International Journal of Multidisciplinary and Current Research ISSN: 2321-3124
Research Article Available at: http://ijmcr.com

Ethnicity Identification based on Fusion Strategy of Local and Global Features Extraction

Shakhawan Hares Wady[1] and Hawkar Omar Ahmed[2]

[1]Mathematics Department, School of Science Education, University of Sulaimani, Sulaimani, Iraq
[2]Statistics and Computer Department, College of Commerce, University of Sulaimani, Sulaimani, Iraq

Accepted 20 Feb 2016, Available online 03 March 2016, **Vol.4 (March/April 2016 issue)**

Abstract

Human facial identification become an active and challenging subject in the area of computer vision, that provides the demographic information such as ethnicity, age, gender, etc. However, even with all the research that has been conducted in ethnic identification, it is still seems to be a difficult and largely problem. In this paper, we propose an approach for robust ethnic identification by fusing two complementary local and global descriptors. In this approach, uniform Local Binary Pattern (ULBP) has been used to extract the local features while Discrete Cosine Transform (DCT) has been performed to extract global features from facial images. The classification of the selected feature vector is computed using the k-Nearest Neighbor (KNN) classifier with the city block distance. We evaluate the performance of the proposed fusion approach, through the extensive experiments conducted on our collected dataset. Our experimental results indicate clearly that the proposed fusion strategy gives a high level of 98.03% ethnicity identification accuracy rate.

Keywords: *Fusion Strategy, Feature Extraction, Ethnicity Identification, Uniform Local Binary Pattern (ULBP), Discrete Cosine Transform (DCT), k-Nearest Neighbor (KNN)*

1. Introduction

Ethnicity identification from face images refers to the process of recognizing the ethnic group that can be used in various areas of application. The basic identification process will consist of selecting, extracting a propagate facial features, and searching for similar facial images from a given dataset. Over the past few dedicates, several approaches have been proposed for ethnicity identification and much progress has been made However, it is hard to recognize faces belonging to a different race swiftly and accurately under the uncontrolled condition. Despite the fact that humans are able to identify and recognize faces in a scene with little or no effort, getting a computer system to accomplish such objectives is very challenging. The challenges generally come from the large variations in visual stimulus due to facial expressions, viewing directions, poses occlusions (eye glasses) and illumination conditions.

2. Related work

In literature, the algorithms used to represent the face images for ethnicity identification are categorized mostly into two groups, namely, global feature extraction and local feature extraction algorithms [1]. Recently, a lot of extraction approaches has been done on ethnicity identification. Faraidoon and Aree [2], presents an approach through hybrid wavelet and discrete cosine transform method for ethnicity identification with a good accuracy. Venkata *et al*. [3] proposed a multi-feature fusion based facial expression classification using DLBP and DCT also with good levels of accuracy. Salah *et al*. [4] describe a fusion system that uses block-based uniform LBP and Haar wavelet transform to concatenate local and global features, the experiment results show good level of accuracy for ethnicity identification. Manesh *et al*. [5] presented a two class ethnicity classification problem Asian and non-Asian using an appearance-based method to determine the confidence of different facial regions using SVM on facial features such as eyes, nose and mouth also with good accuracy. A linear discriminant analysis (LDA) approaches for two class (Asian and non-Asian) ethnicity classification has been developed by Xiaoguang and Anil [6], it is unclear about the efficiency of their approaches in more accurate conditions of multi-ethnic groups. Ghulam *et al*. [7] performed LBP and WLD for three ethnic class (Tibetan, Uygur and Zhuang) using local descriptors. The paper proposes the race recognition system from face images based on Weber local descriptors (WLD) with City block, Euclidean, and chi-square minimum distance classifiers. Hlaing and Myint [8] presented the race identification issue based on facial images. The PCA based scheme has been used for the two-class (Myanmar vs. Non-Myanmar) race classification task. The idea behind this paper is to combine two discriminative and complementary vector facial features. In this approach, uniform Local Binary Patterns (ULBP) is performed to extract local features, while two-dimensional DCT is performed to extract global features of the image. The new feature vector is then applied to similarity measure classifier using the k-Nearest Neighbor (KNN). Experiments with different scenarios are implemented on our collected dataset.

This paper is organized into six main sections. Section 3 briefly describes the algorithms of uniform LBP and DCT feature extraction methods. Section 4 presents the block diagram of our proposed system that have been used in this paper. In Section 5, evaluation method and experimental results are given and Section 6 concludes the paper.

3. Feature extraction methods

Feature extraction is the key process in any ethnicity identification that involves reducing the amount of resources required to define a large dataset accurately [9]. Over the past few decades many methods have been proposed in the literature for facial feature extraction. In this paper We seek to integrate uniform local binary pattern (ULBP) as local feature extractor and discrete cosine transformation (DCT) as global feature extractor for ethnicity identification.

3.1 LBP operator description

The basic local binary pattern operator, introduced by Ojala *et al*. [10], is a powerful feature extraction technique which consider both texture and shape information to represent facial. The LBP operator assigns a label to every pixel of an image by thresholding the intensity values in the 3×3-neighborhood of each pixel with the center pixel intensity value and converts the result into a binary number by using equation (1).

$$\boldsymbol{LBP}_{P,R}\,(x,y) = \sum_{n=0}^{n-1} 2^n\, S\,(i_n - i_c) \qquad (1)$$

$$\boldsymbol{S}\,(\boldsymbol{k}) = \begin{cases} 1 & k \geq 0 \\ 0 & k < 0 \end{cases}$$

where i_c denote the gray level of the central pixel (x_c, y_c), i_n denote the gray level of the 8 surrounding pixels, and s(k) is defined as thresholding operation function. The binary result will be obtained by reading the values clockwise, starting from the top left neighbor. the basic LBP operator is illustrated in Figure 1.

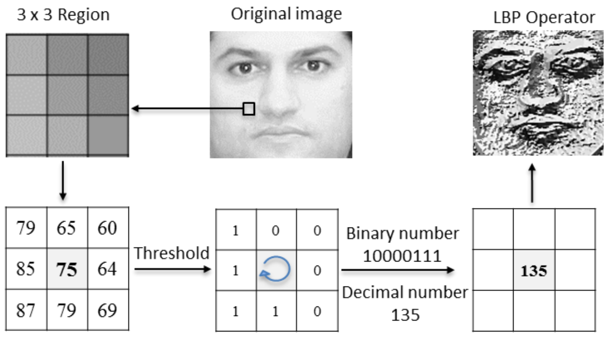

Figure 1: The Basic Local Binary Pattern operator

Another extension to the original LBP operator which uses the property called uniform patterns [11,12] The LBP is identified uniform pattern if it contains at most one (0 – 1) transition and one (1- 0) transition. For instance, the patterns 00000000 (0 transitions), 001111100 (2 transitions) and 11110001 (2 transitions) are uniform patterns. The uniform LBP feature will significantly reduce the dimension of the LBP feature vector which is the advantage for ethnicity identification from face images. The equation of uniform LBP form is shown in equation (2).

$$\boldsymbol{ULBP}_{P,R}(x,y) = |S(i_{n-1} - i_c) - S(i_0 - i_c)| + \sum_{n=0}^{n-1} 2^n\, |S(i_n - i_c) - S(i_{n-1} - i_c) \quad (2)$$

In our experiments, we use uniform local binary pattern operator with 8 sampling points in a local neighborhood region of radius 1 and labelling all remaining uniform patterns with a single label. The reasons for using uniform LBP are twofold. firstly, it is widely accepted that uniform LBP are extremely applicable and most of the image local structures are represented by uniform patterns. It was observed experimentally in [10] that uniform patterns account for a bit less than 90% of all patterns when using the (8,1) neighborhood. Secondly, the reason for using uniform LBPs is the statistical robustness. Using uniform LBP in place of all the possible patterns has produced better identification outcomes in numerous applications [13, 14], also uniform LBP themselves are more stable than the original LBP [15]. After labeling images with uniform LBP operator, a histogram of the labelled images can be defined using equation 3.

$$H(k) = \sum_{x,y} I(f((x,y) = i) \qquad ,i = 1, \ldots .N \qquad (3)$$

Where *I* is the indicator function, $f(x,y)$ is a LBP label at pixel (x, y) and N is the number of bins (number of different labels produced by LBP operator).

3.2 Discrete cosine transform (DCT)

Discrete-Cosine-Transform or DCT is a popular transform that convert an image from spatial domain to frequency domain that most widely used for feature extraction in image processing applications. DCT reduce the amount of bits needed to represent the information in an image, by eliminating the redundancy between neighboring pixel values [16]. The DCT of an image basically involves of different frequency bands low frequency, middle frequency and high frequency sub bands. Each sub band containing some detail and information in an image. The low frequency sub band generally contains the average intensity of an image which is the most intended in feature extraction [17, 18]. Mathematically, the general equation for a 2D-DCT (N x M) of an image is defined by the following equation 4.

$$F(u,v) = \frac{2}{\sqrt{NM}} \sum_{i=0}^{N-1} A(i) \cos\left[\frac{\pi u}{2N}(2i+1)\right] \sum_{j=0}^{M-1} A(j) \cos\left[\frac{\pi v}{2M}(2j+1)\right] \; f(i,j) \tag{4}$$

$$A(i), A(j) = \begin{cases} 1/\sqrt{2} & for\ u,v = 0 \\ 1 & otherwise \end{cases}$$

Where,

- The size of the image is $(N \times M)$.
- $f(i,j)$ is the intensity of the pixel at coordinates (i,j)
- $F(u,v)$ is the DCT coefficient at coordinate (u,v)

The resulting vector has dimensionality MN for image resolution M x N.

In our work, the original image is divided into small blocks of pixels (8×8 pixels), then discrete cosine transform is performed over each block independently and result can be 64 transform coefficients. The image features associated with the DCT is shown in Figure 2. The next step the quantization process on the 8 x8 DCT coefficient then, order the quantized coefficient by applying the zigzag scanning function.

The outcomes of a 64-element DCT transform are 1 DC coefficient and 63 AC coefficients. The low frequency DC coefficient that are located in the upper left corner represent the average energy of cells that contain the most of information and is the most affected by variation in illumination. Therefore, the low frequency DCT coefficients is utilized for selecting more effective subset of features for fusion process with uniform LBP.

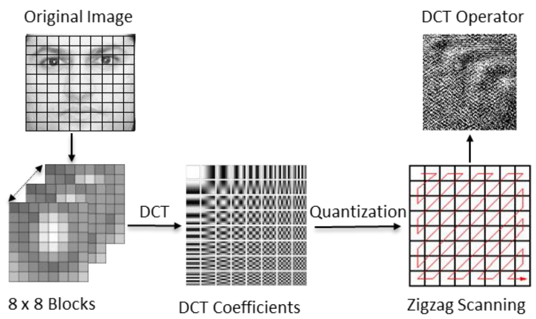

Figure 2: The basic Discrete Cosine Transform (DCT)

4. Methodology

This paper proposes an efficient algorithm for ethnicity classification from face images by combining both local and global feature extraction based on uniform LBP and DCT. The block diagram of our proposed algorithm has been illustrated in Figure 3. the classification system typically contains several steps such as preprocessing, feature extraction, feature concatenation and classification.

The steps in this classification algorithms are explained below.

1) Preprocessing: in this step, the original facial color images are cropped based on the two eyes location, then resize facial images to the size of 128x128 pixels, and convert color images to grayscale image.
2) Local feature extraction: in the second step uniform LBP has been performed for local feature extraction process. In our framework, instead of conventional LBP, uniform LBP is used for local feature extraction. After computing uniform LBP for each pixel the histogram of uniform LBP over local face image region has been computed.
3) Global Feature extraction: this step begins by dividing the face images in to blocks of 8 x 8 pixel then two-dimensional DCT algorithm has been performed to extract the global features vectors from face images.
4) Feature Fusion: after computing uniform LBP and DCT on the face images, the extracted local feature and global features are fused to form a single feature vector for the whole image using concatenation process.
5) Classification: the process of classification features is performed using the K-Nearest Neighbor (KNN) Classifier.

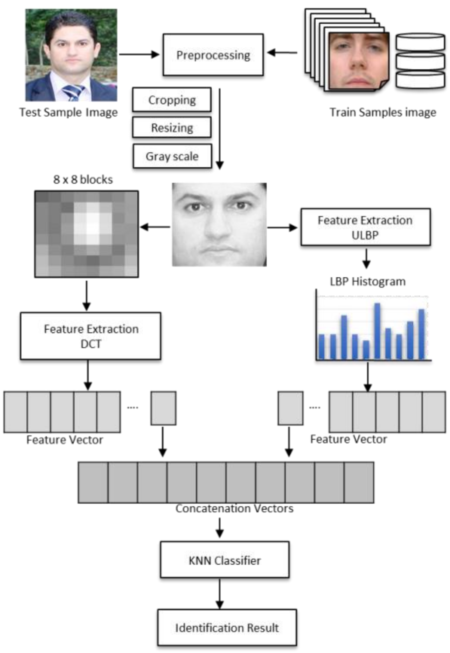

Figure 3: Block diagram of proposed system

5. Experiments and results

5.1 Dataset

Currently, there are a large number of face images datasets available to researchers for face recognition, such as FERET dataset, the GUFD face dataset, the YALE face dataset, and etc. [19, 20]. Most of the face dataset contain a set of face images collected in different pose condition facial expressions, illumination conditions, etc. The dataset presented in Table 1 represents a set of our face images that was collected manually from different dataset source. The experimented dataset consists of 1000 color frontal face images from different subjects and different facial expressions, occlusions (eye glasses and scarf), and illumination conditions. As presented in Table 1 we focus on three major ethnic groups: European, Oriental and African groups with 400, 400 and 200 face images respectively for ethnicity identification system.

Table 1 Collected dataset images from different dataset source

Race Group	Database	No.Subjects	Total Images	Variation
European	FERET	120	400	lighting, expression
	GUFD	170		
	PUT	74		
	Ext. YALE	36		
Oriental	CAS-PEAL	64	400	eye glasses, lighting, expression
	CASIA	110		
	CUHK	100		
	FERET	105		
	JAFFE	11		
African	FERET	95	200	scale, lighting , eye glasses
	PDA	22		
	Other	83		
Total Images			1,000	

5.2 Classifier

Classification include a wide range of decision making approaches that are used in facial recognition or identification system. pixel-based image classification techniques analyze the numerical properties of selected image features vectors and organizes data into categories. Recently, many classification techniques have been developed such as support vector machine (SVM), Decision Tree (DT), and Linear Classifier (LC), etc. [21]. In this work the k-Nearest Neighbor (KNN) has been employed to classify face images using a leave-one-out validation strategy. The KNN classifier is one of the most popular learning and classification techniques that consists of two phases of processing: training and testing feature vectors [22, 23]. In the training phase the image features are isolated then in testing phase the features are used for classification of images. In k-NN classification process, an image is classified by a majority vote of its neighbors, with image being assigned to the category most common between its k nearest neighbors. In our experiments, the value of k is set to five to guarantee majority voting at least one of three ethnic groups. To study the effect on identification accuracy and computational efficiency the city block distance will be used for measuring the distance between image feature vectors.

5.3 Results and Discussion

Here, we present different experiments for ethnicity identification from the facial images. In our study we describe with details a collected dataset of 1000 RGB face images converted into grayscale with a fixed size of 128 x 128 pixels. The dataset images are collected from different dataset images as explained in Section 5.1. In these experiments we implemented local and global feature extraction algorithms for ethnicity identification separately then for our proposed algorithm we implemented feature fusion approach by concatenating the two feature vectors obtained from uniform LBP and DCT to form a new feature vector. Then the new feature vector computed to KNN classifier for classification process. In the feature extraction process, the local features are extracted by computing uniform LBP. Table 2 shows the ethnicity identification average accuracy obtained using different uniform LBP implementations with the dataset images. The first column identifies the uniform LBP implementation. The second column lists the ethnic groups' average accuracy for European, Oriental and African. The third column contains the total average identification accuracy using different number of ULBP feature vector component. The table shows the best average identification accuracy is achieved when using only two transition uniform LBP with 42 bins and the accuracy in this case reach 63.61%. Therefore, the best result of ULBP (42 bins) has been used as local feature extraction for fusion process with DCT.

Table 2 Ethnicity Identification accuracy using different ULBP implementation

Number of bins	Ethnic Groups			Average Accuracy
	European	Oriental	African	
14 (1 transition only)	67.25	65.25	30.00	54.17
42 (2 transition only)	77.25	72.25	41.33	**63.61**
56 (≤ 2 transitions)	74.50	72.00	39.33	61.94
58 (Uniform LBP)	75.00	70.75	34.67	60.14
59 (Uniform and Non-Uniform)	76.00	69.00	35.33	60.11

In global feature extraction, DCT descriptor has been used to extract the global grayscale features of the whole image. In our case we experimented different length of DCT feature vectors including (9, 16, 25, 64, 144 and 256) for block sizes 3 × 3, 4 × 4, 5 × 5, 8 × 8, 12×12, and 16 ×16, respectively. Table 3 illustrations the ethnicity identification average accuracies obtained for European, Oriental and African groups and total average accuracy using different DCT block feature vector size on the dataset images. The overall best performance was given by block size 5 × 5, yielding an average accuracy of 97.11%. A poorer performance was reached in case of the use block size 16 ×16, yielding an average accuracy of 87.42% for global features.

Table 3 Ethnicity Identification accuracy using different DCT block size

DCT block size	Ethnic Groups			Average Accuracy
	European	Oriental	African	
3 × 3	98.00	97.25	92.33	95.86
4 × 4	98.50	98.75	92.67	96.64
5 x 5	99.00	99.00	93.33	**97.11**
8 x 8	99.25	99.25	87.33	95.28
12 x 12	93.75	91.50	90.00	91.75
16 x 16	99.50	98.75	64.00	87.42

After extracting the local and global feature vectors (Uniform LBP and DCT) with the high yielding accuracies, both features are then fused by concatenating them into a single vector for classification task. The KNN classifier has been employed to identify the similarity according to city block distance between feature vectors. In Fusion process, to evaluate the performance of the proposed approach we experimented 42 ULBP features (only two transitions) with different length of DCT feature vectors including (9, 16, 20, 25 and 36) for block sizes 3 × 3, 4 × 4, 4 × 5, 5 ×5, and 6 x 6 respectively. The results for each of the average accuracy rates across the three ethnicities (European, Oriental and African) are given Table 4. With our experimentation we concluded that the overall best performance was given by block size 4 × 5, yielding an average accuracy of 98.03%.

Table 4 Ethnicity Identification accuracy of the proposed approach

No. of Features ULBP + DCT	DCT Features	European	Oriental	African	Average Accuracy
51	3 x 3	97.80	98.00	94.00	96.60
57	4 x 4	98.80	98.80	93.73	97.11
62	4 x 5	99.05	99.55	95.50	**98.03**
67	5 x 5	98.55	99.55	91.07	96.39
78	6 x 6	98.30	98.55	91.04	95.96

As this figure illustrates (Figure 4), the ethnicity identification accuracy rates for different block sizes. all the performed experiments, has been computed by concatenating different (ULBP + DCT) length of the feature vectors with (51, 57, 62, 67, and 78) features. 98.03 % was the best identification accuracy rate achieved with 62 features by concatenating 42 local features of uniform LBP and 20 global features of DCT.

Figure 5 demonstrates the average accuracy rate across the three different ethnicities group. The experiments performed for all ethnic groups show a clear improvement when local ULBP and global DCT are fused together and the average accuracy rates were 99.05 %, 99.55 %, and 95.50 % for the European, Oriental and African groups, respectively. Our results, investigate clearly the fusion of both local (ULBP) and global (DCT) features for ethnicity identification give high level of accuracy rate.

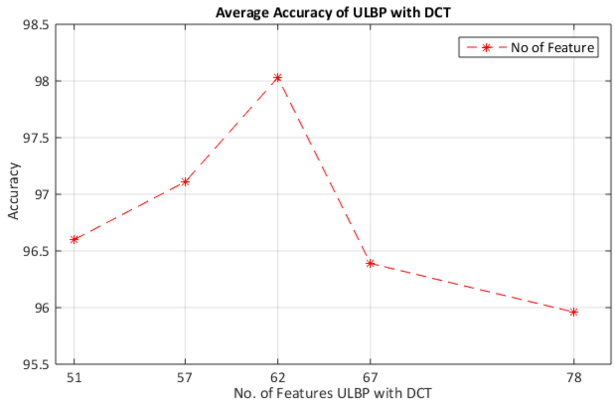

Figure 4: Effect of fusing of No. of features on identification accuracy

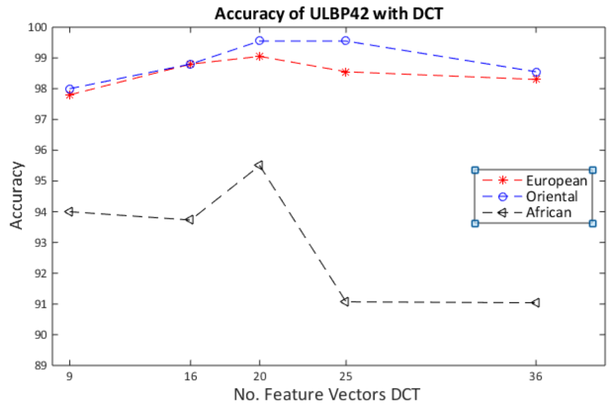

Figure 5: Effect of No. of features on ethnicity identification accuracy

Our results, investigate clearly the fusion of both local (ULBP) and global (DCT) features for ethnicity identification gives a high level of accuracy rate.

Conclusion

Reasonable work has been done on ethnicity identification, but search for improved ethnicity identification performance is still going on. In this paper, issue of ethnicity identification from face images has been investigated. It is being observed that the fusion of both local and global features from facial images improves the identification performance. In this method, uniform LBP is used to extract local descriptors while Discrete Cosine Transform (DCT) are used to extract global features from facial expression images. The proposed method is robust because high identification accuracies can be obtained with very few ULBP and DCT features. With our experimentation we concluded that the overall best performance was given by block size 4 × 5, yielding an average accuracy of 98.03 %. Our results, investigate clearly that the fusion strategy for ethnicity identification gives a high level of identification accuracy rate.

References

[1]. X. Han, H. Ugail and I. Palmar, (2009), Gender classification based on 3D face geometry features using SVM, *IEEE international conference on cyber world*, pp. 114-118.

[2]. Faraidoon H. & Aree A., (2015), Hybrid Wavelet and Discrete Cosine Transform Methods for Ethnicity Identification, *JZS - Volume 17 - No. 1*.

[3]. Ch. Venkata, K.V. Krishna, Debnath and Tai-hoon, (2014), Multi-Feature Fusion based Facial Expression Classification using DLBP and DCT, *International Journal of Software Engineering and Its Applications*, Vol.8, No.9 pp.55-68.

[4]. S. Hma, H. Du, and N. Al-Jawad, (2013), Fusing Local Binary Patterns with Wavelet Features for Ethnicity Identification, *International Journal of Computer, Electrical, Automation, Control and Information Engineering,* Vol:7, No:7.

[5]. Manesh, F.S., Ghahramani, M., and Tan, Y., (2010), Facial part displacement effect on template-based gender and ethnicity classification, Control Automation Robotics & Vision (ICARCV), *International Conference on Control, Automation, Robotics and Vision, Singapore*, pp.1644-1649.

[6]. Xiaoguang Lu and Anil K. Jain, (2004), Ethnicity identification from face images, Proceedings of SPIE *International Symposium on Defense and Security: Biometric Technology for Human Identification, Orlando, Florida*, pp.114-123.

[7]. muhammad h., fatmah A., george B., anwar M., and hatim A., (2015), Race classification from face images using local descriptors, *International Journal of Artificial Intelligence*, DOI: S0218213012500194.

[8]. Hlaing T. and Myint S., (2011), Race Identification for Face Images, *ACEEE Int. J. on Information Technology*, Vol. 01, No. 02.

[9]. P. Thangaraju, and N. Mala, (2015), Effectiveness of Searching Techniques in Feature Subset Selection: A Review IRACST, *International Journal of Computer Science and Information Technology & Security (IJCSITS)*, ISSN: 2249-9555 Vol. 5, No2.

[10]. Ojala, T., Pietikainen, M., and Maenpaa, T., (2002), Multiresolution gray scale and rotation invariant texture classification with local binary patterns*, IEEE Trans. Pattern Anal. Mach. Intell*., vol. 24, no. 7, pp. 971–987.

[11]. H. R. Eghtesad, and M. C. Amirani, (2013), Texture Classification with Local Binary Pattern Based on Continues Wavelet Transformation, *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, (An ISO 3297: 2007 Certified Organization*, Vol. 2, Issue 10.

[12]. Guo, Z., Zhang, D., Zhang, D., (2010), A completed modeling of local binary pattern operator for texture classification, *IEEE Transactions on Image Processing,* 19(6)1657–1663.

[13]. Tapia, J., Perez, C., (2013), Gender classification based on fusion of different spatial scale features selected by mutual information from histogram of LBP, Intensity, and Shape, *IEEE Transactions on Information Forensics and Security,* 8(3)488–499.

[14]. Zhou, H., Wang, R., Wang, C., (2008), A novel extended local-binary-pattern operator for texture analysis. *Information Sciences,* 178(22)4314 – 4325.

[15]. M.Navya, B.Munilakshmi, R.Kalyan, (2015), Fast Texture Data Multi Resolution Approach for Noise Reduction, *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 3, Issue 6.

[16]. Maneesha G., and Amit K., (2012), Analysis of Image Compression Algorithm Using DCT, *International Journal of Engineering Research and Applications (IJERA)*, ISSN: 2248-9622, Vol. 2, Issue 1, pp.515-521.

[17]. Nadia B., Omar A. and Rokia M., Texture feature extraction in the spatial-frequency domain for content-based image retrieval.

[18]. Aman R, Pallavi P., and M. Mani, (2011), Face Recognition Using Discrete Cosine Transform for Global and Local Features, *Proceedings of the 2011 International Conference on Recent Advancements in Electrical, Electronics and Control Engineering (IConRAEeCE), IEEE Explore*: CFP1153R-ART; ISBN: 978-1-4577-2149-6.

[19]. Burton, A., White, D., and McNeill, A., (2010), The Glasgow face matching test. Behavior research methods, 42(1), pp286-291.

[20]. Milborrow, S., Morkel, J. and Nicolls, F., (2010), The MUCT landmarked face database, Pattern Recognition Association of South Africa.

[21]. S. Gutta, J.R.J. Huang, P. Jonathon, H. Wechsler, (2000), Mixture of experts for classification of gender, ethnic origin, and pose of human faces, *IEEE Transactions on Neural Networks*, Vol.11, No.4, pp948-960.

[22]. Bremner D., Demaine E., Erickson J., Iacono J., Langerman S., Morin P., and Toussaint G., (2005), Output sensitive algorithms for computing nearest neighbor decision boundaries, Discrete and Computational Geometry, pp. 593–604.

[23]. Cover, T., and Hart, P., (1967), Nearest-neighbor pattern classification, *Information Theory, IEEE Transactions*, pp. 21-27.