Human Emotion Recognition Based On Wavelet Domain Feature Fusion

E. Pandian, *Research Scholar Manonmaniam Sundaranar University*
Dr. S. Santhosh Baboo, *Associate Professor, Department of Computer Science*
D G Vaishnav College, India.

Abstract:

The affective computing is becoming one of the challenging area in the Human-Computer Interaction. Capturing and evaluating the exact expression of human emotion that is the affective aspect is becoming daunting task for the computer professional's all over the world.In this paper, an efficient novel approach for human emotion recognition system based on the fusion of features extracted from the Discrete Wavelet Transform (DWT) and Undecimated Wavelet Transform (UWT) is presented. The main drawback of DWT is not translation invariant. Translations of an image lead to different wavelet coefficients. UWT is used to overcome this and more comprehensive feature of the decomposed image is obtained. The classification of human emotional state is achieved by extracting the energies from all sub-bands of DWT and UWT decomposed image. Then the features are fused together to classify the emotion of the human image. The robust K-Nearest Neighbor (K-NN) is constructed for classification. The evaluation of the system is carried on using JApanese Female Facial Expression (JAFFE) database. Experimental results show that the proposed energy fusion algorithm system produces more accurate recognition rate than DWT and UWT energies.

Keywords: Affective Computing,Facial expression, wavelet transform, undecimated wavelet transform, human emotion, feature fusion

1. Introduction

The motion or positions of the muscles in the skin of a human face convey the emotional state of the individual to observers. These emotional states are a form of nonverbal communication. The recognition of emotional state of a human face has attracted increasing notice in pattern recognition, human-computer interaction and computer vision. A method for automatic recognition of facial expressions from face images by providing Discrete Wavelet Transform (DWT) features to a bank of five parallel neural networks is presented in [1]. Each Neural Network (NN) is trained to recognize a particular facial expression, so that it is most sensitive to that expression.

A new approach to facial expression recognition based on Stochastic Neighbor Embedding (SNE) is presented in [2]. SNE is used to reduce the high dimensional data of facial expression images into a relatively low dimension data and Support Vector Machine (SVM) is used for the expression classification. A new approach for the 3D human facial expressions analysis is presented in paper [3]. The methodology is based on 2Dand 3D wavelet transforms, which are used to estimate multi-scale features from real a face acquired by a 3D scanner. The different feature extraction techniques with advantage and disadvantage and find the recognition rate by using JAFFE databases is studied in [4]. The Adaboost classifier is used to classify the facial expression and from the JAFFE databases 60% data are used for the training and 40% data are used for the testing purpose.

Various feature representation and expression classification schemes to recognize seven different facial expressions, such as happy, neutral, angry, disgust, sad, fear and surprise, in the JAFFE database is investigated in [5]. A facial expression recognition system based on Gabor feature using a novel Local Gabor filter bank is proposed in [6]. A two-stage classifier for the elastic bunch graph matching based recognition of facial expressions is proposed in [7]. The distinctive similarity between image patterns are obtained by applying optimal weights to responses from different Gabor kernels and those from different fiducial points.

An algorithm based on Gabor filter and SVM is proposed for facial expression recognition in [8]. First, the features of facial expression emotion are represented by Gabor filter. Then the features are used to train the SVM classifier. Finally, the facial expression

is classified by the SVM. A new method of facial expression recognition based on local binary patterns (LBP) and Local Fisher Discriminant Analysis (LFDA) is presented in [9]. The LBP features are firstly extracted from the original facial expression images. Then LFDA is used to produce the low dimensional discriminative embedded data representations from the extracted high dimensional LBP features with striking performance improvement on facial expression recognition tasks.

The performance of different feature extraction methods for facial expression recognition based on the higher-order local auto correlation (HLAC) coefficients and Gabor wavelet is investigated in [10]. An experiment on feature-based facial expression recognition within an architecture based on a two-layer perceptron is reported in [11]. The geometric positions of a set of fiducial points on a face, and a set of multi-scale and multi-orientation Gabor wavelet coefficients' at these points are used as features. A method of facial expression recognition based on Eigen spaces is presented in [12].

In this paper, an automatic classification of human emotion based on UWT and KNN classifier is presented. The remainder of this paper is organized as follows: The methodologies and proposed method is described in sections 2 and 3 respectively. The comparative study between DWT and UWT is given in section 4. Finally, conclusion is given in section 5.

2. Methodology

The proposed system for the classification of human emotion is built based on DWT, UWT and KNN classifier. This section gives the background of the methodologies used in the proposed emotional recognition system.

### 2.1 Discrete Wavelet Transform

Nowadays, wavelets have been used frequently in image processing and used for feature extraction, denoising, compression, face recognition, and image super-resolution. The decomposition of images into different

frequency ranges permits the isolation of the frequency components into different sub-bands. This process results in isolating small changes in an image mainly in high frequency sub-band images.

The 2-D wavelet decomposition of an image is performed by applying 1-D DWT along the rows and then columns. At first, 1-D DWT is applied along the rows of the input image. This is called row-wise decomposition. Then, 1-D DWT is applied again along the columns of the resultant image. This is called column-wise decomposition. This operation results in four decomposed sub-band images referred to as low–low (LL), low–high (LH), high–low (HL), and high–high (HH). For multi resolution analysis, the LL band of previous level is again decomposed by DWT. Figure 1 (a) shows the original image and Figure 1 (b) shows the wavelet transformed image at level 1.

2.2 Undecimated Wavelet Transform

The concept behind the undecimated wavelet transform is no decimation. The procedure for decomposing an image by undecimated wavelet transform is same as discrete wavelet transform. However, the main difference in undecimated wavelet transform is, it omits both down-sampling in the forward and up sampling in the inverse wavelet transform. More precisely, the transform is applied and all detail coefficients are saved and the low-frequency coefficients are used for the next level. The size of all sub-bands does not reduce from level to level. By using all coefficients at each level, it produces more accurate information for the frequency localization.
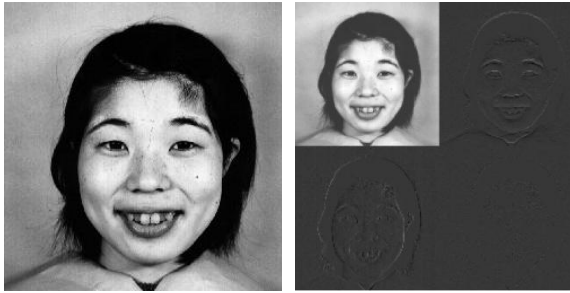
(a)   (b)

Figure 1: (a) Sample image from JAFFE database (b) 2-D Wavelet transformed image at level 1

2.3 K-Nearest Neighbour Classifier

The k-nearest neighbour algorithm (KNN) is a method for classifying objects based on closest training examples in the feature space. KNN is a type of instance-based learning where the function is only approximated locally and all computation is deferred until classification. In KNN, an object is classified by a majority vote of its neighbours, with the object being assigned to the class most common amongst its k nearest neighbours (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of its nearest neighbour. The neighbours are taken from a set of objects for which the correct classification is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. The following distance measures are used in the proposed emotion recognition system to evaluate the performance.

2.3.1 Euclidean distance

If the features have n-dimensions such as $u = (x_1, x_2, x_3, \dots x_n)$ and $v = (y_1, y_2, y_3, \dots y_n)$ then the generalized Euclidean distance formula between the feature points is given by

$$Euclidean\ Distance(u,v) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

(1)

2.3.2 City block Distance

If the features have n-dimensions such as $u = (x_1, x_2, x_3, \dots x_n)$ and $v = (y_1, y_2, y_3, \dots y_n)$ then the generalized Manhattan or city block formula between the feature points is given by

$$MH(u,v) = |x_1 - y_1| + |x_2 - y_2| + \cdots \dots |x_n - y_n|$$

$$= \sum_{i=1}^{n} |x_i - y_i|$$

(2)

2.2.3 Cosine

Let us consider the feature vector X and Y where $X = (x_1, x_2, x_3, \dots x_n)$ and $Y = (y_1, y_2, y_3, \dots y_n)$ then $\cos\theta$ may be consider as the cosine of the vector angle between the feature vectors X and Y in *n* dimension. The cosine of the vector angle between the features X and Y is given by

$$COSINE(X,Y) = \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2} \sqrt{\sum_i y_i^2}}$$

(3)

2.3.4 Correlation

A correlation is a single number that describes the degree of relationship between two feature variables X and Y where $X = (x_1, x_2, x_3, \dots x_n)$ and $Y = (y_1, y_2, y_3, \dots y_n)$. The correlation r between the features is defined by

$$r = \frac{N\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{[N\Sigma x^2 - (\Sigma x)^2][N\Sigma y^2 - (\Sigma y)^2]}}$$

(4)

3. Proposed System

The proposed system for the classifications of human emotion mainly consist of two stages namely features extraction and emotion recognition. In the

proposed emotion recognition system, the fused energies of all wavelet and undecimated wavelet sub-bands are used as features and KNN classifier is used to classify the state of the human emotion. Figure 2 shows the block diagram of the proposed human emotion recognition system.

3.1 Feature Extraction stage

In pattern recognition and machine learning problems, the essential preprocessing step is feature extraction. In the proposed approach, energies are extracted from the sub-bands of wavelet as well as undecimated wavelet decomposed image obtained from different resolution levels. The resolution level used in the proposed system varies from 1 to 8. The extracted energy features from the wavelet and undecimated decomposed image for a particular resolution level are fused together to form the feature vector. This feature vector is used to recognize the emotion of human. The energy is calculated by using the eqn. (5)

$$ENERGY_k = \frac{1}{RC} \sum_{i=1}^{R} \sum_{j=1}^{C} |S_k(i,j)|$$

(5)

where $S_k(i,j)$ is the pixel value of the kth sub-band and R, C is width and height of the sub-band respectively. Emotional Database (ED) is constructed by using the training emotional images. This database is used in the emotion recognition stage to recognize the emotion of the human face.

3.2 Emotion recognition stage

In emotion recognition stage, the proposed DWT and UWT features are extracted from the facial image. Then the DWT and UWT energy features of the unknown facial image are fused together and processed with the features in the ED generated in feature extraction stage. The emotional state of the given face image is classified using a KNN classifier, in which the distance between the features and the

corresponding features in the ED is calculated by various distance measure The distance measure used in the proposed system are Euclidean distance, cosine distance, city block and correlation measures. The classification performance is measured as the percentage of test images recognized into the correct facial expression.
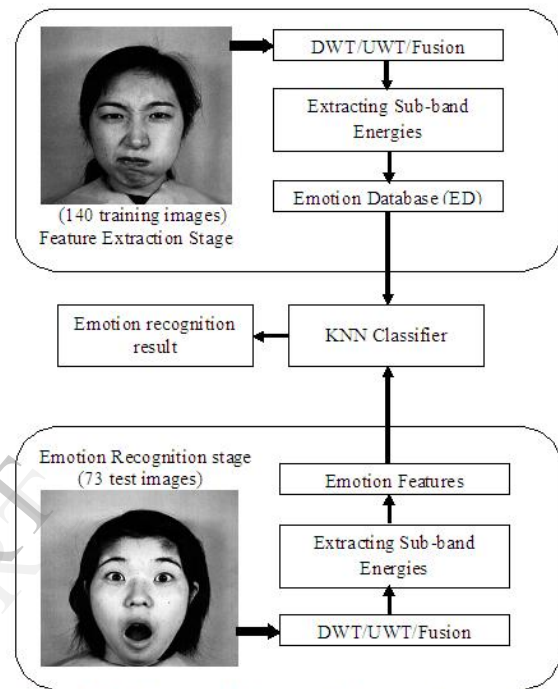

Figure 2 : Block diagram of the proposed human emotion recognition system

4. Experimental Results

The JAFFE database [13] is used to evaluate the performance of the proposed system. The database contains 213 images of 7 facial expressions. The facial expressions in this database are happiness, sadness, surprise, anger, disgust, fear and neutral. The images in the database are grayscale images of size 256x256 in the tiff format. The heads of the subjects in the images are in frontal pose. The eyes are roughly at the same position with a distance of 60 pixels in the final images. The proposed system is implemented in MATLAB version 7.10. Many computer simulations and experiments with JAFFE images are performed.

All the images in the JAFFE database are considered for the emotion recognition test. Among the 213 images

140 images from 7 facial expressions are used for training the classifier and remaining 73 images are used for testing the classifier. Table 1 to table 4 shows the average recognition rates obtained by the proposed emotion recognition system by using Euclidean distance, cosine, city block and correlation distance measure respectively.

Table 1: Average classification rate of the proposed emotion recognition system using Euclidean distance measure

| Level of Decomposition | KNN distance measure = Euclidean measure | | |
| | Average recognition rate (%) | | |
	DWT	UWT	Fusion
1	73.58	76.45	74.05
2	75.00	80.11	75.00
3	73.55	82.00	74.94
4	77.28	82.90	77.31
5	80.08	85.62	80.97
6	84.02	84.75	83.74
7	84.72	84.74	86.62
8	88.49	85.57	89.42

Table 2 : Average classification rate of the proposed emotion recognition system using cosine distance measure

| Level of decomposition | KNN distance measure = cosine distance | | |
| | Average recognition rate (%) | | |
	DWT	UWT	Fusion
1	75.38	77.36	76.80
2	78.61	83.40	81.45
3	79.51	84.30	83.71
4	82.83	84.21	85.64
5	84.78	85.17	87.96
6	85.67	87.50	88.42
7	88.10	83.86	90.27
8	89.39	85.21	91.67

Table 3: Average classification rate of the proposed emotion recognition system using city block distance measure

| Level of decomposition | KNN distance measure = city block | | |
| | Average recognition rate (%) | | |
	DWT	UWT	Fusion
1	74.53	75.97	75.00
2	75.47	80.61	74.51
3	74.51	82.95	76.88
4	78.68	83.80	81.92
5	82.34	85.21	84.23
6	84.24	85.24	87.51
7	86.55	86.63	87.52
8	91.25	88.35	91.27

While using the Euclidean distance measure, the fusion of energy features provides higher classification rate for higher decomposition level of 7 and 8 only. For city block distance measure, the higher classification rate is achieved for 6 to 8 decomposition level. However, the distance measures cosine and correlation provides higher recognition rate for all the decomposition level. From the table it is concluded that the proposed fusion of DWT and UWT based features provides better performance than DWT and UWT features. Figure 3 shows the average recognition rates for various distance measure in fusion of DWT and SWT. Among the 4 distance measure, the cosine distance measure is far better than others used in KNN classifier.

Table 4 : Average classification rate of the proposed emotion recognition system using correlation distance measure

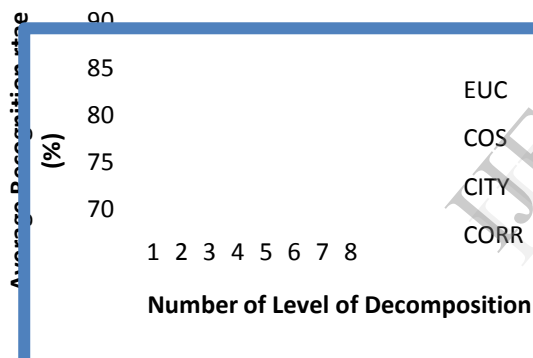| Level of Decomposition	KNN distance measure = correlation measure		
	Average recognition rate (%)		
	DWT	UWT	Fusion
1	72.16	73.15	73.55
2	75.87	81.00	80.00
3	79.05	81.01	82.36
4	82.86	81.95	86.61
5	86.60	85.17	87.09
6	86.13	87.97	89.38
7	88.39	83.87	90.72
8	89.83	85.21	91.67

Figure 3: Average recognition rates for various distance measure in fusion of DWT and SWT

5. Conclusion

In this paper, a novel feature fusion approach for the classification of human emotional state based on DWT and UWT is presented. The proposed method considers the energy features of all sub-bands of DWT and UWT decomposed image. The fusion of these two energy features outperforms DWT and UWT based features in terms of average recognition rate. The proposed system is implemented in MATLAB and JAFFE database is used for evaluation. Experimental results show that the proposed fusion based approach produces higher classification rate approximately 2% over than the DWT based features and 5% over than the UWT based features. Also, as the resolution level increases the classification rate of UWT based system decreases compared to DWT based features.

References:

[1] Sidra Batool Kazmi, Qurat-ul-Ain and M. Arfan Jaffar, "Wavelets Based Facial Expression Recognition Using a Bank of Neural Networks", 5th International Conference on Future Information Technology (FutureTech), May 2010

[2] Mingwei Huang, Zhen Wang and Zilu Ying, "Facial Expression Recognition Using Stochastic Neighbor Embedding and SVMs", International Conference on System Science and Engineering (ICSSE), June, 2011

[3] S. C. D. Pinto, J. P. Mena-Chalco, F. M. Lopes, L. Velho and R. M. Cesar Junior, "3D Facial Expression Analysis by Using 2d And 3d Wavelet Transforms", IEEE International conference Image Processing, 2011

[4] Aruna Bhadu, Rajbala Tokas and Dr. Vijay Kumar, "Facial Expression Recognition Using DCT, Gabor and Wavelet Feature Extraction Techniques", International Journal of Engineering and Innovative Technology (IJEIT), vol. 2, Issue 1, July 2012

[5] Frank y. Shih , chao-fa chuang and Patricks.P.Wang, "Performance Comparisons Of Facial Expression Recognition In Jaffe Database", International Journal of Pattern Recognition and Artificial Intelligence, vol. 22, No. 3 (2008) ,pp 445–459

[6] Hong-Bo Deng, Lian-Wen Jin, Li-Xin Zhen and Jian-Cheng Huang, "A New Facial Expression Recognition Method Based on Local Gabor Filter

Bank and PCA plus LDA", International Journal of Information Technology,vol.11,No. 11,2005

[7] Fan Chen and Kazunori Kotani, "Facial Expression Recognition by SVM-based Two-stage classifier on Gabor Features", IAPR Conference on Machine Vision Applications, May 16-18, 2007

[8] Xue Weimin, "Facial Expression Recognition Based on Gabor Filter and SVM", Chinese Journal of Electronics, vol.15, No.4A, 2006

[9] Shiqing Zhang, Xiaoming Zhao and Bicheng Lei, "Facial Expression Recognition Based on Local Binary Patterns and Local Fisher Discriminant Analysis", WSEAS Transactions on Signal Processing, vol.8,issue.1,January 2012

[10] Seyed Mehdi Lajevardi and Zahir M. Hussain, "Facial Expression Recognition: Gabor Filters versus Higher-Order Correlators", International Conference on Communication, Computer and Power (Icccp'09), February 15-18, 2009

[11] Zhengyou Zhang, "Feature-Based Facial Expression Recognition: Sensitivity Analysis and Experiments with a Multi-Layer Perceptron", International Journal of Pattern Recognition and Artificial Intelligence, 13(6), pp 893-911, 1999

[12] G. R. S. Murthy, R.S.Jadon, "Effectiveness of Eigen spaces for Facial Expressions Recognition", International Journal of Computer Theory and Engineering, vol.1, No.5, December 2009

[13] JAFFE database : http://www.kasrl.org/jaffe.html

[14] Picard, R. W. (1997). *Affective Computing*. Cambridge,MA: The MIT Press.

[15] E.Pandian,Dr.S.Santhosh Baboo, "Undecimated Wavelet Transform based Classification of Human Emotion", Proceedings of 2012 International Conference On Advanced Computer Science and Information Systems, ISBN : 978-979-1421-15-7,IEEE Catalog Number : CFP1219R-PRT.