

A New Discrete Cosine Transform on Face Recognition through Histograms for an Optimized Compression

Zahraddeen Sufyanu, Fatma Susilawati Mohamad and Abdulganiyu Abdu Yusuf
Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin (UniSZA),
21300, Gong Badak Campus, Terengganu, Malaysia

*Corresponding Author: Zahraddeen Sufyanu, Faculty of Informatics and Computing,
Universiti Sultan Zainal Abidin (UniSZA), 21300, Gong Badak Campus, Terengganu, Malaysia Tel: +60146451708*

ABSTRACT

Images with reduced signal level that can be easily stored in a database for recognition purposes are highly demanded. A system with large capacity requires a highly compressive technique for its efficiency. In this study, more efficient image compression algorithm was disclosed which reduced the size of original image drastically. After the necessary preprocessing, the histogram plot of an image was obtained. This histogram was rotated by 45° to produce a reduced image. Therefore, the two-dimensional Discrete Cosine Transform (2DCT) was computed on the rotated histogram to produce a new matrix. However, the proposed framework was tested on ten subjects from Olivetti Research Laboratory (ORL) database and its performance was evaluated using compression parameters such as PSNR and MSE. The new framework results in better reconstruction error and better PSNR values than original JPEG compression level. The new DCT matrix is proposed to proceed for postprocessing in DCT domain, where several coding approaches are employed. It was believed that, this discovery is new in the field of signal processing.

Key words: Compression level, DCT, face recognition, rotated histogram

INTRODUCTION

Image compression is the application of data compression on digital images, meanwhile maintaining the quality of the images. The main purpose of image compression is to reduce redundancy and irrelevancy present in an image, so that it can be stored, transferred and recovered efficiently (Rani and Bishnoi, 2014). The DCT has become one of the most successful transforms in image processing for the purpose of data compression, feature extraction and recognition (Jing and Zhang, 2004; Amornraksa and Tachaphetpiboon, 2006). The transform tends to concentrate the information, making it useful for image compression applications (Jain, 1989; Pennebaker and Mitchell, 1993). It was categorized in the computer vision by Wang (1984), into four slightly different transformations named DCT-I, DCT-II, DCT-III and DCT-IV. It has been successfully used for face recognition in Kao *et al.* (2010), Dabbaghchian *et al.* (2010) and Cui and Zhang (2012a, b).

Furthermore, image compression algorithms are divided into two groups: Lossless and lossy compression (Pratt, 2001; Moerland, 2003; Ahmed, 2011). Lossless image compression techniques compress image without introduction of errors in the decompressed image. It does not give high

compression ratio but ensures no quality loss. Conversely, in lossy image compression techniques some amount of data are discarded. Although, the compression ratio achieved is generally higher but at the expense of image quality (Dutta *et al.*, 2012). The former occurred up to a certain level beyond which errors are introduced.

Intensity histogram provides information about the global appearance (whole pixel) of an image and it is always a solution to comparison of colors (Mohamad *et al.*, 2010). The histogram is used for image enhancement, segmentation or compression. The impact of image compression is shown when the dynamic range is reduced to only little intensity (quantization).

There are various compression techniques, some of which focused on high compression ratio, whereas some on better quality with appreciable compression ratio (Dutta *et al.*, 2012). Feature extracted from DCT was used for face recognition in Hafed and Levine (2001); the paper declared that, DCT gives the near-optimal performance of Karhunen-Loeve Transform (KLT) in facial information compression. However, many approaches have been proposed to boost the accuracy of lossy and lossless image compressions. Another research proposed by Dutta *et al.* (2012) was on compression algorithm for lossy image using histogram based block optimization and arithmetic coding where better compression ratio was reported. In another algorithm suggested by Ferreira and Pinho (2002), a preprocessing technique using histogram packing was discussed which cut down the approximation error by reducing the image variation to improve the lossless compression rates. Moreover, a hierarchical statistical Compressive Sensing (CS) inversion algorithm for wavelet-based and JPEG-based DCT was developed in He *et al.* (2010), where a Tree-Structure in the sparseness pattern was exploited. This yielded computational efficiency while exploiting the structure inherent to each construction.

All these researches and many more in the literature that used DCT especially on faces, concentrate more on extracting low frequency coefficients of the DCT matrix. Nevertheless, despite the energy compaction of the DCT, the acquired images may be too large when it comes to recognition of real life with large users. Hence, highly compressive technique is required even before the feature extraction. A closer approach investigated on the use of histogram recently and the idea of rotated histogram was used on face detection using Hough Transform (HT) in Zahraddeen *et al.* (2014a). This research brought another benefit upon Duda and Hart (1972) on the restricted use of Hough transform. It also expanded the idea of using HT in image processing. To the best of authors' knowledge, DCT has never been applied on the compressed original image.

In this study, a pioneering feature extraction through DCT and histogram is proposed into the literature. The idea is to discover if DCT can be applied on a reduced feature obtained from histogram and to find a suitable rotation angle of the histogram to obtain the appropriate feature representing original image of face. The goals of this framework are two-fold: First, to improve lossy compression techniques and promote recognition accuracy. Second, is to enable creation of a system with large number of users on a given amount of disk or memory space. As a result of that, the problem of data storage during enrollment is minimized. In addition, the proposed algorithm is very simple to implement and it gives better compression ratio.

IDEA OF THE PROPOSED SOLUTION

Figure 1 demonstrates the principle of the new DCT, where the image in spatial domain has to be converted to intensity histogram before the DCT is applied. The definition of the new DCT for an input image f and output image F can be written in Eq. 1.

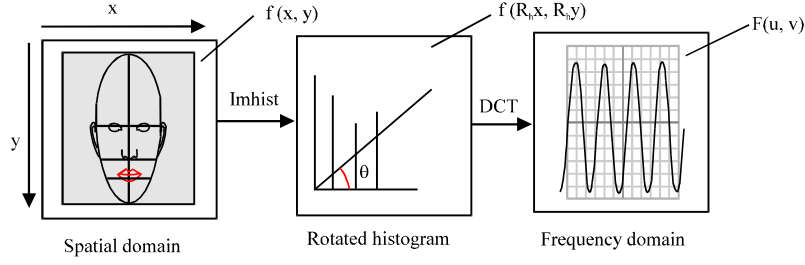


Fig. 1: Image transformation from spatial domain through histogram and to frequency domain

$$F(u, v) = C_{(u)} C_{(v)} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(R_h x, R_h y) \cos \left[\frac{(2R_h x + 1)u\pi}{2M} \right] \cdot \cos \left[\frac{(2R_h y + 1)v\pi}{2N} \right] \quad (1)$$

Where:

$$C_{(u)} = \begin{cases} \frac{1}{\sqrt{M}}, & u = 0 \\ \sqrt{\frac{2}{M}}, & 1 \leq u \leq M-1 \end{cases}$$

$$C_{(v)} = \begin{cases} \frac{1}{\sqrt{N}}, & v = 0 \\ \sqrt{\frac{2}{N}}, & 1 \leq v \leq N-1 \end{cases}$$

and:

$$R_h x = (x - \Delta x); R_h y = (y - \Delta y)$$

The ' Δx ' and ' Δy ' represent the changes induced in the spatial domain by the rotation of the histogram at angle 45° . Whereas ' $R_h x$ ' and ' $R_h y$ ' represent the reduced feature of the original image before computing the DCT.

The ' M ' and ' N ' are the row and column size (of rotated histogram) respectively which also varied according to change of the angle.

MATERIALS AND METHODS

ORL face database was used to test the proposed DCT. The images from this dataset were preprocessed accordingly. And samples of one subject are shown in Fig. 2. In addition, MATLAB R2012b version 8.0 was used for simulation.

Histogram as a reduction technique: The range of pixel intensities of the preprocessed images was stretched (i.e., contrast stretching), the normalized images were converted into histogram using histogram plot function which produce distributed histograms that are easier to extract the features. However, the histogram plot was rotated by 45° to obtain the reduced-sized images of approximately squared dimensions and that reduces complexity during recognition. Figure 3 shows the process of histogram conversion which is an angle-based.



Fig. 2: Samples of ORL face (with varieties of poses in this database)

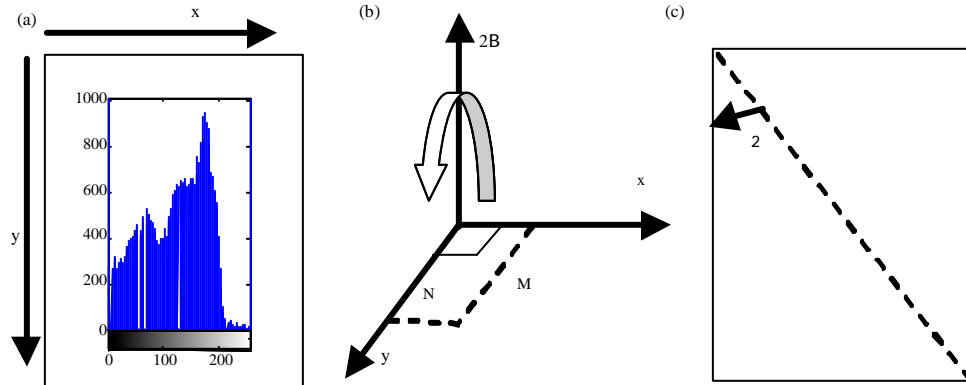


Fig. 3(a-c): Histogram conversion process

Proposed feature extraction using the new DCT: Generally, the upper left corner of the DCT matrix represents the low frequency. This section is related to illumination variations and smooth regions such as forehead and cheeks of the face. Reducing the values of this coefficients diminish the image quality. Conversely, the coefficients situated at the bottom right corner of the matrix, represent the high frequency. This section represents the noise and detailed information about the edges in the image. While the mid frequency region coefficients represent the general structure of the face in the image. In this study, the DCT is applied on the rotated histogram which gives a DCT matrix with different coefficient structure. The coefficients are evenly distributed into two blocks. The white blocks represent the illumination and detailed information of the faces. Whereas the dark blocks represent the edges of the rotated histogram. The block 8×8 DCT is recommended, the reason for choosing a block size of 8×8 is to have enough sized blocks to provide sufficient compression, besides keeping the transform simple (Ekkenel and Stiefelhagen, 2005). In to this study, the DC

component and all AC coefficients containing the highest information will be extracted via zig-zag scan to capture the details of the image. The coding will be achieved by selecting up the $F(u,v)$ coefficients in a zig-zag scanned order and then encoding the “Run- level” pairs of none zero $F(u,v)$. Bit allocation based on the threshold coding of DCT coefficients will be adopted. Finally, suitable distance measure classifier will be used for matching to report the performance of the algorithm on ORL face database images. The framework of the overall encoding scheme using threshold coding is depicted in Fig. 4. Meanwhile, Fig. 5 illustrates the representation of the 8×8 DCT blocks. The block-DCT sample of the process to be adopted is shown in three regions: ‘a’, ‘b’ and ‘c’, through dividing the images into small blocks and then taking the DCT of each sub image (Hafed and Levine, 2001). This idea is described in that Fig. 5 where ‘a+b+c’ regions represent the white diagonal blocks in forward direction. The recognition through these regions is beyond the scope of this study.

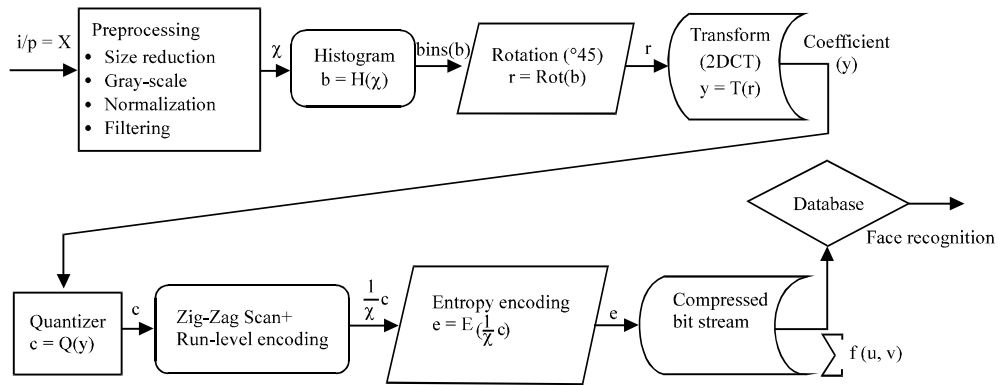


Fig. 4: Proposed block diagram of the new DCT

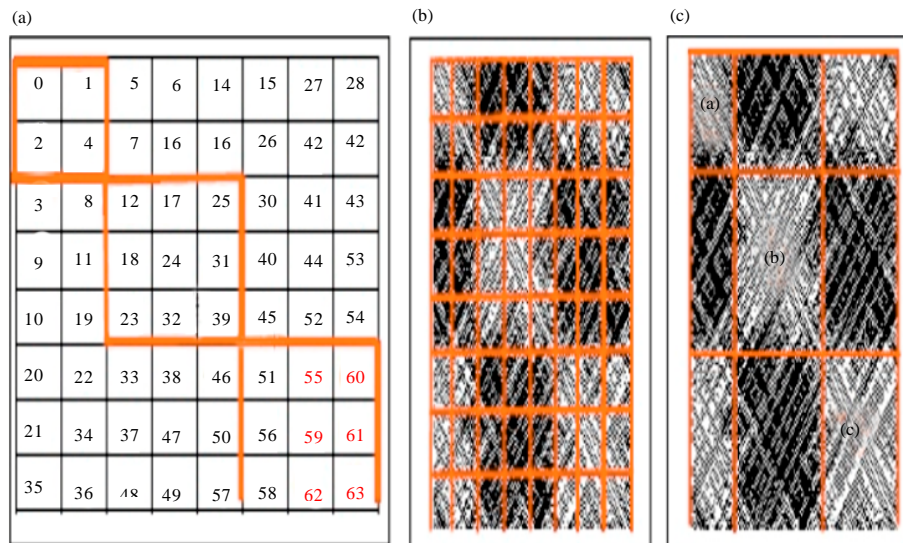


Fig. 5(a-c): 8×8 DCT matrix (new DCT) with regions ‘a, b, c’ containing the most important details

EVALUATION OF THE PROPOSED METHOD

To test the efficiency of the new compression technique, error Metrics are considered.

Relative reconstruction error: The Tree-Structured JPEG-based DCT transform was used, within Compressive Sensing inversion which employs Markov Chain Monte Carlo (MCMC) inference, similar to He *et al.* (2010) to check the reconstruction error of the original images and that of equivalent rotated histogram images. Additionally, Fig. 6 shows one of the results obtained, whereas Table 1 summarizes these results. Figure 7 is the representation of the results in line graph acquired through 40 iterations using 32×32 images. The reconstruction errors were plotted against number of subjects, with the results arranged in ascending order to simply check the errors variations. The error in the new DCT is quite lower than the existing DCT using the original image.

Peak signal to noise ratio (PSNR) and mean square error (MSE): Another error metrics used to compare the various image compression techniques are the MSE and the PSNR. The PSNR mathematical relation in Eq. 3 is a measure of the peak error, whereas MSE in Eq. 2 is the cumulative squared error between the compressed and the original image.

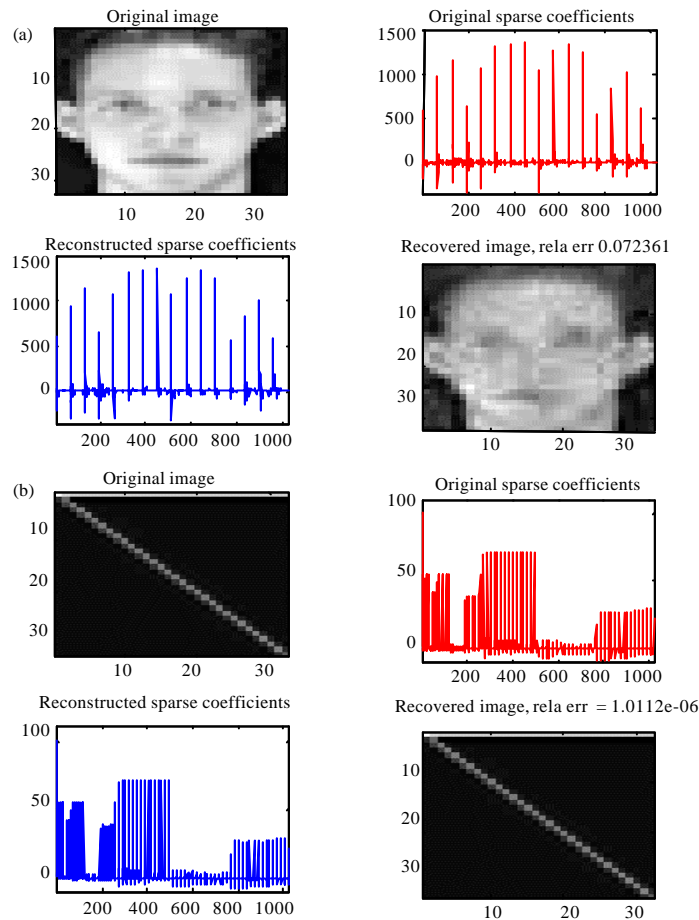


Fig. 6(a-b): Results of 2-dimensional 32×32 images acquired using the (a) Original image and (b) Equivalent rotated histogram

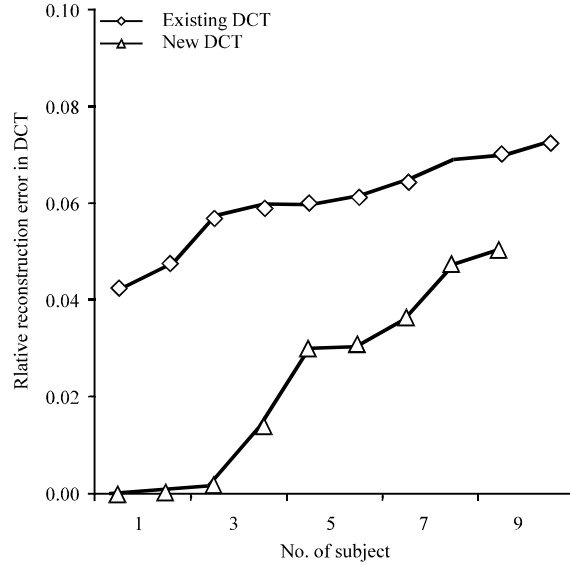


Fig. 7: Line graph representation of the relative reconstruction error per individual

Table 1: Relative reconstruction errors using ten subjects from ORL database

S/N	Existing DCT	New DCT
1	0.042	6.65 e ⁻⁶
2	0.047	1.01 e ⁻⁶
3	0.057	1.48 e ⁻³
4	0.059	1.37 e ⁻²
5	0.060	3.00 e ⁻²
6	0.061	3.05 e ⁻²
7	0.064	3.57 e ⁻²
8	0.069	4.67 e ⁻²
9	0.070	5.00 e ⁻²
10	0.072	6.88 e ⁻²

Table 2: JPEG compression parameters between existing and new DCT

Parameters	Methods	
	Existing DCT	New DCT
Average MSE	0.991	0.154
Average PSNR (dB)	52.978	61.069
Average RSME	0.995	0.392
Average Mean absolute error	1.430	0.211

$$MSE = \frac{1}{MN} \sum_{y=1}^N \sum_{x=1}^M [I(x,y) - I'(x,y)]^2 \quad (2)$$

$$PSNR = 20 \times \log_{10} (255/\sqrt{MSE}) \quad (3)$$

Therefore, in Eq. 2 and 3 the lower value for MSE indicates higher PSNR and the lesser the error. Hence, better compression achieved. The RMSE and Mean Absolute Error were also measured. The evaluation of these metrics was recorded in Table 2 showing the values for original images (existing DCT) and that of rotated histogram representing the faces (new DCT).

The influence of the approach in Fig. 6 using compressive sensing algorithm is to improve reconstruction accuracy. The relative reconstruction error outlined in Table 1, were measured between the original sparse coefficients and the reconstructed sparse coefficients. This gives more accurate results compared to measurements using PSNR. The error metrics PSNR and MSE here measured the original images against the compressed images using JPEG standard. The quality of this result is far less the former. Both the results are in favor of the new DCT and sufficient to prove that, the new DCT through histograms is more efficient than the conventional lossy compression techniques in the literature.

RESULTS AND DISCUSSION

Figure 8 displays three results of the new transform. The first row are the normalized images obtained using stretchlim function at 0.05 and 0.95 (minimum and maximum) pixel points. The

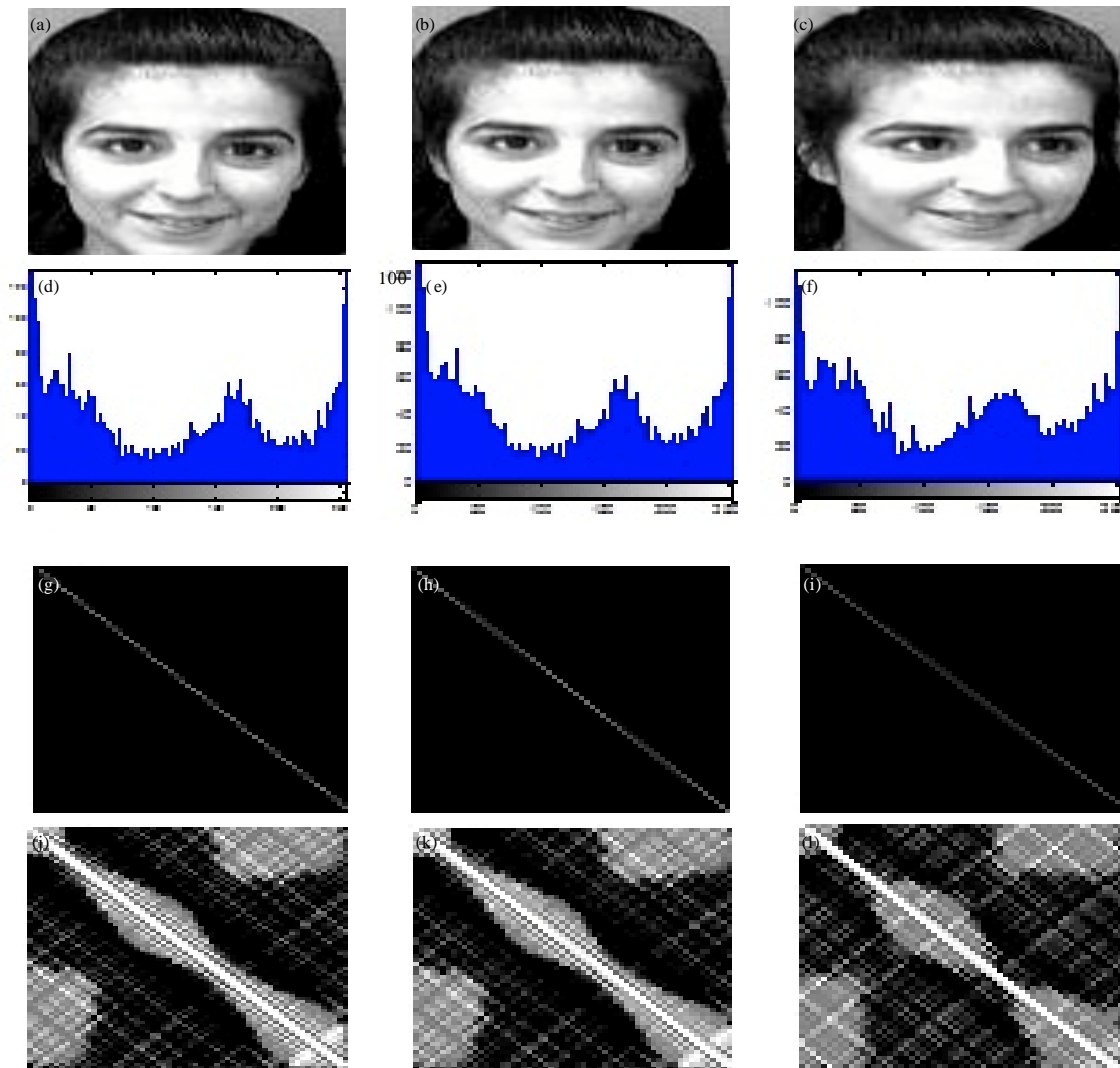


Fig. 8(a-l): Some of the results of the new DCT applied on the rotated histogram

second row shows histogram plots of the normalized images obtained through `imhist` function. And the third row displays the equivalent rotated histograms obtained through `imrotate` function at an angle of 45° . Whereas the fourth row are the corresponding DCT matrices obtained through the DCT2 function.

The rotation was tested from 0° at a step of 15° until an approximately squared image of symmetrical structure was obtained at 45° . The relative reconstruction error significantly reduced in the rotated histogram image compared to that of original image. Furthermore, the following reasons describe the advantages of compressing the original image in this study: The size of the rotated histogram is 2.6 KB while that of the original image is 7.2 KB. This signifies the optimized data. Moreover, each matrix shows significant decorrelation with respect to one another. Therefore, even extracting the 8×8 blocks as shown in Fig. 5 produce an efficient feature matrix during recognition, let alone few blocks from these coefficients. This shows that, the efficiency of image compression algorithm depends on how best and swift the redundancies can be exploited which may be through reduction of image size without any appreciable loss of quality (Dutta *et al.*, 2012). The results in Fig. 7 demonstrate lower relative reconstruction error in the new DCT compared to the existing one. In the same vein, Table 2 concluded that, DCT can be applied on a reduced feature obtained from histogram, at a specified angle with lesser MSE values and higher PNSR values. This further signifies that the error of the reduced images obtained from histogram is insignificant; the reconstruction can be achieved at high quality using the new DCT. Therefore, since dimensionality reduction using the histogram have moved the variables from the original coordinate system to the reduced coordinate system, it ascertains that the resulting models will require fewer parameters and faster to learn (Prince, 2012). Although feature extraction is time consuming but easily achieved using this approach since the required features obtained at 45° were very small. However, the new features create massive storage ability on a given amount of disk; this proves the objectives of the study.

In addition, this framework strongly improves lossy compression techniques, where images need not to be exactly reproduced as loss of information within certain levels is acceptable (Ahmed, 2011). It will also promote the recognition accuracy and many images can be stored in a given amount of disk or memory space. On the other hand, there is an increased computation time during histogram conversion process. Although the DCT matrices obtained via this framework showed significant decorrelation with one another but we are afraid of having high False Accept Rate (FAR) since the initial normalization affects the matrices. This can be considered unexpected finding in this paper. However, the normalization effect cannot be concluded until we extract the coefficients of the new DCT and report the recognition rates. We might decide to avoid the first normalization and apply the Default Normalization Matrix in JPEG or use the benchmark normalization procedure reported in Zahraddeen *et al.* (2014b) using Gabor filtering. This will be fairly justified in another paper following this research.

CONCLUSION

A new feature extraction frame work using DCT through histogram was introduced into the literature. The aim for improving the compression ability of the DCT using the reduced feature obtained from histogram was achieved. The study advances the research in this field by applying the DCT on the extracted features of faces. It was reported that, the relative reconstruction error of the rotated histogram representing the face was amply reduced compared to the errors encountered in the original image. Hence, this concept optimized the compaction level in lossy

compression techniques. In addition, it creates massive storage ability on a given amount of disk or memory space; this is the uniqueness of the study. The resulted matrix showed significant decorrelation with other images. However, the performance of the proposed algorithm was evaluated within a CS inversion algorithm of the Markovian structure used in He *et al.* (2010) and other compression parameters. In the subsequent studies, we are optimistic that, this technique will report appreciable recognition rates using the compressed bit stream of the optimized DCT. It is hoped to implement the new DCT of the entire image using threshold coding and perform the Block-based coding approach to extract few regions for more optimized DCT. The result will be compared with many hybrid systems in the literature. But it will be ensured that, the quality of the images are preserved by closely measuring the accuracy of the subsequent recognition system using ROC and CMC curves considering many subjects. Finally, the undesirable property occurred using the contrast stretching for enhancement, will be justified by selecting the most suitable preprocessor out of the 22 illumination normalization techniques in the literature.

ACKNOWLEDGMENT

Z. Sufyan thanks Isyaku Hassan, Faculty of Languages and Communication of Universiti Sultan Zainal Abidin (UniSZA) Malaysia, for minimizing the grammatical errors in this manuscript. And also thanks his mentor Engr. Bashir Muhammad, Department of Electrical Engineering, Universiti of Technology Malaysia (UTM).

REFERENCES

- Ahmed, B.K., 2011. DCT image compression by run-length and shift coding techniques. *J. Univ. Anbar Pure Sci.*, Vol. 5.
- Amornraksa, T. and S. Tachaphetpiboon, 2006. Fingerprint recognition using DCT feature. *Elect. Lett.*, 42: 522-523.
- Cui, P. and R.B. Zhang, 2012a. A semi-supervised coefficient selection method for face recognition. *J. Harbin Eng. Univ.*, 33: 855-861, (In Chinese).
- Cui, P. and R.B. Zhang, 2012b. Face feature extraction method based on part of labeled data. *J. Optoelectron. Laser*, 23: 554-560, (In Chinese).
- Dabbaghchian, S., M.P. Ghaemmaghami and A. Aghagolzadeh, 2010. Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology.. *Pattern Recognit.*, 43: 1431-1440.
- Duda, R.O. and P.E. Hart, 1972. Use of the Hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15: 11-15.
- Dutta, S., A. Abhinav, P. Dutta, P. Kumar and A. Halder, 2012. An efficient image compression algorithm based on histogram based block optimization and arithmetic coding. *Int. J. Comput. Theory Eng.*, 4: 954-957.
- Ekenel, H.K. and R. Stiefelhagen, 2005. Local appearance based face recognition using discrete cosine transform. *Proceedings of the 13th European Signal Processing Conference*, September 4-8, 2005, Antalya, Turkey.
- Ferreira, P.J.S.G. and A.J. Pinho, 2002. Why does histogram packing improve lossless compression rates? *IEEE Signal Proces. Lett.*, 9: 259-261.
- Hafed, Z.M. and M.D. Levine, 2001. Face recognition using the discrete cosine transform. *Int. J. Comput. Vision*, 43: 167-188.
- He, L., H. Chen and L. Carin, 2010. Tree-structured compressive sensing with variational bayesian analysis. *IEEE Signal Proces. Lett.*, 17: 233-236.

- Jain, A.K., 1989. Fundamentals of Digital Image Processing. Prentice Hall, Englewood Cliffs, NJ., ISBN-10: 0133361659, Pages: 569.
- Jing, X.Y. and D. Zhang, 2004. A face and palmprint recognition approach based on discriminant DCT feature extraction. *IEEE Trans. Syst. Man Cybernetics-Part B: Cybernetics*, 34: 2405-2415.
- Kao, W.C., M.C. Hsu and Y.Y. Yang, 2010. Local contrast enhancement and adaptive feature extraction for illumination-invariant face recognition. *Pattern Recogn.*, 43: 1736-1747.
- Moerland, T., 2003. Steganography and Steganalysis. Universiteit Leiden, Rhone-Alpes, France.
- Mohamad, F.S., A.A. Manaf and S. Chuprat, 2010. Histogram matching for color detection: A preliminary study. *Proceedings of the International Symposium in Information Technology*, Volume 3, June 15-17, 2010, Kuala Lumpur, pp: 1679-1684.
- Pennebaker, W.B. and J.L. Mitchell, 1993. JPEG: Still Image Data Compression Standard. Springer Science & Business Media, USA., ISBN: 9780442012724, pp: 29-32.
- Pratt, W.K., 2001. Digital Image Processing: PIKS Inside. 3rd Edn., John Wiley and Sons, Inc., USA., ISBN-13: 978-0-471-37407-7, Pages: 656.
- Prince, S.J.D., 2012. Computer Vision: Models, Learning and Inference. Cambridge University Press, Cambridge, UK., ISBN-13: 9781107011793, pp: 345-346.
- Rani, N. and S. Bishnoi, 2014. Comparative analysis of image compression using DCT and DWT transforms. *Int. J. Comput. Sci. Mobile Comput.*, 3: 990-996.
- Wang, Z., 1984. Fast algorithms for the discrete wavelet transform and for the discrete fourier transform. *IEEE Trans. Acoust., Speech Signal Proces.*, 32: 803-816.
- Zahraddeen, S., S.M. Fatma and A.Y. Abdulganiyu, 2014a. An efficient discrete cosine transform and gabor filter based feature extraction for face recognition. *Proceedings of the 6th International Conference on Postgraduate Education*, December 17-18, 2014, Melaka.
- Zahraddeen, S., S.M. Fatma, A.Y. Abdulganiyu and A.S. Ben-Musa, 2014b. A new Hough transform on face detection using histograms. *Image Vision Comput. J.* (In Press).