Submitted to the Proceedings of the 2nd International Conference on Information Technology for Application (ICITA 2004)

Application of the DCT Energy Histogram for Face Recognition

Ronny Tjahyadi, Wanquan Liu, Svetha Venkatesh
Department of Computer Science, Curtin University of Technology,
GPO Box U1987, Perth WA 6845, Australia
{tjahyadi, wanquan, svetha}@computing.edu.au

Abstract-- **In this paper, we investigate the face recognition problem via energy histogram of the DCT coefficients. Several issues related to the recognition performance are discussed, in particular the issue of histogram bin sizes and feature sets. In addition, we propose a technique for selecting the classification threshold incrementally. Experimentation was conducted on the Yale face database and results indicated that the threshold obtained via the proposed technique provides a balanced recognition in term of precision and recall. Furthermore, it demonstrated that the energy histogram algorithm outperformed the well-known Eigenface algorithm.**

I. Introduction

Research on face recognition remains highly popular in computer vision literature. Over previous decades, many attempts have been made in the pattern recognition field to mimic the inherent human ability to recognize faces. Many facial representation approaches have been developed (statistical-based, neural networks and feature-based) with the purpose of recognizing individual faces, given various viewing conditions [1, 2]. Many of these techniques have already been utilized in society through Automated Teller Machines, building access systems, law enforcement, airport surveillance and computer login systems [3].

One successful approach to face recognition uses Principal Component Analysis (PCA), originally proposed by Sivorich and Kirby [4]. PCA is an optimal signal representation that extracts the Eigenvectors and Eigenvalues from a covariance matrix constructed from an image database. This technique reduces the number of dimensions to represent images in the database and computes the basis of the space represented by its training vectors. In 1991, Turk and Pentland [15] incorporated PCA into a face recognition system known as Eigenfaces, demonstrating promising results in recognizing frontal images of individuals.

Despite the optimal signal representation produced with PCA, the technique suffers from high computational cost in determining the basis space for a large number of images [6]. In addition, computational cost for PCA increases when adding new images to the facial image database as the basis space for PCA requires recompilation due to the strong data dependency exhibited [5].

The Discrete Cosine Transform (DCT) has been used for feature extraction and has been demonstrated to be superior to PCA due to the lack of data dependency, hence recompilation is not required when adding or removing new images to or from the facial image database [8]. In addition, only a subset of the transform coefficients are needed to preserve the important features of the face images such as hair, eyes, mouth outline and mouth location [7]. Face recognition algorithms incorporating DCT can be found in [7, 8, 9, 10, 11, 12].

Among all these techniques, threshold selection was not implemented thoroughly in the recognition systems. Threshold selection for facial classification is crucial to the performance of the recognition system. As the threshold value increases, the number of individuals that are correctly and incorrectly accepted as known is also increase. Hafed and Levine [8] discussed the Receiver Operating Characteristic (ROC) curve method in finding the classification threshold. An ROC graph illustrated the trade off between correct classifications versus false acceptances where the ROC curve was obtained using a subset of the CIM database. This method as discussed; however, was not implemented in the facial recognition system.

In this paper, we discuss an alternative approach for feature extraction using the energy histogram of the DCT coefficients. The performance of various feature sets with a range of histogram bins were investigated using the Yale face database. In addition, the threshold selection method was discussed which was based on the ROC analysis of the distances gathered from the face database. The performance of the energy histogram face recognition with the selected threshold was analyzed and discussed. To determine the effectiveness of the technique, a comparison was also made with the Eigenface algorithm.

II. Preliminary Background

A. The Discrete Cosine Transform

The DCT is a popular technique in imaging and video compression, which was first applied in image compression in 1974 by Ahmed *et al.*, transforming signals in the spatial representation into a frequency representation. In 1992, the first international standard for image compression, known as the Joint Photographic Experts Group (JPEG), was established with the DCT as the encoder and decoder. JPEG compression uses the DCT to remove the redundancies from images. Each image frame is divided into 8x8 blocks, where each block is transformed independently using the 2D-DCT basis function. The forward 2D-DCT of an 8 x 8 block image is defined by

Forward:

$$F(u,v) = \frac{2}{N} C(u) C(v) \sum_{x=0}^{7} \sum_{y=0}^{7} f(x,y) \cos\frac{(2x+1)u\pi}{2N} \cos\frac{(2y+1)v\pi}{2N}$$

where x and y are spatial coordinates in the image block, and u and v are coordinates in the DCT coefficients block. The C terms are defined as:

$$C(x) = \begin{cases} \frac{1}{\sqrt{2}} & for\ x = 0 \\ 1 & otherwise \end{cases}$$

Figure 1 shows the properties of the DCT coefficients in 8x8 blocks with the zigzag pattern used by JPEG compression to process the DCT coefficients. Although the total energy remains the same in the 8x8 blocks, the energy distribution changes with most energy being compacted to the low frequency coefficients. The DC coefficient is represented by $F(0,0)$ in the forward 2D-DCT equation. As the cosine of zero is one, the equation is simplified to

$$F(0,0) = \frac{1}{8} \sum_{x=0}^{7} \sum_{y=0}^{7} f(x,y)$$

The DC coefficient, which is located at the upper left corner, holds most of the image energy and represents the proportional average of the 64 blocks. The remaining 63 coefficients denote the intensity changes among the block images and are referred to as AC coefficients.

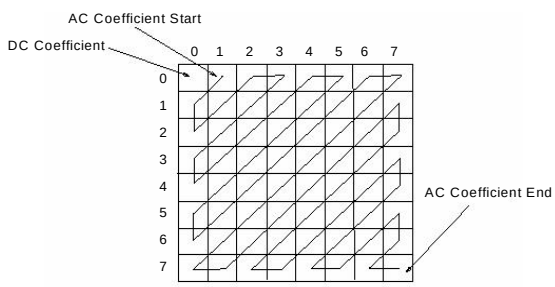

Fig. 1. Block features of the DCT coefficients with the zig-zag pattern.

The DCT was reported to be the second best transformation after PCA in [14] with an energy compaction that closely approximates that of PCA. Although PCA is the optimal transform in an energy packing sense, most practical transform coding systems still apply the DCT. It offers numerous advantages over PCA including producing good quality images at suitable compression ratios and the ability to perform in real time situations due to its computational efficiency. Furthermore, the DCT is easy to implement in both hardware and software, making it a more preferable and affordable technique in image and video applications, compared to PCA.

B. *Energy Histogram*

Histogram analysis is a popular method used in the image retrieval field, which was first introduced, by Swain and Ballard [15] via the use of color histograms. A histogram consists of multiple bins where each bin corresponds to a range of values. A color histogram of an image is obtained by discretising the colors within the image and counting the number of times each color occurs in the domain of the corresponding bin.

The advantages of color histograms are described in [15]. Histograms are invariant to image manipulations such as rotations, translations and they also change slightly with a change in scale, angle of view or with occlusion. Despite these advantages, histograms perform poorly under different imaging or lighting conditions. They are also ineffective in distinguishing different images that have similar color distributions and suffer with inefficient computation due to their dimensionality. Many techniques have been introduced to solve these limitations as described in [16].

An energy histogram is similar to color histograms but instead of counting pixel color, an energy histogram counts the occurrence of the DCT coefficients in the corresponding bins. In comparison, energy histograms incur less computational cost compared to the traditional histogram and they are preferable techniques over traditional histograms as their dimensions are greatly reduced by DCT. An image with N x N pixels requires N^2 dimensions, which is computationally expensive to be applied with the traditional histograms. However, energy histograms are applied in much lesser dimensions as the image redundancies have been removed by DCT. The implementation of energy histogram algorithm in image retrieval has been discussed in [17].

III. SYSTEM DESIGN

A. *Face Recognition via Energy Histogram*

The implemented face recognition system includes feature extraction, computed from the energy histogram in the DCT Coefficients, and recognition classification, which recognizes images based on their feature vectors using the Euclidian distance classifier. Figure 2 shows the face recognition model, incorporating the feature extraction and recognition classification.

The DCT in the energy histogram is implemented through a similar approach as in the JPEG compression. The approach in feature extraction computes the DCT on individual subsets of each facial image where images are first divided into 8x8 locks. If the last few blocks do not fit into an 8x8 block, they are padded with zero. The Forward 2D-DCT is then applied to each block, resulting in 8x8 DCT coefficient blocks.

As recognition speed is an issue, expensive computations resulting from using the whole DCT coefficients were avoided by retaining only certain DCT subsets, which are shown in Figure 3. These subsets were chosen as square-like features and retain the higher energy level of the image. The F1 feature set contains only the DC component, whilst F2, F3 and F4 resemble the 2x2, 3x3 and 4x4 coefficient blocks at the upper left corner of the DCT coefficient block. The DC coefficient is included in all features, as it directly resembles the majority of the image blocks energy. Once the features have been selected, the energy histogram is built and used as the feature vector. When the feature vectors of the image database have been created, the

next step is to select the threshold using the algorithm specified in the next section.

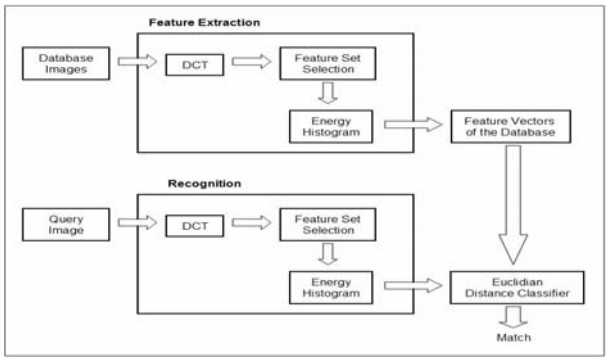

Fig. 2. Face Recognition System with Energy Histogram.

$F1 = [DC]$,
$F2 = [DC, AC_{01}, AC_{10}, AC_{11}]$,
$F3 = [DC, AC_{01}, AC_{02}, AC_{10}, AC_{11}, AC_{12}, AC_{20}, AC_{21}, AC_{22}]$,
$F4 = [DC, AC_{01}, AC_{02}, AC_{03}, AC_{10}, AC_{11}, AC_{12}, AC_{13}, AC_{20},$

Fig. 3. The feature sets.

To recognize a face image, the system compares the image's feature vector to each of the feature vectors in the database using the Euclidian distance classifier. If the feature vector of the query image is v and that of the database is f, then the Euclidian distance between the two is defined by:

$$d(v-f) = \left(\sum_{i=1}^{n} \left\| h_i^v - h_i^f \right\|^2 \right)^{1/2}$$

where d(v-f) represents the distance between feature vector v and f, h_i represents the number of frequency occurrences in bin i, where $i = 1, 2, \ldots, n$ and n denotes the number of histogram bins. If the distance between v and f falls below a threshold, then the images are classified as a matched; otherwise, it is an unmatched.

B. Adaptive Threshold

Query effectiveness is evaluated using precision and recall statistics. Precision measures the percentage of the recognized face images that are relevant, whilst recall measures the percentage of the relevant images that are recognized. An accurate face recognition system maximizes both precision and recall rates. However, in practice, these metrics generally contradict each other. The degree of importance of recall versus precision depends on the specific application. A maximum precision is required for preventing the false positives where the face images are incorrectly classified [18].

Recall and precision rates correlate to the chosen threshold value. Identifying this value is normally done through ROC curve analysis [8, 19] and trial and error. As precision and recall rates always oppose each other, a trade off is made when selecting the threshold. Recognition with a large threshold value generally results in low precision and high recall rates, whilst a small threshold value produces high precision and low recall rates. In this section, we will show how to select a threshold value by analyzing the features of image database. A balanced precision and recall rates will be obtained through the selected threshold value.

ROC curve analysis is a method to compare classifiers on natural datasets using accuracy [20]. The ROC graph shows the trade off between True Positive (TP) and False Positive (FP) and its derived ROC curve is plotted using the TP and FP values from the testing dataset for each given acceptance threshold value. The acceptance threshold values are computed using the images in the training dataset [19]. Figure 4 illustrates the ROC curve with break-even, good and near optimal performance. The optimal performance is obtained when the ROC curve approaching to the upper-left corner.

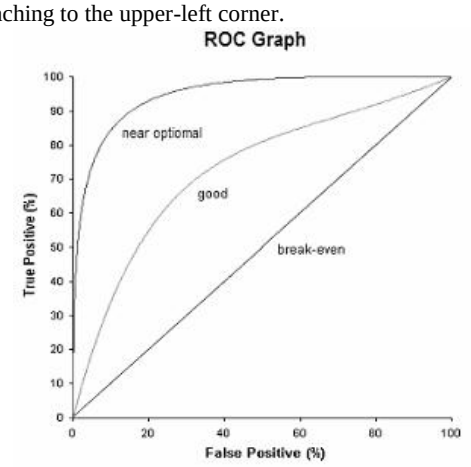

Fig. 4. The ROC Graph.

Threshold selection is based on the distance information gathered from the feature database. The feature database requires two or more images per individual as greater numbers of images contribute more distance information, hence may result in a better approximation for the threshold. Prior to the threshold selection algorithm, tolerated TP and FP values must be defined from the algorithm shown in Figure 5. These values are unique for each face database and are only recalculated when different face databases are used. The threshold selection is based on the tolerated TP and FP values; algorithm is described in Figure 6. The selected threshold is used for each feature from the Energy Histogram, regardless of its bin sizes.

The threshold algorithm is able to predict the performance of the face database on the given algorithm using TP and FP rates. As TP rates approach 100% and FP rates approach 0%, the given algorithm's accuracy is seen to perform near optimal. Separation of intra and inter classes is demonstrated using the selected threshold with intra class and inter class distances near zero and diverging away from zero, respectively. Both TP and FP require the database to have significant numbers of individuals (≥ 30). If the numbers of individuals are too few (<30), the number of intra class distances will also be too few in order to make accurate prediction on the threshold value. A database with n individuals in which each individual has m images, provides n numbers of distances in the intra class and $(n-1)*(m*n)$ numbers of distances in the inter class.

1. For a selected Energy Histogram feature, calculate 5 feature vector databases with histogram bin sizes of 10, 20, 30, 40 and 50. This range of histogram bin sizes was selected as the ROC curve was shown to be optimal within this range. For each feature vector database, calculate two distance classes from its feature vectors. The first is the *intra* class where the distances between feature vectors of the same individual are calculated. This class gives an indication of how similar the images of the same individual are. The other is the *inter* class, in which the distance between each feature vector of an individual is measured against feature vectors of other individuals. These distance values indicate how different each image of an individual is against images of other individuals in the database.
2. Draw the ROC graph for each feature vector database. The distances of the intra class are set as the acceptance threshold values. The ROC curve is plotted using the TP and FP values obtained from each acceptance threshold value. The TP rate is measured by calculating the percentage of the distances of the intra class that are allowed to be less than the acceptance threshold value. The FP rate is measured by calculating the percentages of the distances of the inter class that are allowed to be less than the acceptance threshold value. The TP refers to the face images that correctly matched as known, whilst the FP refers to the face images that incorrectly classified as known.
3. From all the ROC graphs of the feature vector databases, find the optimal ROC curve that shows high TP and low FP values. These values are specified through observation of the optimal ROC curve. They are defined realistically through a point near the ROC curve that gives a high TP value and lower FP value and are

Fig. 5. Algorithm to find the tolerated TP and FP values.

1. For each feature vector database i,
   $\exists$: $\{di_1, di_2, \ldots, di_n\} \in$ intra class i
   and $\{si_1, si_2, \ldots, si_m\} \in$ inter class i,
   where n = number of elements in intra class i and
   m = number of elements in inter class i.
   The *individual threshold* is selected depending on the two scenarios:
   a. If max{intra} < min{inter}
   $$\textit{individual threshold} = \frac{\max\{\text{intra}\} + \min\{\text{inter}\}}{2}$$
   b. If max{intra} $\geq$ min{inter}
   individual threshold is selected from a value in the inter class, that satisfies the conditions, where the TP value must be greater than or equal to its tolerated value and the FP value must be less than or equal to its tolerated value.
2. The minimum of the *individual threshold* values is used as the threshold for future classifications.

Fig. 6. *Individual Threshold* Algorithm.

The adaptive ability of the threshold selection is the result of the DCT's data dependency. Threshold selection can be recalculated after a user-defined number of database manipulations, including addition or deletion of database face images. When adding images to the face database, only the distances of the added images are calculated while the existing distance classes are still valid for selecting the classification threshold. Image deletion is similar to addition except that the distances of the deleted images are removed from the existing distance classes. Following the update to the distance class values, the process of threshold recalculation can then be immediately performed, as the feature vector of each image is independent.

IV. Experimental Results

These experiments were carried out on the Yale Face Database [21], containing of 165 images for 15 individuals with 11 images each. All of the facial images were taken in a frontal upright position with various facial expressions or configurations such as: center-light, left-light, right-light, with glasses, without glasses, happy, unexpressive, sad, sleepy, surprised, and winking. The training database was constructed from happy and sad face expressions of the first 10 individuals. The remaining facial images that were not included in the training database were used to construct the testing database. There were no preprocessing methods applied to enhance the facial images prior to the feature extraction. Face images with light configurations were excluded as the excessive light cast shadow on the background, which would require preprocessing in practice.

A. Threshold Selection

The training dataset consists of 20 face images from 10 individuals, hence, the number of distances in the intra class is 10 and the number of distances in the inter class is 180. In order to obtain realistic TP and FP values, we analyzed the ROC graphs. Figure 7 shows the ROC graph where the ROC curve was calculated from the distances gathered from images in the training database based on F2 feature set with histogram bin sizes of 10. The performance of the face recognition were considered to be near optimal as all of the ROC curves laid closer to the upper-left corner. The best performance was obtained with histogram bin of 10, where the 90% of TP were achieved with scarifying smaller amount of FP (11.67%).

The TP tolerated value was set to 90% as achieving 9 out of 10 images to be correctly classified was reasonable accurate. The FP tolerated value was consequently set to 5% based on the observation that allowing 9 out of 180 images to be misclassified was deemed reasonable.

Table 1 shows the *individual threshold* values of the F2 feature with 5 histogram bins required from 10 to 50 along with their corresponding True Positive and False Positive rates. The threshold for the training database with the F2 feature set was derived from the minimum of the *individual threshold* values (55.7). The same procedure was applied to the other feature sets (F1, F3 and F4) and datasets. Evaluation on the effectiveness of the threshold selection is described in the following section.

B. Evaluation on Threshold Selection and Energy Histogram Face Recognition

This section discusses the effectiveness of the energy histogram algorithm and the recognition performance of each feature set with 10 bin sizes ranging from 10 to 100. The figures below show the recognition performance under various feature sets.

The threshold values for the classification were calculated using the proposed threshold selection algorithm. Judging the recognition ability in term of precision and recall rates depends on the application. Some applications require higher precision rate, as it may be detrimental to misclassify an unknown individual. Our aim with the threshold selection is to find an optimal balance between precision and recall rates.

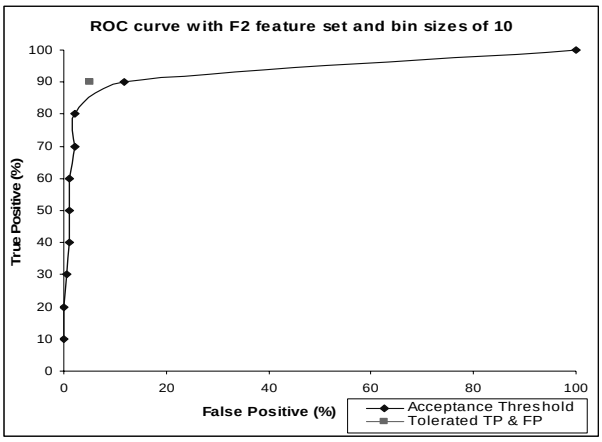

Fig. 7. ROC graph of F2 feature set with histogram bin sizes of 10.

Bin Size	*Individual Threshold*	TP (%)	FP (%)
10	57.8	90	2.2
20	55.7	90	0
30	77.8	90	4.4
40	60.2	90	2.2
50	69.1	90	5

Table 1: *Individual threshold* values for F2 feature under various bin sizes.

As shown in figure 8, threshold selection works well in the F1 feature set. The precision and recall rates are balanced regardless of the histogram bin size. Figure 9 shows that the recall rate in F2 feature sets is slightly reduced as the bin size increased. The threshold selection is still considered to perform well as the recognition rates are still higher in terms of precision and recall. At higher bin sizes, the threshold selection fails to obtain balanced recognition with feature sets F3 and F4. These feature sets are deemed not suitable for face recognition as they produce low recognition rates and the corresponding algorithm performs slower due to the increase number of features used.

The best recognition in the F1 feature set is obtained with a histogram bin size of 50 producing a precision rate of 94% and a recall rate of 85%. Similar recognition rates can be obtained with the F2 feature set using a histogram bin size of 20. Results further demonstrated that F2 outperforms the F1 feature set with a histogram bin size of 30 resulting in a precision rate of 96% and a recall rate of 90%.

C. Comparison between Energy Histogram and Eigenface

In the previous section, we demonstrated that optimal recognition was obtained using the Energy Histogram approach with the F2 feature set. Next, we implement the Eigenface algorithm [5] for a comparison. The trial and error method was used to select the threshold and eigenvectors. The threshold was

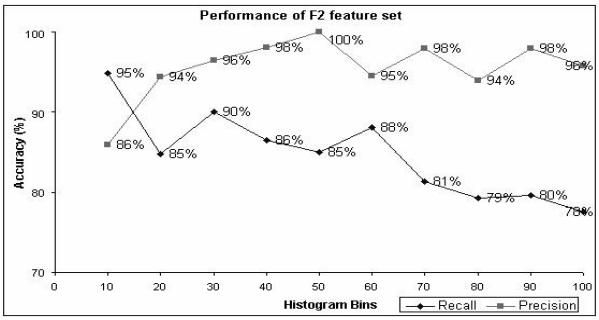

Fig. 8. Performance of F1 feature set.

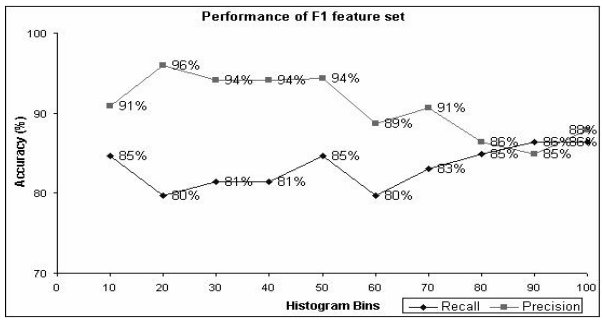

Fig. 9. Performance of F2 feature set.

chosen as 14.64 that gave balanced on the precision and recall rates while the first 10 eigenvectors were used for the recognition as it gave optimal recognition results. Figure 10 shows the comparison of the F2 feature set with the Eigenface algorithm. The F2 feature set with 4 bin sizes ranging from 20 to 60 outperformed the Eigenface algorithm producing a recognition rate higher in both precision and recall. As described earlier, optimal recognition with the Energy Histogram approach occurs when using the F2 feature with a histogram bin of 30. The precision and recall rates with this feature and histogram size are 96% and 90% respectively, whilst the Eigenface algorithm yielded only corresponding rates of 89% and 84%.

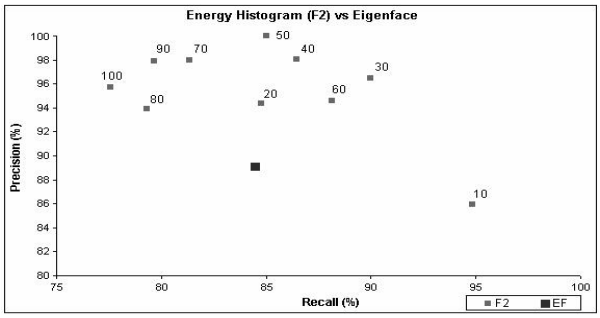

Fig.10. Comparison of the Energy Histogram with F2 feature set and Eigenface

V. CONCLUSION AND FUTURE WORKS

This paper has investigated the issue in selecting the classification threshold using the distance information obtained from the face database and has discussed an approach in feature extraction with the Energy Histogram of the DCT coefficients for face recognition. Some important issues related to the recognition performance are investigated; in particular the issue of histogram bin sizes and feature sets. Experimentation has been conducted on Yale face database. Results have shown that the selected threshold value provides a balance in precision and recall rates and the F2 feature set with a histogram bin size of 30 has yielded an optimal recognition performance. Furthermore, the Energy Histogram approach has shown to outperform the well-known Eigenface algorithm.

A number of possible improvements for future works have been identified through this research. The importance of image preprocessing prior to the application of an energy histogram has been shown to increase recognition performance accuracy in other face recognition techniques. Heseltine *et al.* [23] have presented an evaluation of image preprocessing techniques for Eigenface-based face recognition. The authors reported a significant improvement using simple image preprocessing techniques such as color normalization methods, statistical methods, convolutions methods, and above combinations. Sequential preprocessing using intensity normalization moments within local regions of the image and a contour convolution filter, has been reported to reduce the Equal Error Rate (EER) resulting in increased recognition accuracy with face recognition technique. All these techniques can be applied to the algorithm in this paper.

Secondly, we believe a combination of the Energy Histogram approach with other face recognition algorithms may further increase the recognition rate. Lu *et al.* [26] proposed classifier combination algorithms for face recognition. A combination of three common algorithms (PCA, Independent Component Analysis (ICA), Linear Discriminant Analysis (LDA)) with the sum rule and Radial Basis Function (RBF)-based integration strategies was investigated by Lu *et al.* using a face database containing 206 individuals (2,060 face images). The experimental results with two integration strategies outperformed each individual classifier.

REFERENCES

[1] R. Chellappa, C. L. Wilson and S. Sirohey, "Human and Machine Recognition of Faces : A Survey", *Proceedings of The IEEE*, Vol. 83, No. 5, May 1995.

[2] W. Y. Zhao, R. Chellappa, A. Rosenfeld and P. J. Phillips, "Face Recognition: A Literature Survey", *UMD CfAR Technical Report*, CAR-TR-948, 2000

[3] F. S. Samaria, "Face Recognition using Hidden Markov Models", PhD Thesis, University of Cambridge, 1994.

[4] L. Sirovich and M. Kirby, "Low-dimensional Procedure for the Characterization of Human Faces", *Journal of the Optical Society of America A*, Vol. 4, pp. 519-524, Mar. 1987.

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

[6] M. Kirby and L. Sirovich, "Application of Karhunen-Loève Procedure for the Characterization of Human Faces", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, Jan. 1990.

[7] Z. Pan and H. Bolouri, "High Speed Face Recognition Based on Discrete Cosine Transforms and Neural Networks", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 1999.

[8] Z. M. Hafed and M. D. Levine, "Face Recognition Using the Discrete Cosine Transform", *International Journal of Computer Vision*, Vol. 43, No. 3, pp. 167-188, 2001.

[9] M. Shneier, and M Abdel-Mottaleb, "Exploiting the JPEG Compression Scheme for Image Retrieval", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 8, Aug. 1996.

[10] S. Eickeler, S. Müller and G. Rigoll, "High Quality Face Recognition in JPEG Compressed Images", *IEEE Int. Conference on Image Processing*, Kobe, Japan, Oct. 1999.

[11] S. Eickeler, S. Müller and G. Rigoll, "Recognition of JPEG Compressed Face Images Based on Statistical Methods", *Image and Vision Computing Journal, Special Issue on Facial Image Analysis*, Vol. 18, No. 4, pp. 279-287, March 2000.

[12] A. V. Nefian and M. H. Hayes, "Hidden Markov Models for Face Recognition", *International Conference on Acoustics, Speech, and Signal Processing*, pp. 2721-2724, May 1998.

[13] N. Ahmed, T. Natarajan and K. Rao, "Discrete Cosine Transform", *IEEE Transactions on Computers*, Vol. 23, No. 1, pp. 90-93, 1974

[14] S. Wong, "Energy Compaction by Orthonormal Transform", *EE392c Project Report*, Dec. 1997, available [online]: http://ise0.stanford.edu/class/ee392c/demos/wong/.

[15] M. J. Swain and D. H. Ballard, "Indexing Via Color Histograms", *Proceedings of 3rd International Conference of Computer Vision*, Osaka, Japan, pp. 390-393, 1990.

[16] Y. Alp Aslandogan and C. T. Yu, "Techniques and Systems for Image and Video Retrieval", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 11, No. 1, pp. 56-63, 1999.

[17] J. A. Lay and L. Guan, "Image Retrieval Based on Energy Histograms of the Low Frequency DCT Coefficients", *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 6, pp. 3009-3012, 1999.

[18] J. Vogel and B. Schiele, "On Performance Characterization and Optimization for Image Retrieval", 7th *European Conference on Computer Vision*, Springer, pp. 49-63, 2002.

[19] G. L. Marcialis and F. Roli, "Fusion of LDA and PCA for Face Verification", *in Biometric Authentication*, vol. LNCS 2359, pp. 30-37, Jun. 2002.

[20] F. Provost, T. Fawcett and R. Kohavi, "The Case Against Accuracy Estimation for Comparing Induction Algorithms", *In Proceeding Fifteenth Intl. Conf. Machine Learning*, pp. 445-553, 1998.

[21] Yale University Face Database, available [online] : http://cvc.yale.edu/projects/ yalefaces/yalefaces.html

[22] T. N. E. Heseltine, N. E. Pears and J. Austin "Evaluation of Image Pre-processing Techniques for Eigenface Based Face Recognition" *2nd International Conference on Image and Graphics*, Vol. 4875, pp. 677-685 , 2002.

[23] T. E. Campos, R. S. Feris and R. M. Cesar Jr.; "Eigenfaces versus Eigeneyes: First Steps Toward Performance Assessment of Representations for Face Recognition"; *Lecture Notes in Artificial*