World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:1, No:7, 2007

A Novel Neighborhood Defined Feature Selection on Phase Congruency Images for Recognition of Faces with Extreme Variations

Satyanadh Gundimada and Vijayan K Asari

***Abstract*—**A novel feature selection strategy to improve the recognition accuracy on the faces that are affected due to non-uniform illumination, partial occlusions and varying expressions is proposed in this paper. This technique is applicable especially in scenarios where the possibility of obtaining a reliable intra-class probability distribution is minimal due to fewer numbers of training samples. Phase congruency features in an image are defined as the points where the Fourier components of that image are maximally in-phase. These features are invariant to brightness and contrast of the image under consideration. This property allows to achieve the goal of lighting invariant face recognition. Phase congruency maps of the training samples are generated and a novel modular feature selection strategy is implemented. Smaller sub regions from a predefined neighborhood within the phase congruency images of the training samples are merged to obtain a large set of features. These features are arranged in the order of increasing distance between the sub regions involved in merging. The assumption behind the proposed implementation of the region merging and arrangement strategy is that, local dependencies among the pixels are more important than global dependencies. The obtained feature sets are then arranged in the decreasing order of discriminating capability using a criterion function, which is the ratio of the between class variance to the within class variance of the sample set, in the PCA domain. The results indicate high improvement in the classification performance compared to baseline algorithms.

***Keywords*—**discriminant analysis, intra-class probability distribution, principal component analysis, phase congruency

I. Introduction

Over the past 15 years, research has focused on making face recognition systems more accurate and fully automatic. Significant advances have been made in the design of classifiers for successful face recognition. Among appearance based holistic approaches, eigenfaces [1, 2] and Fisherfaces [3] have proved to be effective on large databases. PCA performs dimensionality reduction by projecting the original n-dimensional data onto the lower dimensional linear subspace spanned by the leading eigenvectors of its covariance matrix. Its goal is to find a set of mutually orthogonal basis functions that capture the directions of maximum variance in the data and for which the coefficients are pair-wise decorrelated. Unlike PCA, LDA encodes discriminating information in a linearly separable space using bases that are not necessarily orthogonal. In addition to these methods there are other methods such as Independent Component Analysis. Kernel methods such as Kernel Principal Component Analysis (KPCA) and Kernel Fisher Discriminant Analysis (KFDA) [4] show better results in face recognition than linear subspace methods. Nonlinear projection based methods have been able to overcome the problem of expressions and lighting in face images to some extent. But there has not been a significant improvement in the recognition accuracy in situations where the face images undergo lot of variations including expressions, partial occlusions and lighting and at the same time not many samples are available to represent the distribution. There are recent publications [5,6] in this direction of expression, occlusion and lighting invariant face recognition. In [6] a weighted distance measure is implemented which reduces the effect of pixels in the test image, which underwent significant movement from the corresponding positions in the training images. The technique presented in this paper is aimed to deal with varying expressions, partial occlusions and extreme lighting variations.

To make the recognition process robust to illumination, the classification technique is carried out on the phase congruency maps of the face images. The feature selection process presented in this paper is derived from the concept of modular spaces [8].

Achieving lighting invariance in face recognition is in itself a huge challenge. Some of the work done in this direction is given in [10-12]. It is important that the features that are selected for classification are invariant to illumination changes. Phase congruency maps are divided into very small sub regions of 4×4 pixels. Modular subspaces of dimension 64 (8×8 pixels) are created by considering various combinations of the smaller sub regions. The combinations are then indexed in the order of their closeness to each other. This indexing strategy is based on the empirical results, which suggest that the modular spaces created from the combination of the smaller sub regions that are closer in locality are more

Manuscript received February 27, 2006.

Satyanadh. Gundimada is with the Department of Electrical Engineering, Old Dominion University, Norfolk, VA 23529 (e-mail: sgund002@ odu.edu).

Vijayan. K. Asari is with the Department of Electrical Engineering, Old Dominion University, Norfolk, VA 23529 (e-mail: vasari@ odu.edu).

effective in the process of classification.

Phase congruency features [7], which are robust to illumination variations, are used for the classification purpose. Modular eigen subspaces [2][8,9] on phase congruency images, combined with a voting mechanism are implemented in this proposed technique.

A criterion function which gives a measure of the discriminating capability of a module, by calculating the ratio of the between class variance to the within class variance is used to arrange the modular subspaces. This arrangement helps in reducing the number of modules that are necessary to obtain the highest accuracy achieved by the classifier.

The proposed technique is computationally efficient and has achieved high accuracy rates on certain freely available standard face databases. The testing strategy is implemented in such a way that the training set consists of face images taken under controlled conditions where as the testing set consists of images captured in uncontrolled conditions. The paper is organized as follows. The second section describes the effect of facial variations on recognition accuracy and the role of modularization in reducing those effects. Section three describes the technique of implementation to obtain phase congruency features from the corresponding intensity images. Fourth section provides the implementation steps of the proposed feature selection process. Section five gives the algorithmic steps for the implementation of the proposed face recognition technique. Section six explains the details about the type of testing strategy that is implemented along with the obtained recognition accuracies on face database.

II. Facial Variations

Variations caused in facial images due to expression, makeup and non-uniform lighting tend to move the face vector away from the neutral face of the same person both in image space and reduced linear subspace. It has been observed that the dimensionality reduction techniques on individual modules or the local regions of the face images improve the accuracy of face recognition compared to applying on the whole image. An experiment is conducted on AR database to show the effect of modularization of the face images on recognition accuracies. Two sets of images, one with expressions mostly affecting the mouth regions and the other set with partial occlusions on the bottom half of the face images. The two sets are tested separately using a leave one out strategy. All the training images are divided into 64 local regions. Each region or the module is projected into a reduced eigen space. The test module is classified using a nearest neighbor algorithm. The accuracies of the individual modules for both the sets are shown in figure 1.

Sample face images of each set are also displayed besides the accuracy results in the figure 1. It can be observed that only the regions that are affected adversely due to facial variations have low accuracy rates. In conventional non-overlapping modular subspaces, pixel dependencies are confined only to the local regions. Or in other words the pixels in one sub-region are considered independent to the pixels in

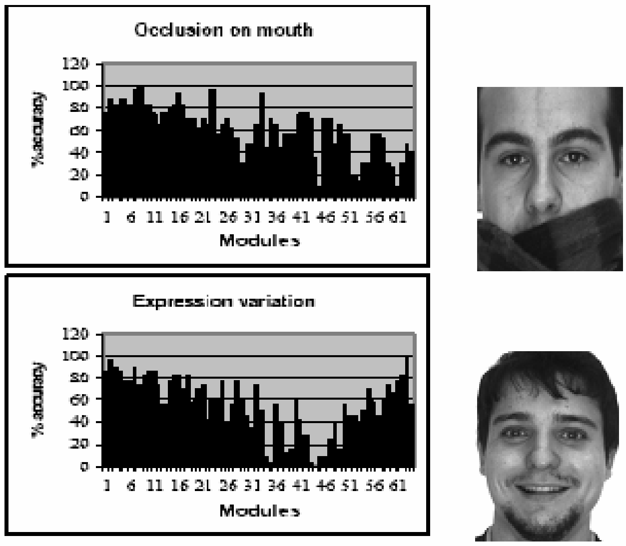

Fig. 1 Illustration of the accuracies of various local regions in the respective modular eigen space domain

other sub-regions. Since the facial variations in real world scenario are confined to local regions, considering additional pixel dependencies across various sub-regions could help in providing additional information, which could help in improving the classification accuracy. A feature selection policy based on the above discussion in which modular spaces are created with pixels from across various local regions taking into account the locality of such regions is provided in section four. A criterion function, which reduces the total number of features required to attain highest accuracy with the full set of features, is also presented. The above-mentioned techniques are carried out on phase congruency maps of the training images. A detailed explanation of the phase congruency technique is given in the following section.

III. Phase Congruency Features

Gradient-based operators, which look for points of maximum intensity gradient, will fail to correctly detect and localize a large proportion of features within images. Unlike the edge detectors, which identify the sharp changes in intensity, the phase congruency model detects points of order in the phase spectrum. According to Opeinheim and Lim [13], phase component is more important than the magnitude component in the reconstruction process of an image from its Fourier domain. There is also physiological evidence, indicating that human visual system responds strongly to the points in an image where the phase information is highly ordered. Phase congruency provides a measure that is independent of the overall magnitude of the signal making it invariant to variations in image illumination and/or contrast. Figure 2 shows phase congruency image and the corresponding intensity image. The phase congruency technique used in this paper is based on the one developed by Peter Kovesi [7].

Phase congruency function in terms of the Fourier series expansion of a signal at some location x is given by.

$$PC(x) = \frac{\sum_n A_n \cos(\phi_n(x) - \bar{\phi}(x))}{\sum_n A_n} \quad (1)$$

Where A_n represents the amplitude of the nth Fourier component, and $\phi_n(x)$ represent the local phase of the Fourier component at position x. $\bar{\phi}(x)$ is the weighted mean of all the frequency components at x. Phase congruency can be approximated to finding where the weighted variance of local phase angles relative to the weighted average local phase, is minimum. An alternative and easier interpretation of phase congruency is proposed in [7]. It is proposed that energy is equal to phase congruency scaled by the sum of the Fourier amplitudes as shown in equation 2.

$$E(x) = PC(x)\sum_n A_n \quad (2)$$

Hence phase congruency is stated as the ratio of $E(x)$ to the overall path length taken by the local Fourier components in reaching the end point. This makes the phase congruency independent of the overall magnitude of the signal. This provides invariance to variations in image illumination and contrast. $E(x)$ can be expressed as $E(x) = \sqrt{F(x)^2 + H(x)^2}$.

If $I(x)$ is the input signal then $F(x)$ is the signal with its DC component removed and $H(x)$ is the Hilbert transform of $F(x)$ which is a 90^0 phase shift of $F(x)$. Approximations to the components $F(x)$ and $H(x)$ are obtained by convolving the signal with a quadrature pair of filters. In order to calculate the local frequency and phase information in the signal, logarithmic Gabor functions are used. If $I(x)$ is the signal and M_n^e and M_n^o denote the even symmetric and odd-symmetric wavelets at a scale n. The amplitude and phase of the transform at a given wavelet scale is given by equation 3 and equation 4 respectively.

$$A_n = \sqrt{e_n(x)^2 + o_n(x)^2} \quad (3)$$

$$\phi_n = \tan^{-1}\left(o_n(x) / e_n(x)\right) \quad (4)$$

where $e_n(x)$ and $o_n(x)$ are the responses of each quadrature pair of filters. Equation 5 illustrates the response vector.

$$\left[e_n(x), o_n(x)\right] = \left[I(x) * M_n^e, I(x) * M_n^o\right] \quad (5)$$

$F(x)$ and $H(x)$ can be obtained from the equations 6 and 7.

$$F(x) = \sum_n e_n(x) \quad (6)$$

$$H(x) = \sum_n o_n(x) \quad (7)$$

And $\sum_n A_n$ at x is given by equation 8.

$$\sum_n A_n(x) = \sqrt{e_n(x)^2 + o_n(x)^2} \quad (8)$$

If all the Fourier amplitudes at x are very small then the problem of phase congruency becomes ill conditioned. To overcome the problem a small positive constant ε is added to the denominator. The final phase conguency equation is given by equation 9.

$$PC(x) = \frac{E(x)}{\varepsilon + \sum_n A_n} \quad (9)$$

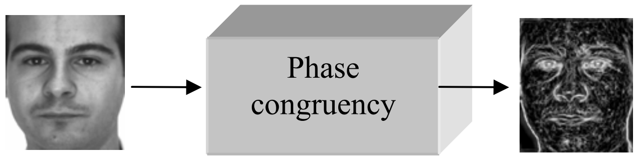

Fig. 2 Phase congruency map obtained from the corresponding intensity image

One-dimensional analysis is carried out over several orientations, and the results are combined to analyze a two dimensional signal (image)[7].

IV. Neighborhood Defined Feature Selection and Arrangement

Facial variations are confined mostly to local regions in reality. Modularizing the images would help in localizing these variations, provided the modules created are sufficiently small. But in doing so, large amount of dependencies among various neighboring pixels are ignored. This can be countered by making the modules larger, but this will result in improper localization of the facial variations. In order to deal with this problem, a novel module creation strategy is implemented in this paper. It has been proved that, dividing an image of size 64×64 into regions of size 8×8 pixels is appropriate in achieving high classification accuracy results [8]. An image under consideration is divided into small non-overlapping sub-regions, here a size of 4×4 pixels. Four of such 4×4 pixel regions are merged to form 8×8 pixel regions (modules). By doing so, we will be able to localize the facial variations and will be able to consider most of the pixel dependencies. But if all such combinations are considered, then a huge set of 29056 modules will be formed, which is not practical. It is emphasized in various modular subspace techniques, that the local dependencies are maximally important. Hence the 4×4 pixel regions considered for creating 8×8 pixel regions are restricted to local regions or in other words local neighborhoods. Here, a region of 16×16 pixel region is termed as neighborhood. A total of 16 such non-overlapping regions are created within a phase congruency image of size 64×64. In such a case a set of 120 such 8×8 modules can be created in each neighborhood. Hence a total of 1920 modules can be created. To further reduce the regions, two 4×4 pixel regions are concatenated to form 4×8 pixel regions called CHB's (Concatenated Horizontal Blocks). Now there are 8 such CHBs in one neighborhood. Combination of these CHBs within each neighborhood are created to form the desired regions of 8×8 pixels. Also 8×8 pixel regions are formed from 8×4 pixel regions within each neighborhood which are called CVBs (concatenated Vertical Blocks). Figure 3 explains the creation of CHBs and CVBs from within each neighborhood. There exist redundant modules because of this process. These are identified and removed. Thus a set of 52 such 8×8 pixel modules are created from each neighborhood to form a total of 832 modules. For each module, an eigen subspace is created from the training data set and the corresponding modules from the test image are projected into the subspaces to classify them.

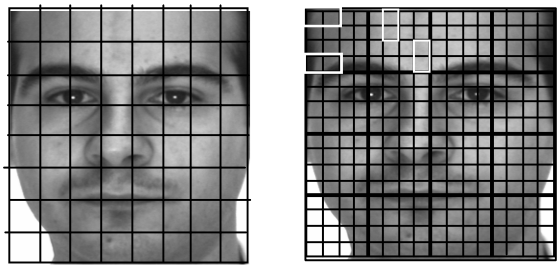

Fig. 3 (a) Left image shows non-overlapping modules, each of size 8×8 pixels (b) The right image shows the creation of CHBs and CVBs, (white rectangular boxes) from two small 4×4 regions. The smaller blocks are 4×4 pixel regions and the thicker, larger blocks are the neighborhoods

A probe image is assigned to a class, which gets the maximum number of modules classified in its favor. In order to reduce the number of modules further, an arrangement strategy based on the locality of the modules (CHBs and CVBs) that are in combination and a criterion function is implemented.

A. *Sorting the modules based on locality*

The total modules that are created are sorted using an indexing strategy first. The indexing is separate to each of the defined neighborhood (16×16 pixels block). The index is given in the increasing order of the distance between the modules that are combining. This can be observed in the figure 4 and figure 5. For example in figure 4, CHB(1) of size 4×8 pixels, can form 7 different combinations with rest of the CHBs in the neighborhood. So, an index of 1 is given to the immediate neighbors of the CHB(1). The CHB diagonal to CHB(1) is given an index of 2 and so on. This process is carried out for all the 16×16 blocks. The modules within each neighborhood, having similar index are grouped together. A criterion function explained in the next sub-section is used to arrange modules of same index in the decreasing order of their discriminating capability.

B. *Selecting the modules based on criterion function*

A criterion function is implemented in this paper, based on which a minimum number of features that are required to achieve the highest accuracy is determined. This criterion function is used to arrange the modules and there by the features needed for classification in the decreasing order of their importance. Fisher criterion is used in many face recognition techniques. It gives a subspace in which the training data belonging to different classes is maximally separable. The same theory of fisher criterion is applied here. A module's importance in classification is determined by its capability in separating the classes on its own. In order to determine that, a criterion function, which is the ratio of the between class variance to within class variance of each module in the PCA domain is proposed. The modules are arranged in the decreasing order of the criterion function. It is observed that the modules covering the mouth regions are less important according to the criterion function. So, after giving first preference to the locality, the next step is to arrange the modules with same index according to the criterion function within each neighborhood. Firstly the best discriminating module of index 1 from each of the larger blocks (neighborhood of 16×16 pixels) is added to the existing set (which is zero initially). Hence in the first round, 64 modules of index 1 are added to the set, where each is best discriminating within each neighborhood. This set of 64 modules is again sorted according to the criterion function. This process is continued until all the modules are considered. Equation 10 explains the procedure for selection of the modules.

$$J(w_{nk}) = \frac{\det\left(W_{nk}^{T} S_{B} W_{nk}\right)}{\det\left(W_{nk}^{T} S_{W} W_{nk}\right)} \tag{10}$$

where n = 1,2,..52 and k= 1,2,..16 represent the indices of combination and neighborhood respectively. $J(w_{nk})$ is the ratio of the between class variance and within class variance of the training set in the eigen sub space represented by W_{nk}. S_B and S_W are the between class and within class variances respectively in the corresponding subspaces. The ratio $J(w_{nk})$ represents the discriminating capability of that module. Hence the modules are sorted according to this ratio. For example, module set 1 consists of $\{M_{n1},M_{n2},..M_{n16}\}$, where M_{nk} is the best discriminating module among the modules with index 1 in k^{th} neighborhood. It is observed that following this technique; only an module set of around 270 is able to achieve the highest accuracy rate of that with all modules.

V. ALGORITHMIC STEPS

Step 1: Normalize all the training images to a size of 64×64 using nearest neighbor interpolation method.

Step 2: Calculate the phase congruency at each pixel position for each of the training images as described in section three.

Step 3: Modularize the 64×64 dimension phase congruency maps into non-overlapping regions of size 4×4.

Step 4: Concatenate two horizontally adjacent modules of size 4×4 to form a 4×8 pixel block called concatenated horizontal block (CHB) as shown in the figure 4.

Step 5: Create combinations of two CHB's within the defined neighborhood of 8 CHBs to form a set of 28 different modules of size 8×8 pixels each..

Step 6: Index the modules (8×8) according to the location of the two CHB's involved in the combination. Figure 4 shows CHB 1 and the indexing given to the other CHBs that would be involved in the combinations.

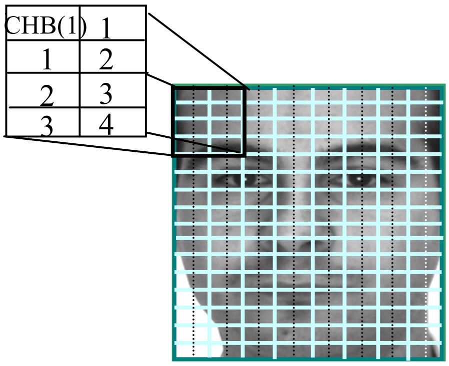

Fig. 4 Image shows creation of CHBs

Step 7: Concatenate two 4×4 blocks into concatenated vertical blocks (CVB) of size 8×4 each.

Step 8: Follow the procedure from step 4 to step 6 in creating 8×8 modules.

Step 9: Combine the set of modules formed in step 5 and 8. Also eliminate the redundant modules (modules that contain same 4×4 pixel modules).

Step 10: Arrange the modules according to the increasing order of the indexing given within the neighborhood. Also arrange the modules with the same index in the decreasing order of the discriminating capability of each module, which can be calculated from equation 10.

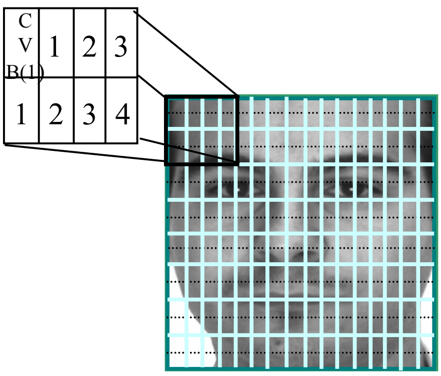

Fig. 5 Image shows the creation of CVB

Step 11: Take one module at a time consecutively (after indexing and sorting) from each neighborhood block to make a set of 64 modules. These are again arranged according to decreasing order of the criterion function.

Step 12: Create eigen subspaces for each of the modules that are created after arrangement.

Step 13: Classify the probe image based on voting procedure. Assign the probe image to the class, which gets the maximum number of modules classified in its favor.

VI. Experimental Results

From the AR database 40 individuals are chosen randomly to create a test database. 13 images of each individual are present in the database.

Fig. 6 Training images of an individual in AR database

Fig. 7 Sample test images of the person in figure 6

Three images of each individual are used in training the proposed technique for classification. Figure 6 shows the training images of an individual. It can be observed that all the three face images are fairly neutral with little expression variations. The rest of the 10 images of each individual in the database are used for testing the proposed technique. The sample test images are shown in the figure 7. The test images are affected due to lighting or (and) expressions or (and) partial occlusions. The graph in figure 8 illustrates the relationship between percentage of accuracy and the dimensionality of the subspace for various methods such as principal component analysis on holistic faces (PCA), modular PCA (MPCA), principal component analysis on phase congruency features (PPCA), modular subspace approach on phase congruency features (MPPCA) and the proposed method of Neighborhood defined module selection on phase congruency features in PCA domain (NPPCA).

It can be observed that the use of phase congruency features improves the face recognition accuracy significantly. Also modular subspaces improve the recognition for both intensity and phase congruency features. It can be observed that accuracy has risen by about 10 % in the case of NPPCA compared to MPPCA.

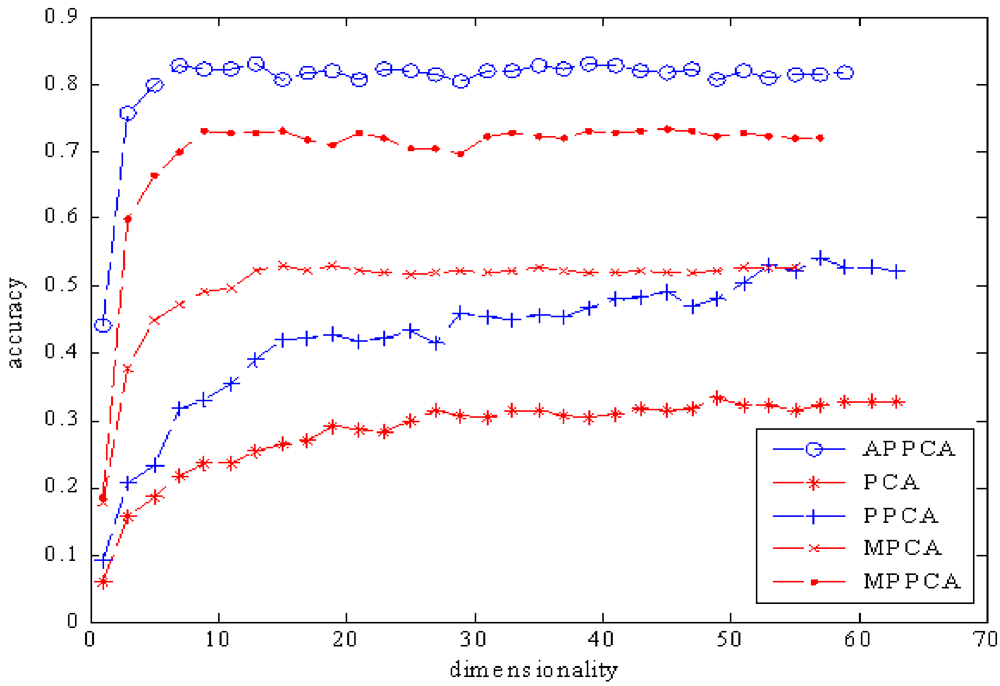

Fig. 8 Accuracies of various methods with respect to increase in dimensionality of the subspace

The concept of importance to locality is illustrated in figure 9. Three different graphs of accuracy vs. dimensions for, holistic PCA, PCA on randomized pixels and Modular PCA are shown in the figure. The pixels in the image are indexed row wise using consecutive numbers from 1 to 4096. These numbers from 1 to 4096 are then randomized to form a random number set. Now the positions of the pixels are changed according to the random number set. The same randomized set of numbers is used to re-arrange the pixels in each of the training images. Figure 10 shows the sample training image and the corresponding randomized image.

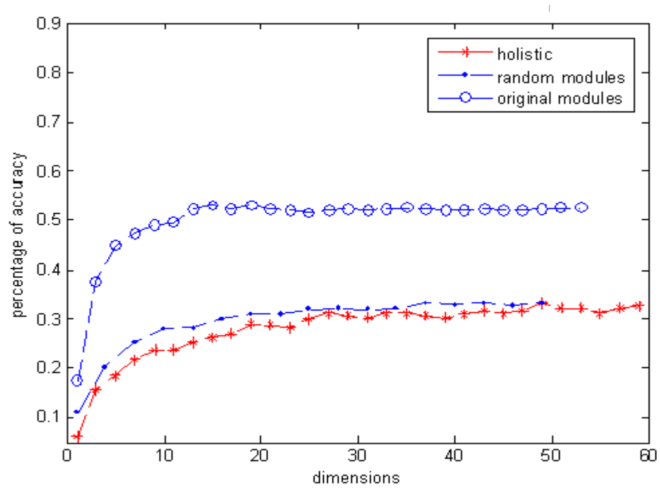

Fig. 9 Comparison of accuracies for holistic, randomized pixel set based modular approach and conventional non-overlapping modular techniques

The same randomized array is also used on the probe image and then MPCA is applied. The result is that the graph representing the accuracy against dimensionality of the subspace almost overlaps that of holistic PCA.

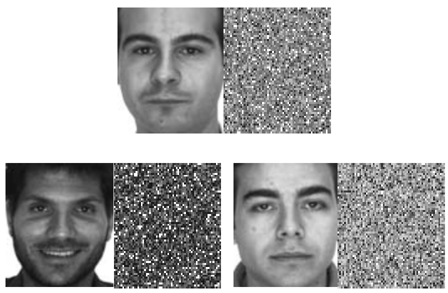

Fig. 10 Original image and the corresponding randomized pixel position images of three different individuals

This indicates that although the randomized pixel images contain all the pixels as the original, due to randomization of the facial variations are being spread across the image, as opposed to confining to a local region. Hence modularization does not have any effect. Thus, based on the assumption that the variations in face images are local, creating more features from within the defined neighborhood, does improve the classification ability and it also suggests that, the importance of dependency between pixels decreases with increase in the distance between them

Figure 10 illustrates the arrangement of modules with different criterion functions. Only the conventional non-overlapping regions of 8×8 pixels are used in this case. It can be observed that arranging the modules in the order of the proposed criterion function yields an accuracy of a conventional method with a lower number of modules.

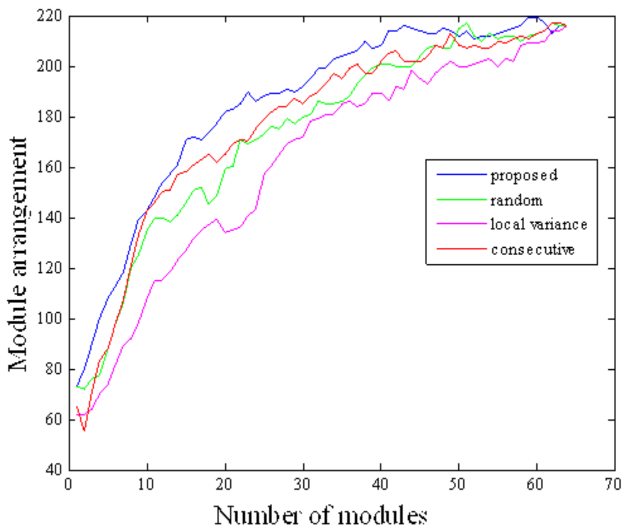

Fig. 11 Accuracy results for different types of module arrangements

The worst result is observed to be that of using the local variance as the criterion for arrangement. The highest accuracy of 83% has been achieved with just 270 modules as compared to more than 460 modules required to achieve the same result using random arrangement of the modules. This also shows that the most important dependencies among the pixels are the local dependencies.

VII. Conclusion

The paper presented a technique of face recognition aimed at improving the recognition accuracies of the faces that are affected due to varying illuminations, partial occlusions and varying expressions. Modular subspaces on phase congruency images are used in this technique. It has been shown empirically, that the type of feature selection based on the locality of the modules from the predefined neighborhood, improves the recognition accuracy. Arranging the modular spaces according to the criterion function, which is based on the discriminating capability of each of the subspace, helps in achieving the same accuracy with less number of modules. This is extremely useful in improving the computational efficiency of the face recognition algorithm.

References

[1] M. Turk, A. Pentland,"Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, pp. 71–86, March 1991.

[2] A. Pentland, B. Moghaddam, T. Starner, "View-based and modular eigenspaces for face recognition," Proceedings of the *IEEE Computer Society Conference on Computer Vision an Pattern Recognition*, 1994, pp.84-91.

[3] W. Zhao, A. Krishnaswamy, R. Chellappa, D. Swets, L. Weng, "Discriminant analysis of principal components for face recognition," *Int. Conference on Automatic Face and Gesture Recognition*, pp.336–341 March 1998.

[4] J. Huang, P. Yuen, C. Chen, W. Sheng, J.H. Lai, "Kernel subspace LDA with optimized kernel parameters on face recognition," *IEEE International conference on Automatic Face and Gesture Recognition*, 2004, pp. 327-332.

[5] A. Martinez "Recognition of partially occluded and/or imprecisely localized faces using a probabilistic approach," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.712–717, Jan 2000.

[6] Y. Zhang and A.M. Martinez "Recognition of Expression Variant Faces Using Weighted Subspaces," *Proceedings of International Conference on Pattern Recognition*, 2004.

[7] P. Kovesi "Edges Are Not Just Steps," *The 5th Asian Conference on Computer Vision*, pp. 23-25, January 2002.

[8] Rajkiran Gottumukkal, Vijayan K. Asari "An improved face recognition technique based on modular PCA approach," *Pattern Recognition Letters*, pp. 429-436, 2004.

[9] K. Tan, S. Chen "Adaptively weighted sub-pattern PCA for face recognition," *Neurocomputing*, v 64, pp.505-511, March 2005.

[10] A.U. Batur,"Segmented linear subspaces for illumination-robust face recognition", *International Journal of Computer Vision*, v 57, n 1, April, 2004, p 49-66

[11] H. Wang, Z. Stan Li, Y Wang, "Face recognition under varying lighting conditions using self quotient image", *Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, 2004, p 819-824.

[12] K.C Lee, J. Ho, D.J. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v 27, n 5, May, 2005, p 684-698

[13] A.V. Oppenheim, J.S. Lim,"The importance of phase in signals", *IEEE Proceedings*, v 69, May 1981, pp 529-541.

Satyanadh Gundimada is a PhD student in Electrical and Computer Engineering at ODU. Satyanadh received the Bachelor's degree in Electronics & Communication Engineering from Osmania University, India, in 2001, the Master's degree in Electrical Engineering from ODU, in 2003.

Vijayan Asari is an Associate Professor in Electrical and Computer Engineering at Old Dominion University, Norfolk, Virginia. He is the Director of the Computational Intelligence and Machine Vision Laboratory (ODU Vision Lab) in the ECE department. Dr. Asari received the Bachelor's degree in Electronics & Communication Engineering from the University of Kerala (College of Engineering, Trivandrum), India, the M. Tech and Ph. D degrees in Electrical Engineering from the Indian Institute of Technology, Madras. His research activities are in the areas of image processing, computer vision, pattern recognition, neural networks, and digital system architectures for application specific integrated circuits. Dr. Asari is a senior member of the IEEE.