR 0.42 using the same training samples but without the assistance of the HMM module for defining the candidate set.
2. The performance of the fusion model degrades if only the F and U (upward) poses are available. It might imply that the L and R poses may better interpolate the U and D poses, but not vice versa; and this needs more experiments to validate. When more poses are available, the performance of all algorithms improves. When the available poses are more than four, fusion model gives the best performance.

Although the 68 subjects in the dataset may not be good enough for studying the influences of gallery sizes on open-set recognition, the false rejection rates (FRR) with 3 gallery sizes, 10, 20, and 34 were still computed following the Protocol-2 and 10 randomized selections on the subjects enrolled to the gallery. The result is given in Fig. 6. The performance was shown in terms of the FRR with a pre-selected FAR at 0.005 for the fusion model and SVM-only. The HMM-only test was with FAR 0.18, as described in the Protocol-1 part of test. Fig. 6 shows the following:

1. The larger the gallery, the worse the FRR. This trend shows some similar observations to those reported in FRVT 2002 [1], reflecting the fact that some more advanced face recognition is needed to handle large galleries. Although better than the HMM-only and SVM-only algorithms, the proposed fusion model needs to be further improved for such a task.
2. The poor performance given by the HMM-only algorithm has again proven that HMM is not appropriate for open-set face recognition, although it was reported to have performed well for close-set identification.

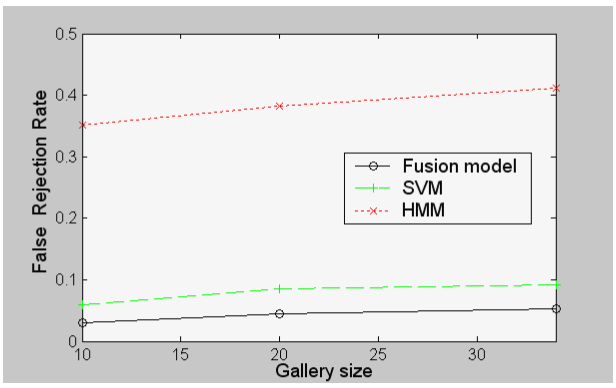

Fig. 6 False rejection rates of the fusion model, SVM-only, and HMM-only algorithms with different sizes of galleries

It should be noted, in addition to the result shown in Fig. 6, that the model size of the fusion model from this experiment is around 227±20 KB in average regardless of gallery sizes, but that of the SVM is 326 KB in average for galleries with 10 subjects, and increases to 361 KB for galleries with 34 subjects.

V. Conclusion

The important features and the continuing phase of the proposed fusion model are summarized below:

1. The fusion model maintains the advantage of HMM in recognizing faces with pose variations in the settings of closed-set identification, and substantially suppresses the high FAR for open-set recognition using a special fusion that combines HMM and SVM.
2. The fusion model can reach a right balance between recognition performance, model size, and processing time. This model is especially effective in coping with the cases in which the subjects with similar faces may lead to a high FAR, and such cases can be common when the gallery set is large.
3. Its two-fold scheme can be extended to other applications or ways of fusing two or more different classifiers. Many different types of classifiers have been made available in the last decade for research upon pattern recognition and computer vision. The ways of putting them together are yet to investigate, and the author believes that this can lead to some different perspectives and potentials of using these classifiers.
4. Validated in the Protocol-1 test upon the PIE database, the performance of the fusion model can be improved when the enrollee's (pose) samples increase. This implies that if a system with the fusion model embedded can take in more sample images, its capability of open-set face recognition can be improved over time. This is an on-going study conducted by the author, and the result will be reported in a following paper.

References

[1] P.J. Phillips, P. Grother, R.J. Micheals, D.M. Blackburn, E. Tabassi, M. Bone, "Face Recognition Vendor Test 2002: Evaluation Report," available at http://www.frvt.org.

[2] V. Blanza and T. Vetter, "Face recognition based on fitting a 3D morphable model," *IEEE Trans. Pattern Analysis and Machine Intelligence* (*PAMI*), vol. 25, no. 9, pp. 1063-1074, 2003.

[3] H.C. Choi; S.Y Kim; S.H. Oh; S.Y. Oh; S.Y. Cho; "Pose invariant face recognition with 3D morphable model and neural network," *Proc. IEEE Int'l J. Conf. Neural Networks* (IJCNN), 2008, pp. 4131 – 4136.

[4] Q. Chen, J. Yao, W.K. Cham, "3D model-based pose invariant face recognition from multiple views," *Computer Vision, IET,* vol.1, i.1, pp. 25 – 34, March 2007.

[5] S. Feng, H. Krim, I. Gu, M. Viberg, "3D Face Recognition Using Affine Integral Invariants," *Proc. IEEE Int'l Conf. Acoustics, Speech and Signal Processing* (*ICASSP*) 2006, vol.2, pp. 14-19.

[6] A.F. Abate, M. Nappi, D. Riccio, G. Sabatino, "3D Face Recognition using Normal Sphere and General Fourier Descriptor," *Proc. 18th Int'l Conf. Pattern Recognition* (*ICPR*), 2006, vol.3, pp. 1183-1186.

[7] L. Wang, L. Ding, X. Ding, and C. Fang, "Improved 3D assisted pose-invariant face recognition, *Proc. ICASSP* 2009, pp. 889 – 892.

[8] X. Liu; T. Chen, "Pose-robust face recognition using geometry assisted probabilistic modeling," *Proc. IEEE Conf. Computer Vision and Pattern*

Recognition, (*CVPR*) 2005, vol.1, pp. 502 – 509.

[9] T. Vetter and T. Poggio, "Linear object classes and image synthesis from a single example image," *PAMI*, vol.19, no.7, pp.733–742, 1997.

[10] T. Vetter, "Synthesis of novel views from a single face image," *Int'l J. Computer Vision* (*IJCV*), vol.28, no.2, 1998, pp.103–116.

[11] X. Chai, S. Shan, X. Chen and W. Gao, "Locally Linear Regression for Pose-Invariant Face Recognition," *IEEE Trans. Image Processing*, vol. 16, Iss.7, pp. 1716-1725, 2007.

[12] A.V. Nefian and M.H. Hayes III, "Hidden Markov Models for Face Recognition," *Proc. ICASSP*, 1998, pp. 2721-2724.

[13] A.V. Nefian and M.H. Hayes III, "An Embedded HMM Based Approach for Face Detection and Recognition," *Proc. ICASSP*, vol.6, 1999, pp. 3553-3556.

[14] S. Eickeler, S. Müller, and G. Rigoll, "Improved Face Recognition using Pseudo 2-D Hidden Markov Models," in *Workshop on Advances in Facial Image Analysis and Recognition Technology* (*AFIART*), Freiburg, Germany, 1998.

[15] F. Samaria and S. Young, "HMM-based architecture for face identification," *Image and Vision Computing*, 12(8), pp. 537-543, 1994.

[16] P.J. Phillips, "Support Vector Machines Applied to Face Recognition," in *Advances in Neural Information Processing Systems*, vol.11, M.J. Kearns et. al., eds., MIT Press, 1999.

[17] K. Jonsson, J. Matas, J. Kittler, and Y.P. Li, "Learning Support Vectors for Face Verification and Recognition," *Proc. IEEE Int'l Conf. on Automatic Face and Gesture Recognition* (*FG*), 2000, pp. 208-213.

[18] B. Heisele, P. Ho, and T. Poggio, "Face Recognition with Support Vector Machines: Global versus Component-Based Approach," *Computer Vision and Image Understanding* (*CVIU*), vol.91, no. 1/2, pp. 6-21, 2003.

[19] B. Heisele, T. Serre and T. Poggio, "A Component-based Framework for Face Detection and Identification," *IJCV*, vol.74, no.2, pp. 167-181, 2007.

[20] P.H. Lee, Y.W. Wang, J. Hsu and Y.P. Hung, "Facial Features Extracted by 2-D HMM for Face Recognition with Pose Variations," *Proc. of IAPR Conference on Machine Vision Applications* (*MVA*), pp. 392~395, 2007.

[21] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Se-lected Applications in Speech Recognition," *Proc. of IEEE*, vol. 77, no. 2, pp. 257-286, 1989.

[22] C.J.C. Burges, "Simplified Support Vector Decision Rules," *Proc. 13th Int Conf. on Machine Learning*, pp. 71-78, 1996.

[23] T. Sim, S. Baker, and M. Bsat, "CMU pose illumination and expression(PIE) database," *PAMI, IEEE Trans*, vol.25, NO.12, Dec 2003. pp. 1613 – 1618, 2003.

[24] Frgc P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, M. Jaesik, W. Worek, "Overview of the Face Recognition Grand Challenge," *CVPR 2005*, vol.1, 20-25, pp.947-954.

[25] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, W. Worek, " Preliminary Face Recognition Grand Challenge Results," *Proc. 7th Int'l Conf Automatic Face and Gesture Recognition*, pp. 15-24, 2006.

[26] A.M. Martinez and R. Benavente, "The AR Face Database," *CVC Technical Report #24* , June 1998.

[27] K. Messar, J. Matas, and J. Kittler, "XM2VTSDB: The Extended M2VTS Database, " *Proc. 2nd Int'l Conf. Audio and Video-based Biometric Person Authentication* (*AVBPA'99*), pp. 2 – 14, 1999.