FACE RECOGNITION BY EMBEDDING OF DT-CWT COEFFICIENT USING SOM AND ENSEMBLE BASED CLASSIFIER

Gauri Agrawal[1] and Sanjay Kumar Maurya[2]

[1]*Department of Electronics and Communication Engineering, B.S.A. College of Engineering and Technology, India*
E-mail: gauri.ag90@gmail.com
[2]*Department of Electrical Engineering, GLA University, India*
E-mail: skmaurya@ieee.org

Abstract

In real world, applications designing of a robust face recognition system have always been a big challenge. This paper presents an approach to face recognition using embedding of dual tree complex wavelet transform, Self-organizing map and ensemble of weak classifier. The DT-CWT is applied on the images to obtain dynamic and multi scale informational characterization of the face images. Thus, a multidimensional feature vector is formed by combining DT-CWT coefficients. Self Organizing Map (SOM) is further used in embedding the feature vector which also results in reduction of feature vector. Finally, ensembles of k-Nearest Neighbor (k-NN) weak classifiers are used for classification of recognition system. The proposed approach is tested on image of ORL database. The experiment shows an impressive recognition result that culminated during the testing.

Keywords:

Face Recognition, Wavelet Transform, Self-Organizing Map (SOM)

1. INTRODUCTION

Recently, face recognition has gained considerable attention and it is significantly inspired by the needs of military, commercial and public security systems. There are various challenges in the field of face recognition due to shift, pose and illumination. Moreover, occlusion due to various reasons makes it more challenging. So, a good face recognition system must account these restrictions effectively before classification [1]. Further, after the extraction of the robust and invariant features, researcher aimed to the performance of the classifiers. The basic model of the classification techniques involves training with multiple faces where several face images of a single person is used for training the classifier. Therefore, the increase in the number of class size adds to the burden on classifier, that, in turn, increases the size of processor.

Further, occlusion has become the challenge for face recognition model. Addressing partial occlusion in face recognition system is done by effective modeling of occluded face features. There are two popular approaches those are evident in delivering the occlusion; first one is statistics based and second one is structural based method. In statistics based method, face is defined as a whole and in structure based method recognition is based on the relationship among facial features e.g. eyes, nose, mouth and face boundary. Various literatures have been devoted to deal with partial occlusion in face recognition methods.

Different holistic methods such as local descriptor analysis (LDA), principal component analysis (PCA), fisher discernment analysis (FDA), locality preserving projection (LPP) [2] , non-negative graph embedding [3] etc. have been studied widely. In these conventional methods, global features are used to recognize the face images but all these approaches involve high computational cost for training data. Generally, these features are affected by occlusions or noise which reduces the robustness of these methods [4].

To overcome the disadvantages of above approaches, structure based method is used to deal with challenges such as occlusion, illumination and pose variations. Pentland et al. [5] proposed Eigen feature method in which information is extracted from local regions and features are received by performing PCA to local face regions independently. Kernels of local spatial support are used to extract information about local facial component in local feature analysis. Wiskott et al. in 1997 [6] proposed an approach in which elastic bunch graph matching is used with neural network to recognize the face images. In this approach, faces are stacked as flexible graphs and at every node of the graph Gabor features are connected. Gabor features are associated with wavelet transform. Gabor features are insensitive to noise, illumination etc. but the size of Gabor kernels is very large. Another method to use structural based approach is to divide images into sub-blocks and model these blocks separately. [7] To solve this problem, weighted local probabilistic subspace approach is used. Singh et al proposed 2D log Gabor phase features [8] and used one single image for per person to train the classifier. In this, Gabor wavelet transform is used to extract features and phase information is used to provide texture information to a face. In [9] local, Gabor binary pattern is used to recognize face images and in this, Gabor wavelet is used to extract the features and local binary pattern is used to give texture information. Since, the size of feature vectors formed by using Gabor wavelet is very large, hence, many researchers used local complex binary pattern on the coefficients of it while some uses another approach in face feature formation such as GLCM [10].

N.G. Kingsbury and Ivan Seleinc [11][12] develop DT-CWT din image processing application. It maintains the attraction for feature formation such as small feature vector in comparison with Gabor wavelet, illumination invariant feature [13] etc. Further, many researchers have been continuing their work in reducing the feature vector up to optimum level for reducing the burden on machine. In [14] used extra dyadic down sampling to reduce the feature vector significantly.

Many researchers focused to improve the performance of classifier. In the early days, researcher classifier was trained with multiple images of one class. As the number of the class increased, the burden on classifier added to more, which finally resulted in sluggish performance. Therefore, to reduce the burden on classifier, Kohonen et al. in 1991 [15] proposed an unsupervised algorithm which could extract all significant information of local facial features using Self Organizing Map (SOM). SOM helped much in reducing the burden on classification because it mapped high dimension data in low dimensional output space. Tan et al. in 2006 [16] proposed an algorithm in which one sample per person was used for training the classifier using self organizing map.

The structural features are affected by the partial occlusion but recognition methods can be made robust if these structural features are merged with statistical features intelligently. Kumar et al. in 2005 [17] proposed a model in which integration of two techniques SOM and PCA were used for dimensionality. The results were in appreciation in spite of low dimensionality. Some of the researchers use Gabor wavelet transform and self organizing map in face recognition model. It becomes challenging in creating the recognition model because firstly, it is computationally very complex and second, it requires of a very high memory to store Gabor features.

In this paper, dual tree complex wavelet transform is used for extracting the face information and SOM is used to further dimensionality reduction. Further, soft k-Nearest Neighbor (KNN) ensemble classifier is used for classification.

The remaining parts of the paper are organized as follows: Section 2 and 3 give the basic information of DT-CWT for feature vector extraction and embedding it using SOM. Section 4 describes proposed model of face recognition, section 5 describes results of the analysis and the section 6 concludes the paper.

2. BACKGROUND OF DT-CWT IN FEATURE EXTRACTION

Complex valued extension of the discrete wavelet transform is represented as complex wavelet transform. Kingsbury's [11] dual tree complex wavelet transform with important additional properties is an improvement to the discrete wavelet transform. DT-CWT is composed of two parallel real DWTs which produce the imaginary and real parts of the complex wavelet and satisfies perfect reconstruction condition. It is represented by Eq.(1),

$$\Psi(t) = \Psi_l(t) + j\Psi_h(t). \tag{1}$$

The wavelet $\Psi_l(t)$ and $\Psi_h(t)$ and corresponding scaling function $v(t)$ with filter bank pair can be defined as,

$$\Psi_l(t) = 2\sum_n l_0 v(2t-n);\ v_l(t) = 2\sum_n l_1(n)v(2t-n)$$

$$\Psi_h(t) = 2\sum_n h_0(n)v(2t-n);\ v_h(t) = 2\sum_n h_1(n)v(2t-n).$$

Here l_0 and l_1 are the low-pass and high pass filter pair respectively for the real part while h_0 and h_1 are the low-pass and high pass filter pair respectively for the imaginary part. The 1-D complex functions $\Psi(t)$ and $v(t)$ produces six wavelet for two dimensional DT-CWT, Therefore, when an image is subjected to DT-CWT; it is characterized by the translations and dilations of a complex scaling function and six complex wavelet functions. Consequently, the orientations of 2D dual-tree complex wavelet transform are in six different directions. The six sub bands are oriented in ±15°, ±45°, ±75° and are shown in Fig.1.

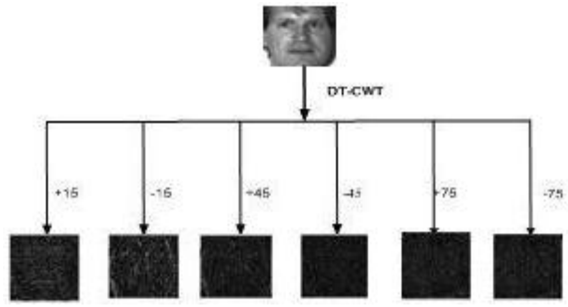

Fig.1. Six Sub bands of DT-CWT decomposed image

The properties of the DT-CWT like good directional selectivity, shift invariance bounded redundancy and efficient computation make it suitable for face feature extraction.

2.1 INTRODUCTION TO SELF ORGANIZING MAP

For converting the high dimensional data into lower dimensional data, SOM is an effective tool. It helps in visualization of the large size data. In SOM, an organized mapping of a high-dimensional distribution is done onto a regular low-dimensional grid. Thus, it is able to convert complex, nonlinear statistical relationships between high-dimensional data items into simple geometric relationships on a low-dimensional display by preserving the topology structure of the data [18]. SOM cuts down dimensions by developing a map of usually 1 or 2 dimensions which plots the resemblances of the data by grouping like data items together. Thus, SOMs accomplish two things: they reduce dimensions and display similarities [15].

The SOM is trained iteratively. In the beginning, weights are randomly assigned. When the n-dimensional input vector x is sent through the network, the distance between the w weight neurons of SOM and the input vector is calculated. It is preferable to calculate this distance by using Euclidian distance given by Eq.(2),

$$\|x - w\| = \sqrt{\sum_{i=1}^{n}(x-w)^2} \tag{2}$$

Best matching unit (BMU) measure the closest match to the present input for given weight. In the iterations t, the BMU and its neighboring neurons are allowed to learn by changing the weights which consequently reduces the distance between input vector and weights.

$$w(t+1) = w(t) + \alpha(t)h_{lm}(x - w(t)) \tag{3}$$

where, α varies from 0 to 1.

The neighborhood function which is function of best matching neuron is given in terms of position of winning neuron and its neighboring output nodes. The most commonly used neighborhood function Gaussian function given by Eq.(4).

$$h_{lm} = \exp\left(\frac{-\|l-m\|^2}{2\sigma(t)^2}\right) \tag{4}$$

where, h_{lm} is neighborhood function of the best matching neuron at iteration t and l and m are position of winning neuron and its neighborhood output node. The SOM network is constructed after the training is complete.

3. PROPOSED MODEL OF FACE RECOGNITION

The overall proposed approach is shown in Fig.2. The local face image similarity is achieved by dividing the DT-CWT filtered images of all sub-bands into smaller sub blocks. As the size of sub block increases the local information of face features becomes weak and, hence, the recognition rate decreases. These sub blocks are mapped into SOM topological space for reduced feature representation. The weak classifier, *k*-nearest neighbor is used for classification of the each sub block. For final recognition result, the ensemble of weak classifier is used.

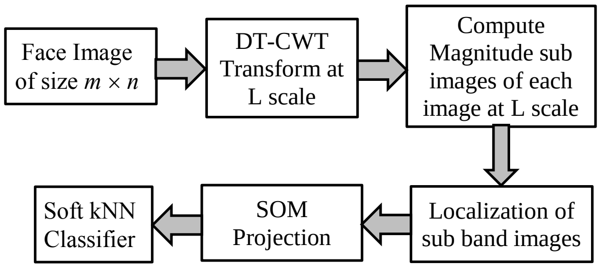

Fig.2. Block diagram of the proposed approach to recognize face images

The details of proposed model are explained in the subsections.

3.1 FEATURE VECTOR USING DT-CWT

Face image of size $n \times m$ is taken and DT-CWT is applied up to L levels and six directions to generate a series of sub-bands. In our experiment we have taken a face image of size 128 × 128 pixels, and DT-CWT has been applied up to three scales and six directions to generate a series of eighteen sub-bands. There are six-sub bands obtained in each of the scales.

For each scale, sub-band computes the magnitude of its complex coefficient. Spectral energy can be accurately measured by it. It is very less sensitive to small shift in face image. The magnitude of complex coefficient is used as a feature.

Only third scale sub-bands have been taken because it gives better performance in our experiment. The sub-band whose scale is smaller than three is considered as noise because of the variation in the environment. One more advantage of this operation is that the dimensions of the face image will be reduced.

3.2 LOCALIZATION OF SUB BAND IMAGES

Six magnitude sub-images that were obtained in the first step are divided into N sub- blocks. Local features of magnitude sub-images are extracted from the N sub-blocks of the face images. N = L/D, where L = Dimensionalities of whole magnitude sub-image and D = Dimensionalities of each sub-block.

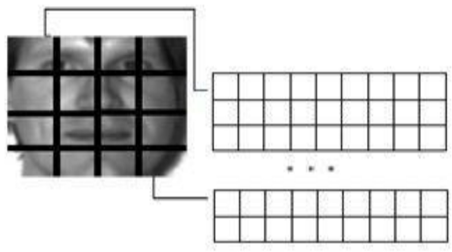

Fig.3. Representation of face images using set of sub-block vectors

So, one full dimensional feature vector can be converted into several low dimensional local feature vectors as shown in Fig.3. After localization of the face images large number of sub-blocks will be generated. Self organizing map is used to find the pattern of sub-images [16].

3.3 DT-CWT COEFFICIENT EMBEDDING USING SOM

There are two types of SOM face approaches for the SOM projection: (i) Single SOM face approach and (ii) Multiple SOM face approach.

In single SOM face approach, for the projection of SOM average of every sub-block vectors is taken, after localization of the face images. And then by using nearest neighbor classifier, average of sub-blocks of each training face is mapped to best matching unit. The disadvantage of this method is that if new face images are taken then weight vector in SOM will be recomputed and it is a very time consuming process.

In multiple SOM face approach, for the projection of SOM average of sub-blocks is not used. Instead, after localization of face images, local feature vectors are obtained to classify these local feature vectors, soft KNN classifiers are used and then outputs of these classifiers are ensemble to get the final output as shown in Fig.4.

3.4 SOFT KNN ENSEMBLE CLASSIFIER

The size of the database becomes very large if we use multiple images per person for training and the burden on the classifier also increases. So, to reduce the burden on the classifier, single image per person is used for training. K-Nearest Neighbor (KNN) is an efficient classification algorithm. It classifies testing feature vector based on the similarities in the feature vector in the training data. K is the number of the nearest neighbor used for classification.

K = 1 means single nearest neighbor used for classification. The best selection of K depends on the data by using heuristic techniques that has been already used for this. K nearest neighbor algorithm can work efficiently against noisy data. The effect of noise can be reduced if larger value of K is chosen. Classification does not involve fitting problem and curse of dimensionality. In KNN algorithm generally one out of Euclidean, Manhattan and Cosine is used as a distance measure. Recognition rate depends on the correctly classified classes.

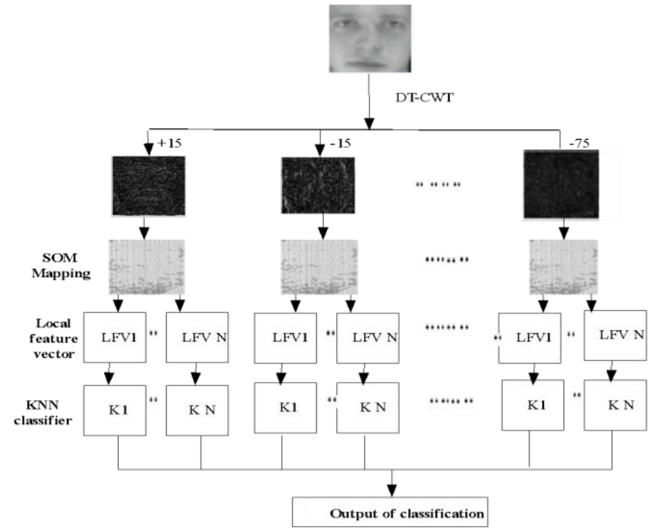

Fig.4. Flow chart of the proposed approach

To improve the accuracy rates of the classification mechanism ensemble classifier have been used. In ensemble classifier outputs of multiple classifiers (KNN) are combined to recognize the face images.

4. RESULT AND DISCUSSION

In order to evaluate the proposed face recognition system, our experiments are performed on the benchmark face database:

4.1 ORL DATABASE

The ORL database contains 400 face images of 40 individuals. By varying the facial details like glasses or without glasses lighting, expressions of faces (like closed and open eyes, smile and sad etc.) many images are taken at different times for some subjects. For the experiment purpose 400 images of 40 individuals were taken and there were 10 images per subject.

Fig.5. Sample images of ORL database of different angles

In Fig.5(A)-Fig.5(E) represent the face images of ORL database at different angles. But the size of database becomes very large if we use multiple images per person for training, and, consequently the burden on the classifier also increased. So to reduce the burden on the classifier single images per person is used for training. In this section, we present the performance of SOM based approach under different variations, especially shift, illuminations and angle occlusion and compare with the other popular approaches.

Experiments were conducted on MATLAB R2013a to compare the proposed DT-CWT+SOM+KNN method with SOM+KNN method. Rate of recognition means the rate of accuracy to correct the top response of the methods.

Results of two experiments on images B and C is shown in Fig.6. Image B is at 45^{o} right of the reference image and image C is at 45^{o} to the left of the reference image. Here, horizontal axis is the rank and vertical axis is the recognition rate and rank one means top match score. For training, frontal images and for testing different angle occluded images have been taken. In this paper, only one reference image per individual has been used for training so the size of training data set has reduced.

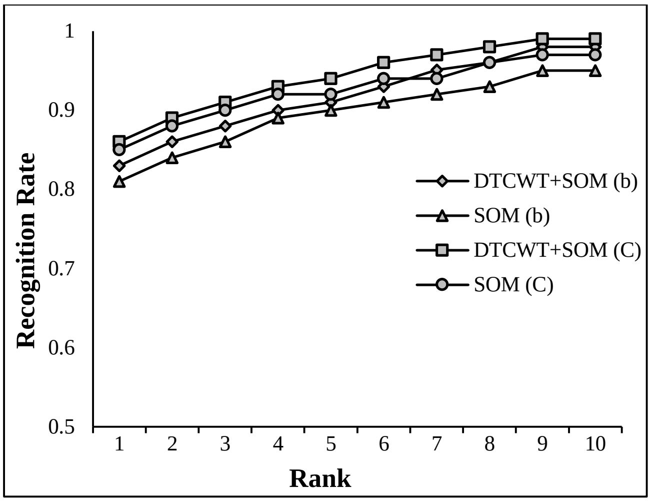

Fig.6. Comparative performances of DTCWT+SOM+KNN and SOM+KNN

The Fig.6 shows better recognition result using DTCWT+SOM+KNN in comparison with SOM+KNN method.

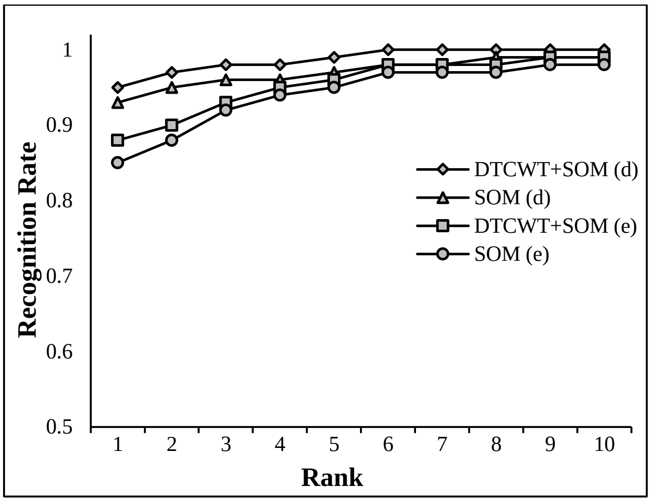

Fig.7. Comparative performance of DTCWT+SOM+KNN and SOM+KNN

Result of two experiment on images D and E is shown in Fig.7. And we can see that better recognition result is obtained by using DTCWT+SOM+KNN in comparison to SOM+KNN method. DTCWT makes image insensitive to shift and illumination.

For the soft KNN classifier to choose an appropriate value of K is a challenging issue. So to overcome this problem sub block size of 4×4 has been chosen and SOM is trained by using 40 images and the recognition rate of these images have been measured by varying the value of K. From Fig.8, we can see that larger value of K does not necessarily give better result. So we can choose the small value of K instead.

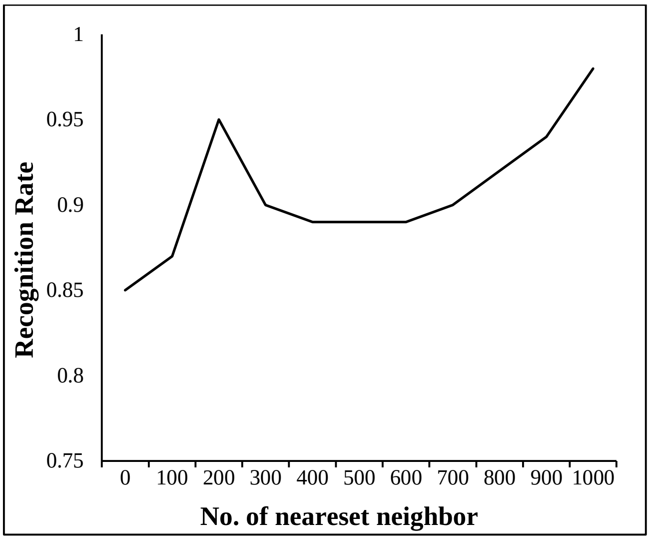

Fig.8. Recognition rate of the proposed approach with varying the number of nearest neighbors

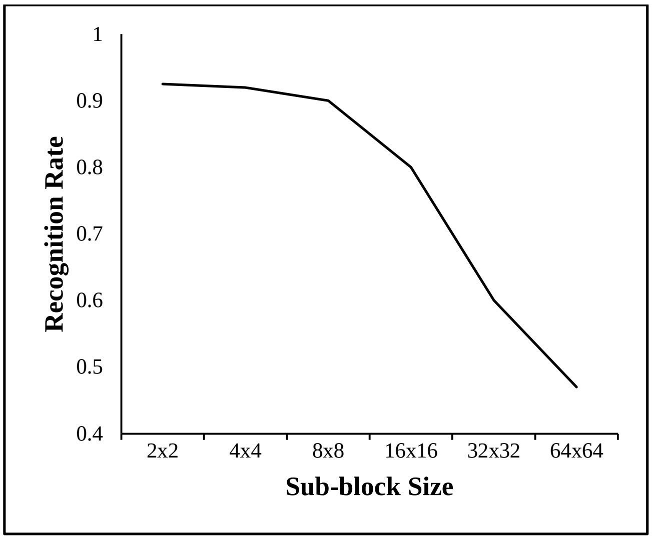

Fig.9. Recognition rate of the proposed approach with varying sub block size

The performance of the proposed approach is varied by using different sub-blocks size. This experiment was repeated for different sub blocks with fixed K value. From Fig.9, we can see that as the size of the sub blocks increase, the recognition rate decrease. So size of sub block should be small for robust recognition.

5. CONCLUSION

In this paper a face recognition model which meets the challenges of the recognition such as luminance, handling large feature vector, burden on the classifier and occlusion is presented. Further, the recognition model uses single face for training the classifier. The classification process uses weak classifier ensembles for recognition. The model is tested on ORL database. The result shows the impressive face recognition results under angular occluded image.

REFERENCES

[1] C. Devi Priyai and A. Srinivasan, "A Novel Design for Face Recognition System using Dual Tree Complex Wavelet Transform", *Proceedings of 3rd International Conference on Electronics Computer Technology*, pp. 153-156, 2011.

[2] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, pp. 103-108, 1990.

[3] Du Yijun, Xiaobo Lu, Wujun Chen and Qianzhou Xu, "Gender Recognition using Fisherfaces and a Fuzzy Iterative Self-Organizing Technique", *Proceedings of 10th International Conference on Fuzzy Systems and Knowledge Discovery*, pp. 196-200, 2013.

[4] Hanwang Zhang, Zheng-Jun Zha, Shuicheng Yan, Meng Wang and Tat-Seng Chua, "Robust Non-Negative Graph Embedding: towards Noisy Data, Unreliable Graphs, and Noisy Labels", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2464-2471, 2012.

[5] Matthew Turk and Alex Pentland, "Eigenfaces for Recognition", *Journal Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991.

[6] L. Wiskott, J M. Fellous, N. Kuiger and C. Von Der Malsburg, "Face Recognition by Elastic Bunch Graph Matching", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 775-779, 1997.

[7] A.M. Martinez, "Recognition of Partially Occluded and/or Imprecisely Localized Faces using a Probabilistic Approach", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 712-717, 2000.

[8] Xuegang Hu and Wenyu Li, "Multi-Scale Optical Flow Estimation of the Video Based on Gradient Optimization", *Proceedings of 3rd International Congress on Image and Signal Processing*, pp. 335-339, 2010.

[9] Reza Azmi and Samira Yegane, "Facial Expression Recognition in the Presence of Occlusion using Local Gabor Binary Patterns", *Proceedings of 20th Iranian Conference on Electrical Engineering*, pp. 742-747, 2012.

[10] Alaa Eleyan and Hasan Demiril, "Co-occurrence Matrix and its Statistical Features as a New Approach for Face Recognition", *Turkey Journal of Electrical Engineering and Computer Science*, Vol. 19, No. 1, pp. 97-107, 2011.

[11] Nick Kingsbury, "Image Processing with Complex Wavelets", *Philosophical Transactions of the Royal Society A Mathematical, Physical and Engineering Science*, Vol. 357, No. 1760, pp. 2543-2560, 1999.

[12] Ivan W. Selesnick, Richard G. Baraniuk and Nick G. Kingsbury, "The Dual-Tree Complex Wavelet Transform", *IEEE Signal Processing Magazine*, pp. 123-151, 2005.

[13] Haifeng Hu, "Multiscale Illumination Normalization for Face Recognition using Dual-Tree Complex Wavelet Transform in Logarithm Domain", *Computer Vision and Image Understanding*, Vol. 115, No. 10, pp. 1384-1394, 2011.

[14] Gauri Agrawal and Sanjay Kumar Maurya, "Reduced Robust Facial Feature Descriptor using DTCWT and PCA", *Recent Advances and Innovations in Engineering*, pp. 1-5, 2014.

[15] T. Kohonen, "Self-Organizing Maps: Optimization Approaches", *Artificial Neural Networks*, Vol. 2, pp. 981-990, 1991

[16] Xiaoyang Tan, Songcan Chen, Zhi Hua Zhou and Fuyan Zhang, "Face Recognition from a Single Image per Person: A Survey", *Pattern Recognition*, Vol. 39, No. 9, pp. 1725-1745, 2006.

[17] D. Kumar, C.S. Rai and S. Kumar, "Face Recognition using Self-Organizing Map and Principal Component Analysis", *Proceedings of International Conference on Neural Networks and Brain*, pp. 1469-1473, 2005.

[18] Samuel Kaski, "Data Exploration using Self Organizing Maps", Ph.D. diss., Department of computer science and technology, Acta Polytechnica Scandinavica, 1997.