ROBUSTNESS OF A FACE-RECOGNITION TECHNIQUE BASED ON SUPPORT VECTOR MACHINES

Prashanth Harshangi[1] and Koshy George[2]

[1]PES Centre for Intelligent Systems, Bangalore, India
E-mail: pharshangi@yahoo.com
[2]Department of Telecommunication Engineering, PES Institute of Technology, Bangalore, India
E-mail: kgeorge@pes.edu

Abstract
The ever-increasing requirements of security concerns have placed a greater demand for face recognition surveillance systems. However, most current face recognition techniques are not quite robust with respect to factors such as variable illumination, facial expression and detail, and noise in images. In this paper, we demonstrate that face recognition using support vector machines are sufficiently robust to different kinds of noise, does not require image pre-processing, and can be used with rather low image resolutions.

Keywords:
Face Recognition, Support Vector Machines

1. INTRODUCTION

The recognition capability of human beings has no parallel even with current sophisticated computational power. The high degree of connectivity, and the capacity to learn through adaptation, are fundamental to the central nervous system. Consequently, the numerous highly interconnected biological neurons can outperform super computers in facial feature extraction, face detection and face recognition. For example, it takes very little time for an infant to recognise its parents, and, even for a child, picking a known face from a group photograph is a trivial task. Achieving similar capabilities using machines is typically rather tedious. Nonetheless, recent advancements in computing capability have created inroads into such face recognition technology.

Face recognition approaches can broadly be classified into two categories. Early face recognition algorithms used simple geometric feature matching such as eyes, nose, mouth, and skin colour [1]–[3]. Later techniques use distances between the features as descriptors of faces [4]. This reduces the computational complexity drastically. However, such techniques rely heavily on feature extraction and measurement of facial features.

The second class of methods use template matching and generally operate directly on an image-based representation of faces; i.e., pixel intensity array [3, 5]. Thus, the need for the more cumbersome geometric feature extraction is eliminated. This class of methods is more practical and can operate in near real-time.

Successful face recognition technologies ought to have the capacity to deal with various changes in face images that include changes in orientation and expression. Surprisingly, the mathematical variations between the images of the same face due to illumination and viewing direction are almost always larger than image variations due to change in face identity. This presents a great challenge to face recognition [6, 7].

Support Vector Machines (SVMs), developed by Vapnik [8], have been used in a number of applications including pattern recognition [9], and isolated handwritten digit recognition [8]. SVMs have also been used in face recognition singly as in [10], or jointly with other techniques such as kernel principal component analysis (PCA) [11], Gabor features [12], binary edge map [13], scale-invariant feature transform [14], and multi-scale PCA based on Gabor wavelets in [15]. In the context of such applications, the robustness of singular value decomposition with PCA is discussed in [16], and the robustness of combining global and local features compared with PCA in [17]. Although the robustness of SVMs per se have been discussed in the past, to the best knowledge of the authors, the robustness of SVMs for face recognition applications has not been discussed.

The objectives of this paper are three-fold: First, we demonstrate that in terms of face-recognition accuracy, a face-recognition procedure that uses SVMs works best when not used in conjunction with typical image pre-processing procedures. Second, SVMs perform better when it is has access to the whole image, rather than the features extracted via, for instance, PCA. Third, when SVMs are used alone for face-recognition, it is quite robust to reasonable levels of different kinds of noise, and to changes in face details.

The paper is organised as follows: In Section 2, we introduce support vector machines, and principal component analysis in Section 3. Typical image pre-processing steps used in face recognition are discussed in Section 4. The efficacy of using only SVMs, and the robustness of such a technique, are brought out through the simulation results in Section 5.

2. SVMs FOR FACE-RECOGNITION

Face recognition is a multi-class problem. In this paper we study the robustness of support vector machines (SVMs) to achieve this multi-classification. Essentially, SVMs are maximum margin classifiers. In its simplest form, the linear SVM classifier determines a hyperplane that separates two classes, C_1 and C_2. If $\{(x_i, d_i)\}$, $1 \le i \le N$, is a set of training examples, where $x_i \in R^n$ belong to one of two classes, and the class labels d_i take one of two possible values, the basic idea here is to determine a separating hyperplane $f(x) = w^T x + b$ if one exists. Here, $w \in R^n$ is the weight vector, and b the bias. For any training sample x_i, the sign of $f(x_i)$ indicates on which side of the hyperplane x_i is positioned; that is, the class of x_i. The discriminant $f(x) = 0$ defines the hyperplane, and if the bias b= 0, the hyperplane passes through the origin of R^n.

For the above linear classifier a hyperplane $f(x) = 0$, when it exists, separates the two classes C_1 and C_2. The closest data

point is referred to as the margin of separation. An SVM linear classifier maximizes this margin of separation. Therefore, the separating hyper plane is determined such that $f(x) \geq \eta$ if $x_i \in C_1$ and $f(x) \geq \eta$ if $x_i \in C_2$ for some fixed η. Without loss of generality, $\eta = 1$; otherwise, the weight vector w and the bias b can always suitably be scaled.

The problem of finding the optimal hyperplane that maximizes the margin of separation can be restated as the following quadratic optimization problem:

$$\min \tfrac{1}{2} w^T w \qquad (1)$$

subject to the constraints

$$d_i \left(w^T x_i + b \right) \geq 1, \quad 1 \leq i \leq N$$

This leads to the following solution: $w = \sum_{i=1}^{N} \alpha_i d_i x_i$

where $\sum_{i=1}^{N} \alpha_i d_i = 0$

These concepts can be extended to the non-separable case. Here, the constraints are modified as follows:

$$d_i \left(w^T x_i + b \right) \geq 1 - \xi_i, \quad 1 \leq i \leq N$$

Where ξ_i s are the slack variables. The data points lie inside the region of separation if $\xi_i \in [0, 1]$. The problem of finding the optimal hyperplane is now stated as follows:

$$\min w^T \phi(x) + M \sum_{i=1}^{N} \xi_i \qquad (2)$$

subject to the aforementioned constraints, where $\phi(x)$ is the feature vector. The solution to this is $w = \sum_{i=1}^{N} \alpha_i d_i \phi(x_i)$, and the hyperplane described by $\sum_{i=1}^{N} \alpha_i d_i \phi^T(x_i)\phi(x_i) = 0$. More generally, the optimal hyperplane can be defined by

$$\sum_{i=1}^{N} \alpha_i d_i K(x, x_i) = 0 \qquad (3)$$

where $K(x, x_i)$ is an inner-product kernel which is a symmetric function of its arguments. There are several choices of such kernels; in this paper we choose the polynomial learning machine:

$$K(x, x_i) = \left(x^T x_i + 1 \right)^2 \qquad (4)$$

Since SVMs are essentially binary classifiers, multi-classification is achieved using either one of the following strategies [18]-[20]: (i) one-against-one and (ii) one-against-all. If M is the number of classes, the former strategy requires $\tfrac{1}{2} M (M - 1)$ discriminant functions, and hence as many SVMs, and the latter strategy, only $(M - 1)$ SVMs are required. Therefore, structurally the latter strategy is simpler to implement, and has been widely used for face recognition; for instance, [13]. It has been frequently commented that both strategies are equally effective. In this paper, we follow the one-against-one strategy and demonstrate that, at the cost of extra complexity; this strategy either outperforms, or is as effective as,

other strategies for face recognition that uses SVM.

During the training phase of an SVM for two classes, the images are mapped into a higher dimensional space via an inner-product kernel to linearly separate the data. As mentioned earlier, the kernel used is a polynomial kernel of degree 2, present with an inherent bias. The optimum hyperplane is obtained by a quadratic optimization of the weight vector norm. The test sample is presented to all learning machines at the bottom of the tree, which selects an appropriate winning class on a layer-by-layer basis, till the winning class is obtained at the top. In this paper we use the SVM-KM Matlab toolbox [21].

3. PRINCIPAL COMPONENT ANALYSIS

The basic idea of Principal Component Analysis (PCA) is to transform a number of possibly correlated quantities into a smaller number of uncorrelated ones referred to as principal components. When applied to digitised images of human faces, these components are called *eigenfaces*. To generate these eigenfaces, the images, taken under the same lighting conditions, are normalized so that the eyes and mouths are lined up. Subsequently, these are re-sampled at the same pixel resolution.

The created eigenfaces appear as light and dark areas arranged in a specific pattern. These patterns determine the manner in which the different features of a face are singled out to be evaluated and scored. For instance, if there is any facial hair there is a pattern to evaluate symmetry. Similar remarks hold for the nose or mouth. However, other eigenfaces may have patterns that are less simple to identify, and, consequently, may have very little resemblance to a face.

The steps involved in creating a set of eigenfaces are as follows: First, the training set is pre-processed so that all the images in the set have the same resolution, and the faces roughly aligned. Further, using the vec operation, the columns of the two-dimensional matrix that forms the image is stacked to form a single column vector. These are then transposed and then arranged into a matrix, denoted T; thus each row of T corresponds to an image. Secondly, the average value of the images is removed. In addition, the eigenvalues, and corresponding eigenvectors of the related covariance matrix S are then computed. These eigenvectors are the eigenfaces, and are the directions in which the images in the training set differ from the mean image. Thirdly, those eigenvectors associated with the larger eigenvalues are retained, discarding the others.

By projecting a new mean-subtracted image on to the eigenfaces, we can then determine the extent to which it differs from the mean face. The loss of information caused by projecting it on to a subset of all the eigenvectors is kept at an optimum level, and is related to those eigenvalues of S that have been discarded. In practical applications, an image of size 100x100, and hence having 10,000 eigenvectors, requires only about 20 to 60 eigenfaces.

4. IMAGE PRE-PROCESSING

In this paper, we use the database created by AT&T laboratories, Cambridge, UK – formerly referred to as the ORL Database – and available with the Speech, Vision, and Robotics Group of Cambridge University, UK. This database consists of

40 subjects with varying illumination, facial expressions (open or closed eyes, smiling or not smiling) and facial details (glasses or no glasses). The size of each image is 92x112 pixels. There are 10 images for each person. Of these, 5 images of a subject are used for training, and the remaining 5 for testing. Samples of images are shown in Fig. 1.

Fig.1. Samples of faces of four individuals in the database

An important phase in a typical face-recognition technique is pre-processing the images. The advantages of such pre-processing are reduction in the illumination variations by normalising with respect to lighting conditions, and a reduction in the overall computational complexity. Some of the pre-processing steps are histogram equalisation, median filtering and bi-cubic interpolation.

Histogram equalisation usually increases the global contrast of many images, especially when the usable data of the image is represented by close contrast values. Through this adjustment, the intensities can be better distributed on the histogram. This allows for areas of lower local contrast to gain a higher contrast without affecting the global contrast. Histogram equalisation accomplishes this by effectively spreading out the most frequent intensity values. The method is useful in images with backgrounds and foregrounds that are both bright or both dark. The calculation is not computationally intensive. A disadvantage of the method is that it is indiscriminate. It may increase the contrast of background noise, while decreasing the usable signal. It has been pointed out in [7] that histogram equalisation combined with median filtering reduces the effect of lighting variations in the face images. The process of histogram equalisation is shown in Fig. 2.

Fig.2. Histogram Equalisation

The key advantage of resizing through bi-cubic interpolation is that it produces smoother surfaces than any other interpolation technique. Bi-cubic interpolation takes into account 16 pixels in a rectangular grid, takes weighted average of pixels, and replaces then with a single pixel; accordingly, that pixel has the flavor of all the replaced pixels. This reduces redundant information.

Median filtering reduces the effects of illumination variations

in a face database [7, 18]. The median filter is often used to remove noise from images or other signals. The idea is to examine a sample of the input and decide if it is representative of the signal. The edge preserving nature of median filter is claimed to be particularly helpful in improving the recognition accuracy.

5. RESULTS

The objectives of the study are three-fold. As mentioned earlier, we study the effect of pre-processing of images on face recognition accuracy (FRA) when multi-classification is achieved with SVMs using the strategy outlined earlier. In addition, we also discuss the effect on FRA when the features of the image are extracted, and these features passed on to an SVM-based face recognition technique. Finally, we study the robustness of such a technique toward noise in the images, and toward changes in the details of the face in the images.

We test our strategy of face recognition using SVMs on the ORL database mentioned earlier. Two hundred samples of images of faces (five for each individual) are randomly chosen as the training set. The remaining two hundred samples are used as the test set. For all the experiments considered here, we use a polynomial kernel with degree 2.

5.1 SVMs WITH PRE-PROCESSING

We consider here the effect of histogram equalization and median filtering on image resolutions. Accordingly, we consider the following scenarios:
(i)   Images pre-processed with histogram equalisation and median filtering.
(ii) Images pre-processed with only median filtering.
(iii) Images not pre-processed by either histogram equalization or median filtering.
All experiments are conducted with four different image resolutions obtained by bi-cubic interpolation: 50x40, 20x20, 10x10, and 5x5.
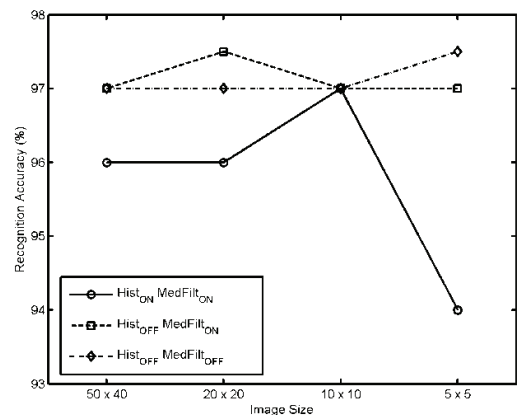

Fig.3. Recognition accuracy with SVM for different pre-processing of images

The results are as shown in Fig. 3. Evidently, the achieved face recognition accuracy (FRA) is between 94% and 97.5%. Moreover, the figure suggests that SVMs work best without any pre-processing, and FRA is independent of the resolution size;

approximately 97% accuracy is achieved for all resolutions. (The minor anomaly of an increase of 0.5% in the case without any pre-processing for a resolution of 5x5 is perhaps due to numerical errors.) Further, there is very little effect of median filtering, and histogram equalisation has the worst effect on FRA. We recall that the database consists of faces of persons with varying illumination, facial expressions (open or closed eyes, smiling or not smiling) and facial details (glasses or no glasses). Despite these variations, our procedure of using only SVMs with the one-against-one strategy for multi-classification achieves reasonable FRA. Thus, our procedure is robust to changes in illumination, facial expressions and detail.

5.2 SVMs WITH PCA

In most face-recognition applications, SVMs have been used in conjunction with another technique; of these, PCA has been used more often. Here, we contrast the results presented earlier with those obtained when SVMs are used along with PCA. These are depicted in Fig. 4. Evidently, the overall FRA is poorer compared to when only SVMs are used. Further, pre-processing has a more significant effect on the FRA when SVM is used together with PCA. Although better accuracy is again obtained when no pre-processing is used, the values are lower than that shown in Fig. 3.
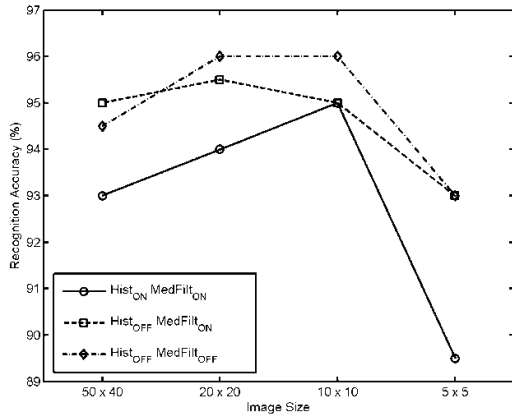

Fig.4. Recognition accuracy with SVM-plus-PCA for different pre-processing of images

5.3 NOISE HANDLING CAPABILITY OF SVM

We study here the robustness of SVMs toward noise in the context of face recognition. We consider the following two scenarios:

5.3.1 Noisy Images in the Test Set

In the first set-up (referred to as *Scenario A* in the sequel) we assume that the images used for training has been obtained in a controlled environment but that the test images are not; e.g., in a biometric system setup in an organisation. Thus, the images used for training the system have very little noise, but the images used for testing are quite noisy. For these experiments, the noisy images were resized to 50x50, 40x40, 30x30, 20x20 and 10x10.

We first consider the effect of salt-and-pepper noise. It may be noted that salt-and-pepper noise affects the whole image, and this is the major type of noise to be dealt with when low-resolution cameras are used, especially in airports. Here, salt-and-pepper noise of different intensities are added to the images. (A noise intensity of, for instance, 0.25 implies that 25% of the face is affected with salt-and- pepper noise.) The results are shown in Fig. 5.
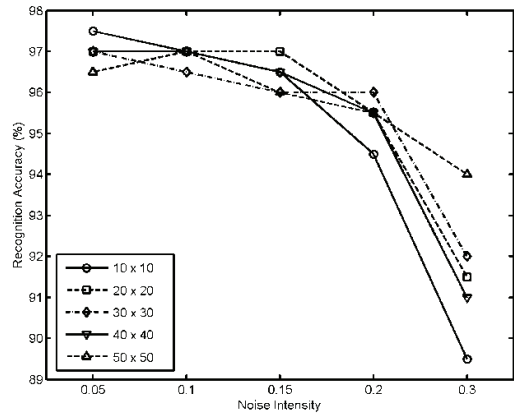

Fig.5. Salt-and-pepper noise: FRA with SVM (Scenario A)

The maximum FRA achieved is still better than 97%, and is comparable to the no-noise case. Thus, there is very little effect on the FRA for reasonable levels of noise intensity, and this is relatively independent of the resolution size. Of course, it is natural to expect a drop in the FRA with higher levels of noise intensities, and noise to affect more the lower resolution images. Both effects can be observed in Fig. 5. The results of similar experiments when SVM is used together with PCA are shown in Fig. 6. It is clear from this figure that this procedure lacks robustness even with very low noise intensities. The results are also summarized in Tables 1 and 2. Thus SVMs, when used alone, are satisfactorily robust toward reasonable levels of salt-and- pepper noise.

Since median filter is the best to remove salt-and-pepper noise, the test images are pre-processed using such a filter. The accuracy achieved is consistently 97% across various noise levels. Along with median filter, when histogram equalization is also applied on the images, there is a slight decrease in accuracy from 97% to 95.5%.
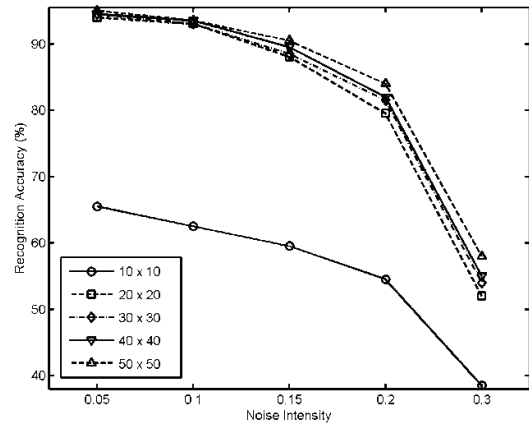

Fig.6. Salt-and-pepper noise: FRA with SVM-plus-PCA (Scenario A)

Table.1. Salt-and-pepper noise: FRA with SVM

Intensity / Resolution	0.05	0.1	0.15	0.2	0.3
10x10	97.5	97.0	96.5	94.5	89.5
20x20	97.0	97.0	97.0	95.5	91.5
30x30	97.0	96.5	96.0	96.0	92.0
40x40	97.0	97.0	96.5	95.5	91.0
50x50	96.5	97.0	96.0	95.5	94.0

Table.2. Salt-and-pepper noise: FRA with SVM-plus-PCA

Intensity / Resolution	0.05	0.1	0.15	0.2	0.3
10x10	65.5	62.5	59.5	54.5	38.5
20x20	94.0	93.0	88.0	79.5	52.0
30x30	94.5	93.0	88.5	81.5	54.0
40x40	94.5	93.5	89.5	82.0	55.0
50x50	95.0	93.5	90.5	84.0	58.0

The effect of Gaussian noise is considered next. We add to the images zero-mean Gaussian noise with varying variance. The results are shown in Fig. 7 using SVM and Fig. 8 using PCA with SVM, and summarised in Tables 3 and 4.
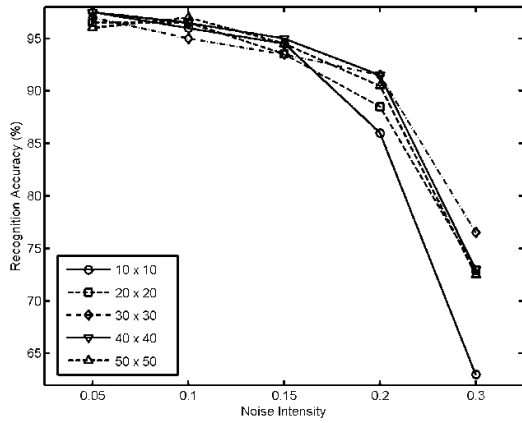

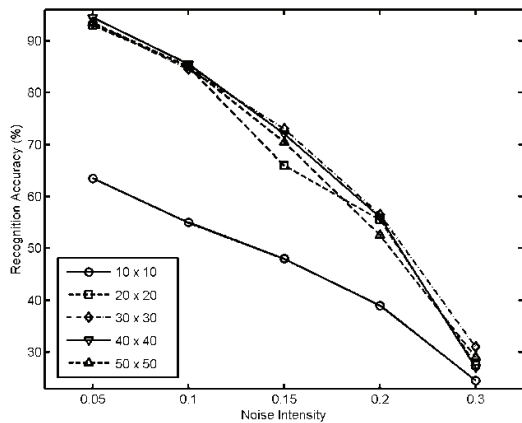Fig.7. Gaussian noise: FRA with SVM (Scenario A)

Fig.8. Gaussian Noise: FRA with SVM-plus-PCA (Scenario A)

Table.3. Gaussian noise: FRA with SVM

Intensity / Resolution	0.05	0.1	0.15	0.2	0.3
10x10	97.5	96.0	94.5	86.0	63.0
20x20	96.5	96.5	93.5	88.5	73.0
30x30	97.0	95.0	93.5	91.5	76.5
40x40	97.5	96.5	95.0	91.5	73.0
50x50	96.0	97.0	94.5	90.5	72.5

Table.4. Gaussian noise: FRA with SVM

Intensity / Resolution	0.05	0.1	0.15	0.2	0.3
10x10	63.5	55.0	48.0	39.0	24.5
20x20	93.0	85.0	66.0	55.5	27.5
30x30	93.5	84.5	73.0	56.5	31.0
40x40	94.5	85.5	72.0	56.0	27.0
50x50	93.5	85.0	70.5	52.5	29.0

As is quite clear from Figures 7 and 8, the effect of Gaussian noise on FRA is similar to that of the effect of salt and pepper noise. Again, the face-recognition is better when SVMs are not used along with PCA. Similar comments can be made for Poisson noise. For brevity we do not include any figures; however, the results are summarized in Table 5.

Table.5. Poisson noise: FRA with SVM and SVM-plus-PCA

Method / Resolution	SVM	SVM + PCA
10x10	97.5	65.0
20x20	97.0	94.5
30x30	97.0	95.0
40x40	97.0	94.0
50x50	97.0	94.5

We conclude here that SVMs are robust to different kinds of noise, and for reasonable levels of noise intensities for the first scenario.

5.3.2 Noisy Images in the Training Set

An example for *Scenario B* would be from forensics. Various images of a person are available from different sources, taken in different locations and different conditions. When the identity of a captured person is to be verified, his image, taken under controlled environment, is compared with the database. Thus, in this scenario, the images used for training is noisy, but the test images may have only very little noise. The studies conducted for this scenario is similar to that for the earlier set-up. The results with only SVM, and PCA with SVM, are respectively shown in Figures 9 and 10 for salt and pepper noise, and Figures 11 and 12 for Gaussian noise.
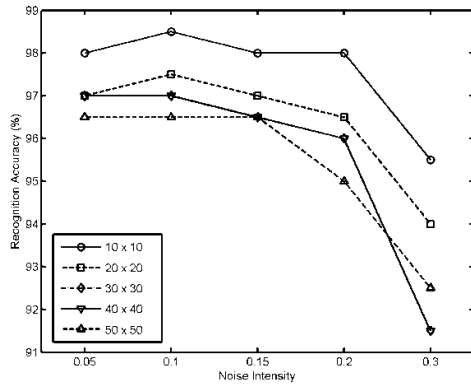
Fig.9. Salt-and-pepper noise: FRA with SVM (Scenario B)

It is quite evident from these figures that the face recognition accuracy for Scenario B with SVM is better than with SVM-plus-PCA for both salt-and-pepper noise and Gaussian noise. As expected, the FRA for Scenario B is lower than that of Scenario A.
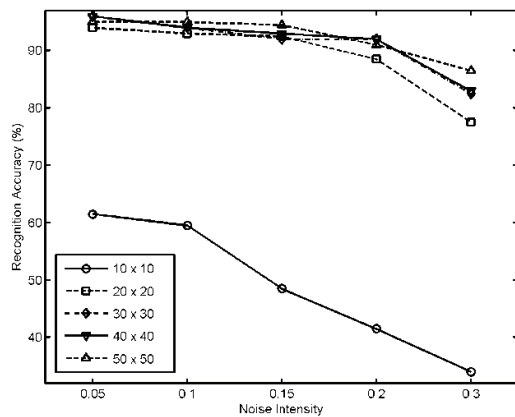

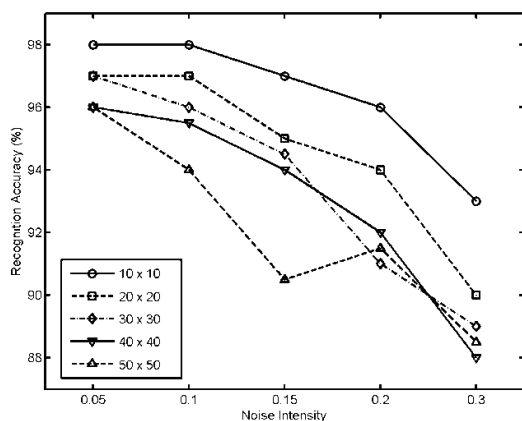Fig.10. Salt-and-pepper noise: FRA with SVM-plus-PCA
(Scenario B)

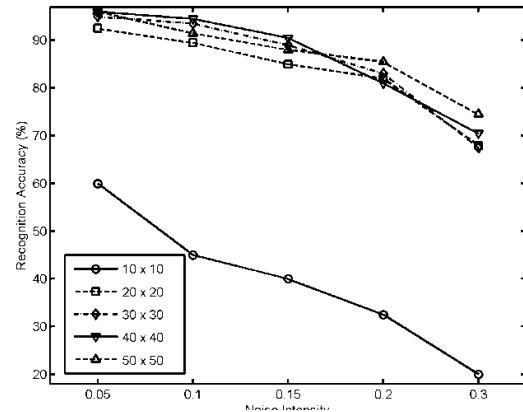Fig.11. Gaussian noise: FRA with SVM (Scenario B)

Fig.12. Gaussian noise: FRA with SVM-plus-PCA (Scenario B)

5.3. *Comparison*

We now compare our results with those strategies that used SVMs in the past [10]-[15], [17]. In general, the FRA obtained with our strategy is typically much better than those reported elsewhere. In addition, the performance of SVM with pre-processing is comparable with some of the results in these references. We note that the robustness of SVM to noise have not been discussed in these references.

6. CONCLUSION

The robustness of an SVM-based face recognition technique is studied here. Here, SVMs are used for multi-classification with a one-against-one strategy. At the cost of higher complexity, it has been observed that in terms of recognition accuracy, this procedure performs as well as, or better than, other procedures that use SVM, either singly or together with other techniques. Moreover, no pre-processing of images (with reasonable resolutions) is required to enhance the recognition accuracy. Further, the procedure is robust towards changes in illumination, facial expression, and facial details. Furthermore, it is also robust to reasonable levels of noise intensities of different noise such as salt-and-pepper noise, Gaussian noise and Poisson noise.

REFERENCES

[1] Goldstein, A.J., L.D. Harmon, and A.B. Lesk, 1971, "Identification of human faces". Proceedings of the IEEE Vol. 59, No. 5, pp.748–760.

[1] Samal, A., and P.A. Iyengar, 1992, "Automatic recognition and analysis of human faces and facial expressions": A survey. *Pattern Recognition*, Vol. 26, pp. 65–77.

[2] Brunelli, R., and T. Poggio, 1993, "Face recognition: Features versus templates". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 1, pp.1042–1052.

[3] Cox, J.G.J., and P. Yianilos, 1996, "Feature-based face recognition using mixture-distance," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 209–216.

[4] Turk, M., and A. Pentland, 1991, "Eigenfaces for recognition". *Journal of Cognitive Neuroscience*, Vol. 3, No.1, pp.71-86.

[5] Moses, Y., Y. Adini, and S. Ullman, 1994, "Face recognition: The problem of compensating for changes in illumination direction," in *Proceedings of the European Conference on Computer Vision*, pp. 286–296.

[6] Lu, Y.-M., B.-Y. Liao, and J.-S. Pan, 2008, "Face recognition algorithm decreasing the effect of illumination," in *Proceedings of the International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 378–381.

[7] Cortes, C., and V. Vapnik, 1995, "Support vector networks". *Machine Learning*, Vol. 20, pp.273–297.

[8] Burges, C.J, 1998, "A tutorial on support vector machines for pattern recognition". *Knowledge Discovery and Data Mining*, Vol. 2: pp.121–167.

[9] Lam, A., and C. R. Shelton, 2008, "Face recognition and alignment using support vector machines," in *Proceedings of the 8th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 1–6.

[10] Dai, G., and C. Zhou, 2003, "Face recognition using support vector machines with the robust feature," in *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication*, pp. 49–53.

[11] Qin, J., and Z.-S. He, 2005, "A SVM face recognition method based on Gabor-featured key points," in *Proceedings of the Fourth IEEE International Conference on Machine Learning and Cybernetics*, pp. 5144–5149.

[12] Qiu, X., W. Wang, J. Song, X. Zhang, and S. Liu, 2007, "Face recognition based on binary edge map and support vector machine," in Proceedings *of the First IEEE International Conference on Bioinformatics and Biomedical Engineering*, pp. 519–522.

[13] Zhang, L., J. Chen, Y. Lu, and P. Wang, 2008, "Face recognition using scale invariant feature transform and support vector machine," in *Proceedings of the Ninth IEEE International Conference for Young Computer Scientists*, pp.1766–1770.

[14] Zhang, G. and J. Zhang, 2003, "Face recognition using multi-scale PCA and support vector machine," in *Proceedings of the 7th World Congress on Intelligent Control and Automation*, Chongqing, China, pp.5906–5910.

[15] Noushath, S., A. Rao, and G.H. Kumar, 2007, "SVD Based algorithms for robust face and object recognition in vision based applications," in Proceedings of the 24th International Symposium on Automation and Robotics in Construction, paper, pp.5-7.

[16] Kathavarayan, R.S., and M. Karuppasamy, 2010, "Preserving global and local features for robust face recognition under various noisy environments". *International Journal on Image Processing*, Vol. 3, No. 6, pp.328–340.

[17] Hsu, C., and C. Lin, 2002, "A comparison of methods for multi-class support vector machines". *IEEE Transactions on Neural Networks*, Vol. 13, No. 2, pp.415–425.

[18] Weston, J., and C. Watkins, 1998, "Multi-class support vector machine", Technical Report, CSD-TR-98-04, Department of Computer Science, University of London.

[19] Abe, S, 2003, "Analysis of multiclass support vector machines," in Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation (CIMCA'2003), pp. 385–396.

[20] Canu, S., Y. Grandvalet, V. Guigue and A. Rakotomamonjy, 2005, "SVM and kernel methods Matlab toolbox," in *Perception Systemes et. Information*, INSA de Rouen, France.