Australian Journal of Basic and Applied Sciences, 11(5) April 2017, Pages: 1-11

AUSTRALIAN JOURNAL OF BASIC AND APPLIED SCIENCES

ISSN:1991-8178 EISSN: 2309-8414
Journal home page: www.ajbasweb.com

Emotion Recognition Based on Texture Analysis of Facial Expressions Using Wavelets Transform

[1]Suhaila N. Mohammed and [2]Loay E. George

[1]*Assistant Lecturer, Computer Science Department, College of Science, Baghdad University, Baghdad, Iraq,*
[2]*Assistant Professor, Computer Science Department, College of Science, Baghdad University, Baghdad, Iraq,*

Address For Correspondence:
Suhaila N. Mohammed, Baghdad University, Computer Science Department, College of Science, Baghdad, Iraq.

A R T I C L E I N F O

Article history:
Received 18 January 2017
Accepted 28 March 2017
Available online 15 April 2017

Keywords:
Facial Emotion, Face Detection, Template Based Methods, Texture Based Features, Haar Wavelets Transform, Image Blocking, Neural Network.

A B S T R A C T

Background: The interests toward developing accurate automatic facial emotion recognition methodologies are growing vastly and still an ever growing research field in the region of computer vision, artificial intelligent and automation. Auto emotion detection systems are demanded in various fields such as medicine, education, driver safety, games, etc. Despite the importance of this issue it still remains an unsolved problem **Objective:** In this paper a facial based emotion recognition system is introduced. Template based method is used for face region extraction by exploiting human knowledge about face components and the corresponding symmetry property. The system is based on texture features to work as identical feature vector. These features are extracted from face region through using Haar wavelets transform and blocking idea by calculating the energy of each block The feed forward neural network classifier is used for classification task. The network is trained using a training set of samples, and then the generated weights are used to test the recognition ability of the system. **Results:** JAFFE public dataset is used for system evaluation purpose; it holds 213 facial samples for seven basic emotions. The conducted tests on the developed system gave accuracy around 90.05% when the number of blocks is set 4x4. **Conclusion:** This result is considered the highest when compared with the results of other newly published works, especially those based on texture features in which blocking idea allows the extraction of statistical features according to local energy of each block; this gave chance for more features to work more effectively.

INTRODUCTION

Due to the rapid development of technologies, it is being required to build a smart system for understanding human emotion (Ruivo *et al.*, 2016). There are different ways to distinguish person emotions such as facial image, voice, shape of body and others. Mehrabian explained that person impression can be expressed through words (verbal part) by 7%, and 38% through tone of voice (vocal part) while the facial image can give the largest rate which reaches to 55% (Rani and Garg, 2014). Also, he indicated that one of the most important ways to display emotions is through facial expressions; where facial image contains much information (such as, person's identification and also about mood and state of mind) which can be used to distinguish human inspiration (Saini and Rana, 2014).

Facial emotion recognition is an active area of research with several fields of applications. Some of the significant applications are: feedback system for e-learning, alert system for driving, social robot emotion recognition system, medical practices...etc (Dubey and Singh, 2016).

Human emotion is composed of thousands of expressions but in the last decade the focus on analyzing only seven basic facial expressions such as happiness, sadness, surprise, disgust, fear, natural, and anger (Singh and

To Cite This Article: Suhaila N. Mohammed and Loay E. George., Emotion Recognition Based on Texture Analysis of Facial Expressions Using Wavelets Transform. ***Aust. J. Basic & Appl. Sci.**, 11(5): 1-11, 2017*

Kushwaha, 2016). Many approaches have been introduced for building and developing facial expression systems. Most of these approaches include three primary stages: (1) image preprocessing and face region detection, (2) facial feature extraction and representation, and (3) facial expression classifier (Kandemir and Ozmen, 2013).

In order to recognize class label for a given pattern, there is a need to extract suitable information from the pattern and produce feature values. Feature information is generated in two ways: (1) Appearance based features that using color or texture information from face region to infer the facial expression, and (2) Geometry based features in which the geometric relationship between certain key points (fiducial points) on the face is calculated when making its decision (Gosavi, 2015).

The appearance-based classification methods use features that appear temporarily in the face during any kind of facial expression (for example: the presence of specific facial wrinkles, bulges, forefront and the texture of the facial skin in regions surrounding the mouth and eyes). Transforms filters, such as Haar wavelets and integral image filters, are applied to ROI region, which either forms the overall face area or specific parts of the face region, to extract feature vector (Verma and Sharma, 2013).

In 2011, Thai *et al*. (Thai *et al*., 2011) have proposed facial based emotion recognition system. In preprocessing stage, Canny filter is used for local regions detection. Then, each of local region's features is presented based on Principal Component Analysis (PCA). Finally, Artificial Neural Network (ANN) is applied for facial expression classification. Japanese Female Facial Expression (JAFFE) database is used for testing and the best tests results up on the seven basic emotions gave accuracy equal to 85.7%.

In 2011, Novakovica *et al*. (Novakovica *et al*., 2011) have used Principal Component Analysis (PCA) for feature selection and dimensionality reduction in input data and neural network to classify the given facial image into one of the basic categories of facial expressions. Simulation experiment results showed that the use of PCA and neural networks in emotion recognition using JAFFE database, gave a recognition rate of approximately 85% when testing six emotions.

In Das (Das, 2014) proposed a person independent facial expression recognition system through using the modified Active Shape Model (ASM) to extract 68 landmark points from the face. These points are used to segment face region into seven sub-regions. After that each of these sub-regions is effectively represented by Census Transformation (CT) that based on feature histogram. Support Vector Machine (SVM) classifier with exponential chi-square weighted merging kernel was used to assign a weight based on the importance of each sub-region for final classification purpose. The proposed method was evaluated on JAFFE database using the seven basic emotions. The experimental results showed that the recognition rate of this approach was 86.3%.

In 2016, Sim and Cho (Sim and Cho, 2016) have proposed facial recognition system in which "Haar-like" feature was used for face detection, then Local Binary Pattern (LBP) Histogram and Canny edge detection were used for Feature Extraction. The extracted feature vector is then supplied to ANN in order to classify it to certain emotion type. 30 data items of specific expressions of anger, happiness, and expressionless (neutral) was selected from JAFFE dataset for system accuracy measurement and the attained recognition rate was 75%.

In this paper, texture based features are generated using Haar wavelets transform for emotional class identification in which these features are extracted from each block of each band of Haar wavelets result, this gave an effective role on increasing the system ability for recognizing the correct emotion for the inputted facial image.

The Proposed System:

The proposed system consists of three primary stages which are: preprocessing and region of interest (ROI) extraction, feature extraction, and classification stage. Figure (1) shows the basic stages that composed the system.

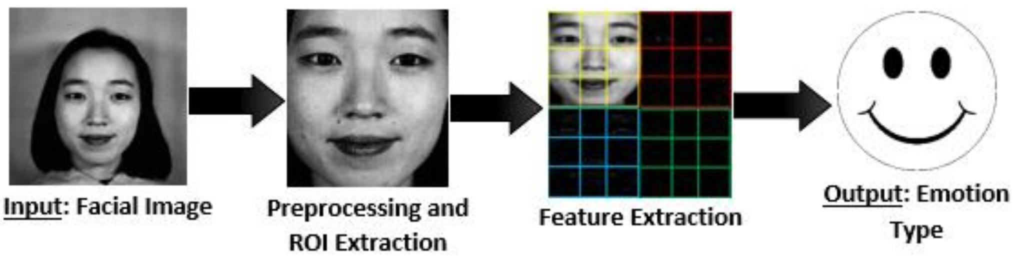


Fig. 1: Basic system design

Image enhancement and preprocessing steps have been used as a first stage in the proposed system to facilitate the process of face detection step. Template based method has been used for face region (ROI) extraction.

After that, the allocated face region is passed into Haar wavelet transform to convert it into scale-shift sub-bands; each sub-band contains certain texture information about certain face region. Then, each sub-band of Haar transform is portioned into NxM blocks and certain statistical features are extracted from each block to assemble the final feature vector. Finally, ANN has been used in classification stage to give the classification result of the system.

Preprocessing and ROI Extraction Stage:

To extract the face region from whole facial input image, the image must be passed through many steps which are based on some image processing methods; and as the following:

Firstly, the shadow that is resulted from bad lighting is reduced by subtracting the input image from its smoothed variant (i.e., the image that is resulted after applying mean filter of size 11x11).

The result of subtraction step involves noise of slat type, so that, mean filter is applied again for noise reduction purpose. Mean filter blurs the input image; this blurring is useful for background texture removable.

Smoothed image range is stretched to be within [G_{min}, G_{max}] range where

$$G_{min} = \mu - \alpha * \sigma \qquad (1)$$

$$G_{max} = \mu + \alpha * \sigma \qquad (2)$$

Where: μ, σ are the mean and standard deviation values, respectively. Stretching process is done using contrast stretching method; i.e., by applying the following equation:

$$I_s(x,y) = 255\left(\frac{I(x,y) - G_{min}}{G_{max} - G_{min}}\right) \qquad (3)$$

Finally, the binary version of the stretched image is generated through using threshold value (T) to split pixels' values to be either 0 (background region) and 1 (foreground region) and as following equation:

$$B(x,y) = \begin{cases} 1 & I_s(x,y) > T \quad \text{if} \\ 0 & \text{otherwise} \end{cases} \qquad (4)$$

To detect face region, the template based method is used to achieve this goal. The final binary image must be segmented into isolated segments before searching about the pattern that represent face region. The seed filling technique which belongs to region growing type of segmentation methods is used to divide the binary image into isolated regions based on pixels' values (pixels' colors); in which the set of connected pixels with same color and extended within bounded region is considered as an isolated segment.

Finally, the knowledge-based method for allocating face region in a single face image is applied. This method requires extraction of facial components (such as eyes, nose, mouth, etc.) which are based on human face morphology and symmetry property (i.e., the face always contains two eyes and one mouth) to determine the face region accurately. Figure (2) illustrates the face pattern that used in face detection step. Face region is then determined by searching for face template within resulted segments and then extracting the final face region. Figure (3) shows an example for preprocessing steps' outputs, respectively.

Feature Extraction Stage:

In this stage, the features required to recognize different facial expressions are determined. The features must have high discrimination ability to classify a particular emotion. The number of features used to achieve high classification decisions must be as small as possible.

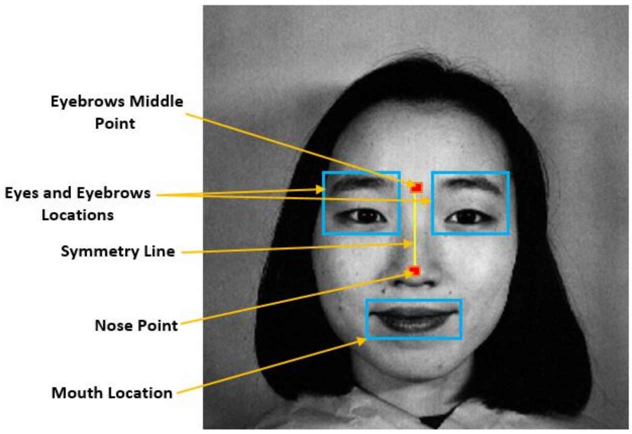


Fig. 2: Face Pattern

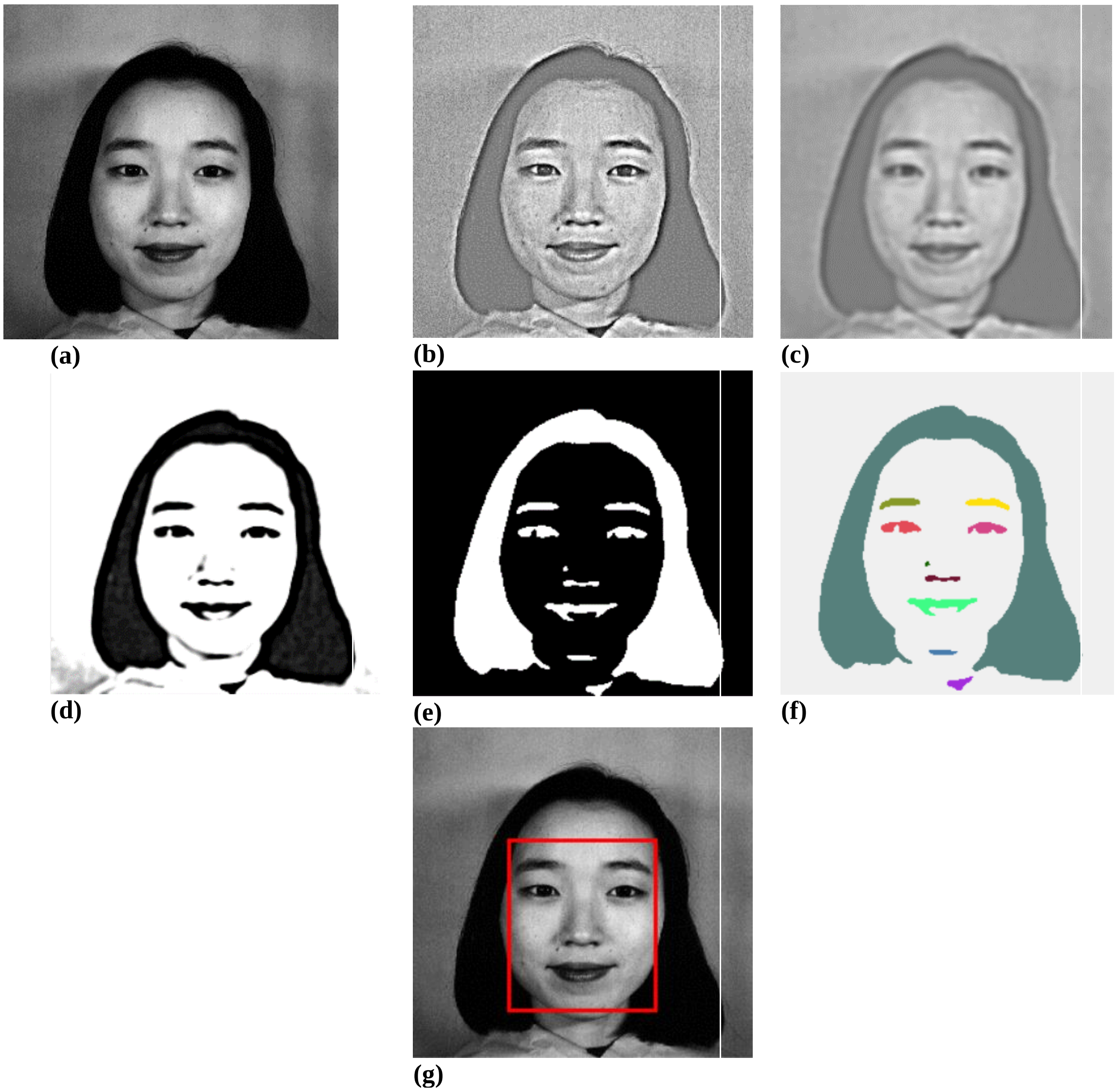


Fig. 3: Preprocessing and face region detection steps' results

In the proposed system, texture based features are extracted from face region through using wavelets transform. Feature extraction stage is composed of three primary steps: (i) Haar wavelets transformation, (ii) feature vector generation, and (iii) feature vector normalization.

Haar Wavelets Transformation:

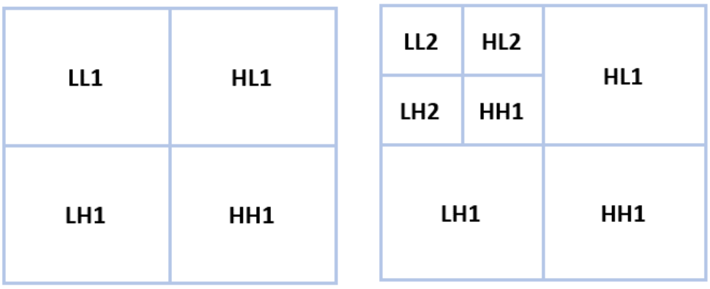
Discrete wavelet transform (DWT) is a mathematical tool that has a great interest in the field of image processing due to its good features. The introduction of DWT had improved some specific applications of image processing by replacing the existing traditional tools with this new mathematical transform by exploiting wavelets characteristics. Some of these characteristics are:

(1) It allows image multi resolution representation in a natural way because more wavelet sub bands are used to progressively enlarge the low frequency sub bands;

(2) It supports wavelet coefficients analysis in both space and frequency domains, thus the interpretation of the coefficients is not constrained to its frequency behavior, and we can perform better analysis for image vision and segmentation; and

(3) For natural images, the DWT achieves high compactness of energy in the lower frequency sub bands, which is extremely useful in applications such as image compression (Goyal and Aggarwal, 2012).

By applying DWT, the image is actually divided (i.e., decomposed into four sub bands) as shown in Figure (4.a). These four sub bands arise from separable applications of vertical and horizontal filters. The sub-bands labeled LH, HL and HH represent the finest scale wavelet coefficients (i.e., detail images) while the sub-band LL corresponds to coarse level coefficients (i.e., approximation image). To obtain the next coarse level of wavelet coefficients, the sub band LL alone is further decomposed and critically sampled. This leads to two-level wavelet decomposition are shown in Figure (4.b) (Singh and Kushwaha, 2016).



(a) **(b)**

Fig. 4: DWT bands; **(a)** one- level wavelet decomposition, **(b)** two- level wavelet decomposition

The Haar wavelet is a certain sequence of rescaled "square-shaped" functions which together form a wavelet family. The Haar wavelet's mother wavelet function ψ (t) can be considered as (Goyal and Aggarwal, 2012):

$$\psi(t) = \begin{cases} 1 & 0 \le t < \frac{1}{2} \\ -1 & \frac{1}{2} \le t < 1 \\ 0 & otherwise \end{cases} \qquad (5)$$

Haar wavelets transform with one level is applied on the detected face region to decompose it into four sub bands and extract the texture features as shown in Figure (5).

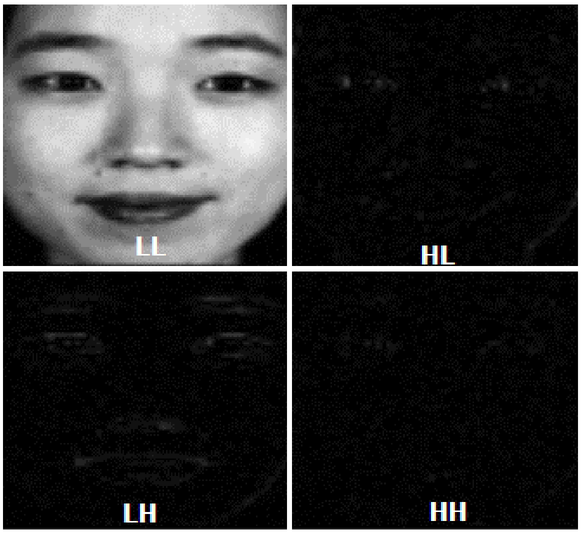


Fig. 5: Haar wavelets (with one level) result

Feature Vector Generation Step:

To generate feature vector with emotion recognition ability, the following steps were applied:

Each band of wavelets transform is divided into NxM blocks (as shown in Figure 6 where each sub band is divided into 3x3 blocks).

Statistical features are then generated from each block by calculating the energy of the block through averaging its pixels' values. By working upon block region we can take more effective local features. Table (1) shows the extracted features of an example image with number of blocks equal to 4x4 for each band.

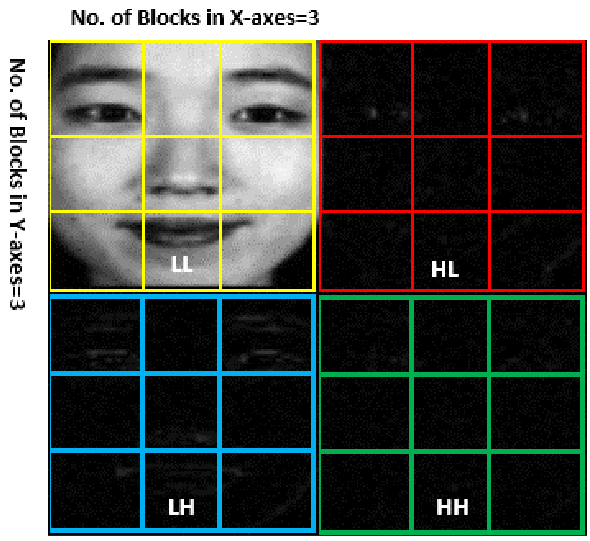


Fig. 6: Blocking idea, with 3x3 blocks

Table 1: The feature vectors for a sample from the used dataset within the seven basic emotional classes

fNo	Angry	Disgust	Fear	Happy	Normal	Sad	Surprise
f_1	76.057	78.633	86.512	109.56	100.701	94.158	105.9
f_2	116.954	132.365	139.126	168.774	138.918	143.893	184.758
f_3	122.764	130.756	98.22	101.632	127.315	137.805	155.242
f_4	51.338	63.155	40.722	45.197	69.518	70.311	85.701
f_5	154.519	124.839	114.648	142.109	173.255	154.636	144.85
f_6	138.248	152.327	143.07	166.016	162.107	167.133	183.521
f_7	139.981	145.362	110.234	115.281	158.211	153.911	161.64
f_8	73.18	79.485	81.506	99.548	86.45	81.796	104.732
f_9	177.811	167.327	154.655	182.954	193.814	170.944	185.622
f_{10}	170.571	171.807	138.534	169.326	189.631	179.63	189.078
f_{11}	170.673	157.864	109.54	135.959	174.908	171.231	171.588
f_{12}	107.707	88.649	97.505	126.09	105.255	99.403	113.723
f_{13}	139.611	128.045	104.125	124.905	160.488	141.397	110.407
f_{14}	171.063	149.215	186.448	205.042	152.41	153.371	212.664
f_{15}	212.83	177.896	142.947	149.322	179.563	164.839	194.636
f_{16}	127.889	99.172	81.956	117.345	125.308	133.926	134.879
f_{17}	5.243	5.84	5.369	4.648	5.651	5.346	6.188
f_{18}	4.685	4.545	4.288	3.605	5.541	5.026	3.83
f_{19}	4.676	3.463	4.041	4.288	4.392	4.098	3.97
f_{20}	2.945	3.912	3.282	3.795	4.234	3.93	3.876
f_{21}	4.081	5.02	7.039	6.052	4.455	4.199	6.562
f_{22}	4.686	6.107	5.331	5.142	5.981	5.329	4.18
f_{23}	4.662	4.991	4.99	5.863	4.329	4.419	4.717
f_{24}	4.075	4.188	3.908	3.965	3.967	3.98	4.413
f_{25}	4.356	5.692	5.958	4.914	4.187	4.149	5.534
f_{26}	5.46	6.546	4.819	4.516	6.328	6.324	4.18
f_{27}	5.635	5.148	4.857	5.844	4.596	4.672	4.484
f_{28}	4.319	5.335	4.364	4.174	4.313	5.431	4.016
f_{29}	4.018	4.92	4.364	4.516	4.659	4.756	6.632
f_{30}	4.03	4.472	3.567	3.206	4.832	4.452	2.826
f_{31}	3.261	3.733	3.472	3.282	3.447	3.778	3.246
f_{32}	4.473	4.228	4.307	5.104	3.904	3.795	3.736
f_{33}	5.333	6.943	5.435	7.356	7.366	7.471	8.01
f_{34}	4.414	4.611	2.84	3.188	5.688	4.958	3.269
f_{35}	2.392	2.962	2.799	3.555	3.467	3.154	3.596
f_{36}	2.153	2.879	2.636	3.147	3.643	3.828	3.97
f_{37}	6.359	6.908	5.619	6.6	5.56	5.684	8.056
f_{38}	5.213	4.661	3.923	4.23	5.512	4.621	3.666
f_{39}	5.944	6.19	5.496	8.663	6.439	5.835	5.838
f_{40}	5.086	4.409	4.72	4.168	4.649	6.375	7.636
f_{41}	5.401	4.944	3.698	3.514	4.362	4.857	4.18
f_{42}	4.004	3.77	4.536	5.129	4.633	4.469	3.339
f_{43}	6.39	7.035	7.172	11.933	6.008	5.313	6.165
f_{44}	6.148	6.044	6.252	5.701	6.119	6.611	8.827
f_{45}	6.556	9.771	8.05	9.91	8.356	7.994	12.236

f_{46}	4.962	5.636	3.535	3.371	7.19	5.194	3.339
f_{47}	2.847	3.328	3.024	4.536	3.307	3.204	3.106
f_{48}	3.782	3.888	4.23	3.984	3.643	3.862	4.11
f_{49}	2.592	4.021	4.117	3.776	3.818	4.199	4.974
f_{50}	2.644	3.563	3.491	3.074	4.354	4.149	3.596
f_{51}	2.18	3.618	3.396	3.415	3.788	3.457	3.503
f_{52}	1.532	3.205	3.017	2.732	3.042	3.255	3.363
f_{53}	2.625	3.967	3.738	4.686	3.728	3.66	5.628
f_{54}	2.769	4.011	3.358	3.776	4.19	3.913	3.62
f_{55}	2.446	3.865	3.586	4.762	3.996	3.609	4.203
f_{56}	2.655	4.133	3.719	3.548	3.743	3.845	3.946
f_{57}	3.059	4.513	4.554	4.345	3.639	4.183	3.76
f_{58}	3.498	4.266	3.719	3.491	4.369	4.064	3.526
f_{59}	2.648	3.668	4.06	4.364	3.594	3.626	3.923
f_{60}	2.715	4.247	4.44	3.453	3.445	4.284	3.876
f_{61}	3.304	4.156	3.074	4.402	3.624	4.048	5.184
f_{62}	2.308	3.65	2.998	2.751	3.504	4.149	3.339
f_{63}	2.298	3.174	2.599	3.017	2.908	3.069	3.526
f_{64}	2.528	2.943	3.055	3.169	2.818	3.255	3.503

Feature Normalization Step:

If the features are of different scales then the large scaled features will override the features of small scale and prevent them from participating, significantly, in final classification decision. So, performing the normalization process is very important since the feature vector that is proposed in this system contains features of different scales. The normalization process is done through applying flowing equation (Cowan, 2013):

$$f_n = \frac{f_r - \mu}{\sigma} \qquad (6)$$

Where, μ & σ are the mean and standard deviation of all classes' samples, f_r is the raw value of a specific feature before normalization, and f_n is the normalization result. Table (2) shows the mean and standard deviation values of each feature within all emotional classes.

Table 2: Mean and standard deviation values of feature vector within all classes

f_{No}	μ	σ	f_{No}	μ	σ
f_1	89.81769	18.89546	f_{33}	5.808004	0.95848
f_2	134.9484	26.07044	f_{34}	4.593112	0.96363
f_3	123.6944	29.90117	f_{35}	3.132928	0.424511
f_4	66.93924	22.13704	f_{36}	3.321405	0.73042
f_5	157.3608	23.82414	f_{37}	5.855473	1.068198
f_6	151.6115	18.51557	f_{38}	5.156853	0.835957
f_7	134.8428	25.80167	f_{39}	5.467602	1.025594
f_8	98.8828	20.21427	f_{40}	4.9887	0.902043
f_9	191.0449	17.32839	f_{41}	4.526886	0.762457
f_{10}	183.2038	14.27383	f_{42}	4.522194	0.651938
f_{11}	154.6823	23.43038	f_{43}	6.983904	1.53834
f_{12}	124.0142	23.09002	f_{44}	6.421633	1.189106
f_{13}	150.9992	19.53583	f_{45}	8.551916	1.232238
f_{14}	181.1815	17.62733	f_{46}	5.583135	1.45291
f_{15}	189.8678	27.70782	f_{47}	3.425482	0.786481
f_{16}	133.862	23.09843	f_{48}	4.286271	0.742482
f_{17}	4.874906	0.647705	f_{49}	3.686437	0.423837
f_{18}	4.701339	0.702805	f_{50}	3.646062	0.53481
f_{19}	4.330125	0.562435	f_{51}	3.293926	0.569037
f_{20}	4.023401	0.683892	f_{52}	3.097864	0.671609
f_{21}	4.718923	0.934417	f_{53}	3.731645	0.688536
f_{22}	5.780051	0.755463	f_{54}	3.909842	0.567234
f_{23}	4.783207	0.587479	f_{55}	3.479375	0.509947
f_{24}	4.245031	0.5194	f_{56}	3.437732	0.491929
f_{25}	4.542833	0.932703	f_{57}	3.696339	0.661959
f_{26}	5.214406	0.634191	f_{58}	3.764776	0.50395
f_{27}	4.861217	0.666192	f_{59}	3.635839	0.609164
f_{28}	4.583667	0.608402	f_{60}	3.631148	0.586989
f_{29}	5.386118	0.758532	f_{61}	4.104847	0.614304
f_{30}	4.53977	0.929042	f_{62}	3.629863	0.688491
f_{31}	3.930825	0.969338	f_{63}	3.265	0.693972
f_{32}	5.078738	1.027997	f_{64}	3.577465	0.683184

Classification Stage:

In this stage, the extracted feature pattern should be classified to an appropriate class. Since each class has more than one feature template, and the number of emotion classes is limited; the neural network classifier is adopted as proper classification tool; because it has good dynamic data clustering abilities, and high robustness against partial input variability.

The feed forward neural network with three layers was used in the proposed system. The number of nodes in input layer is equal to the number of features extracted during the feature extraction stage; taking into consideration the total number of used features is 64. The output is represented using binary coding; that is three output bits have been used to encode the index numbers of facial emotion classes [0...7]. The number of nodes in output layer is 3, as shown in Figure (7). The type of activation function used is the tangent hyperbolic function to weighting the input to hidden nodes and the sigmoidal function used to update the hidden signal output that sent to output nodes.

The applied classification process stage was passed through two phases: (i) training step and (ii) testing step. Training a neural network is the process of finding a proper set of weights' values for neural network nodes which makes it capable of making classification decisions very close to target values. The most common algorithm used to train feed-forward neural networks is called back-propagation. The supervised learning had been used for training purpose. A set of feature patterns belong to all classes was used for enrollment of the neural network.

In testing step, the trained neural network is tested using the training and testing set of the extracted feature vectors. The aim of conducted test is to assess the system performance.

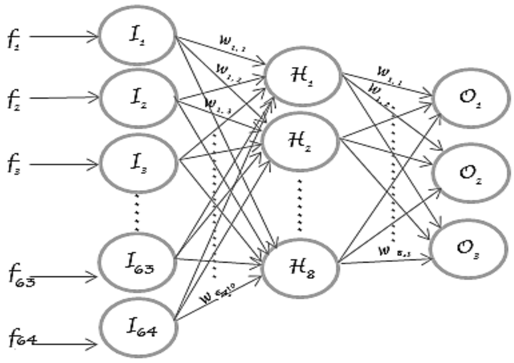


Fig. 7: Feed forward model of neural network

Experimental Results:

In this section, the results of some conducted tests are presented and discussed to evaluate the performance of the established system. In addition, a part of the tests was directed to explore the effects of the involved system parameters on the overall system performance. The programming language C sharp (Microsoft Visual Studio 2008) was utilized to develop the programs.

The dataset used for training and testing the system that proposed in this paper is Japanese Female Facial Expression (JAFFE) public dataset. It consists of 213 gray scale images for 10 Japanese female models, obtained in front of a semi reflective mirror. The samples have been downloaded from the standard reference database (Michael *et al*., 1998) in which a seven different facial expressions, such as neutral (NE), happy (HA), angry (AN), disgust (DI), fear (FE), sad (SA) and surprise (SU) are taken for each subject. Each female has posed from two to four times for each expression. Each image has been rated on 6 emotion adjectives by 60 Japanese subjects. The images were taken at the psychology department in Kyushu University. JAFFE dataset images are of tiff format with 8 bits/pixel (bit depth), and the size of each image is 256x256 pixels taken with resolution 100 dpi.

The weights of neural network are adjusted using the training dataset. The learning was performed first using all dataset samples to check maximum neural network classification ability. Then a sub-set of images was used as training set in order to test the system performance; here a subset of the overall database images was used as training material. The used training set consists of about 75% samples taken, randomly, from each class, and the remaining samples have been treated as testing samples.

The purpose of training step is to determine the proper neural network configuration parameters; which are: number of nodes in hidden layer (HL), learn rate (LR), momentum (Mom), and Epoch (Ep). Various combinations of these parameters have been tested to find the best recognition rate can be reached. Table (3) shows these parameters' values when tested with feature vectors resulted from different blocks size along with recognition rate (RR) for each case when training the system with all dataset samples. Figure (8) shows a chart bar for the effect of different blocks number on the proposed system accuracy.

Table 3: The optimal ANN configuration that selected when testing it with feature vectors that resulted from different blocks number along with RR for each case.

Number of blocks in X-axes	Number of blocks in Y-axes	RR	ANN configuration
1	1	59.72 %	HL= 14, LR= 0.1, Mom= 0.1, Ep=4000
1	2	82.46 %	HL= 15, LR= 0.1, Mom= 0.4, Ep=4000
2	1	83.89 %	HL= 14, LR= 0.1, Mom= 0.6, Ep=4000
2	2	94.79 %	HL= 15, LR= 0.1, Mom= 0.4, Ep=4000
1	3	91.47 %	HL= 14, LR= 0.1, Mom= 0.1, Ep=4000
2	3	97.16 %	HL= 11, LR= 0.2, Mom= 0.2, Ep=4000
3	1	94.79 %	HL= 15, LR= 0.1, Mom= 0.3, Ep=4000
3	2	96.68 %	HL= 15, LR= 0.2, Mom= 0.9, Ep=4000
3	3	99.05 %	HL= 10, LR= 0.2, Mom= 0.6, Ep=4000
4	1	94.79 %	HL= 14, LR= 0.1, Mom= 0.6, Ep=4000
4	2	98.58 %	HL= 10, LR= 0.2, Mom= 0.6, Ep=4000
4	3	99.53 %	HL= 12, LR= 0.2, Mom= 0.3, Ep=4000
4	4	100 %	HL= 8, LR= 0.2, Mom= 0.4, Ep=4000

As shown in Table (3) the system has recognition ability equal to 100% when learning it with all dataset samples and using number of blocks 4x4 (i.e., with number of features equal to 64 features)

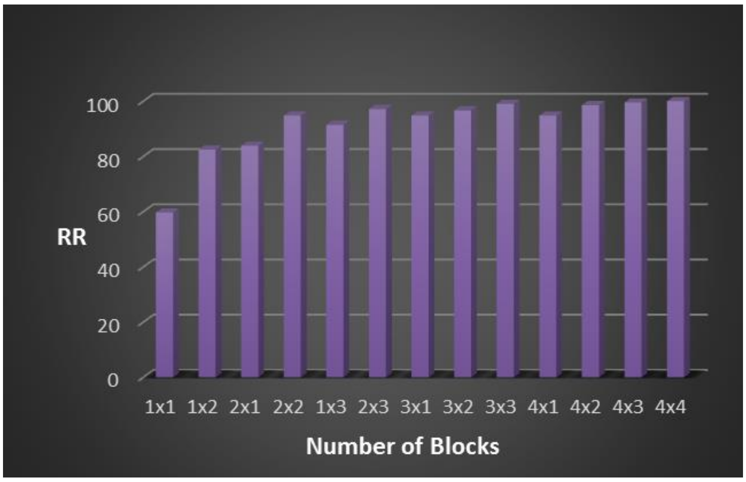


Fig. 8: The effect of blocks number on RR

Table (4) shows the confusion matrix that resulted when learning the system with 75% randomly selected samples from each emotion class and using the resulted neural weights for testing system recognition ability. The best RR was 90.05% with HL=8, LR=0.1, Mom=0.2 and Ep=4000. Figure (9) shows the error flow chart that is resulted from training the neural network.

Table 4: Confusion matrix for the distribution of samples within emotion classes

Class	AN	DI	FE	HA	NE	SA	SU
AN	93.3%	0%	6.25%	0%	3.45%	0%	0%
DI	3.33%	86.2%	3.13%	0%	0%	0%	0%
FE	0%	6.9%	81.3%	0%	6.9%	6.67%	0%
HA	0%	0%	0%	96.8%	3.45%	0%	3.33%
NE	0%	0%	6.25%	3.23%	86.2%	0%	0%
SA	3.33%	3.45%	3.13%	0%	0%	90%	0%
SU	0%	3.45%	0%	0%	0%	3.33%	96.7%

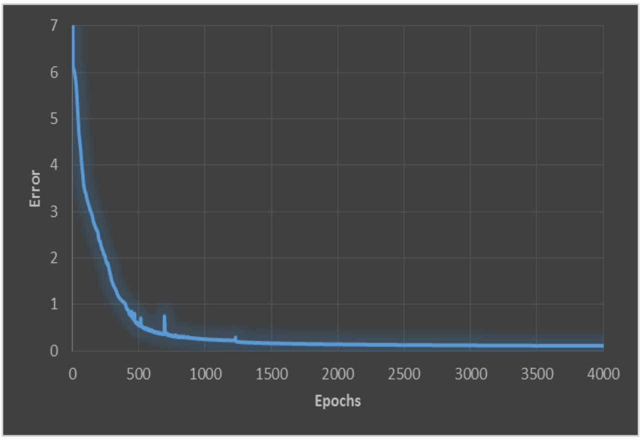


Fig. 9: Error flow chart that resulted when training the neural network

Table (5) shows the results of comparison that is made with some other researches related to person independent environment and applied on JAFFE dataset. The results listed in the table indicate the outperformance of the proposed system.

Table 5: Comparison of recognition rate with some other person's independent systems on JAFFE dataset

Authors	Feature Extraction Method	No. of Emotions	RR
Thai *et al.* (Thai *et al.*, 2011)	Canny filter and PCA	6 emotions	85.7%
Novakovica *et al.* (Novakovica *et al.*, 2011)	PCA	6 emotions	85%
Rojas *et al.* (Rojas *et al.*, 2012)	Local Sign Directional Pattern (LSDP)	7 emotions	89.2%
Das (Das, 2014)	Census Transformation (CT)	7 emotions	86.3%
Sim and Cho (Sim and Cho, 2016)	Local Binary Pattern (LBP) histogram and Canny edge detection	3 emotions	75%
The proposed system	Statistical features derived using Haar Wavelets transform	7 emotions	90.05%

Conclusions And Future Work:

In this paper a facial based emotion recognition system has been proposed. Template based method was used for face region detection through building a facial pattern and searching for it within the inputted image (after applying pre-processing steps up on it), when a match is found then this region is considered as face region (ROI). To extract texture based features from the extracted face region, Haar wavelets transform was used for image domain transformation and for producing wavelets sub-bands labeled LH, HL and HH those represent the finest scale wavelet coefficients (i.e., detail images) and the sub-band LL corresponds to coarse level coefficients (i.e., approximation image). Blocking idea allows the extraction of statistical features according to local energy of each block; this gave chance for more features to work more effectively through focusing on the texture that each emotion could make within the facial components and reflecting that on the proposed system effectiveness. The normalization step that applied on all elements of the feature vector has significant impact on the recognition ability of the proposed system. This normalization prevents the features having small scale from being overridden by high scaled ones. Neural network classification method has a great effect on emotion recognition task due to its ability to work with partial input variability that has more than one template. The test results showed the effectiveness of the proposed system where it gave a recognition rate equal to 100% when training it with all dataset samples and 90.05% when training the system with part of the dataset samples and then testing system ability using the generated weights. As a future work, texture features can be combined with geometric based features to assemble hybrid based features with more accurate recognition ability through improving the feature vector that proposed in the system to contain features generated by using geometrical measurements such as the distances between facial components, angles conveyed among facial components, etc. This improvement will increase the recognition ability of the system. Beside that another form of classifiers can be used rather than feed forward neural network classifier, for example: multilayer neural network, support vector machine, fuzzy classifier, etc. and their effects on the system accuracy can be studied.

REFRENCES

Cowan, M., 2013. Machine Learning. Creative Commons Attribution.

Das, D., 2014. Human's Facial Parts Extraction to Recognize Facial Expression. International Journal on Information Theory (IJIT), 3(3): 65-72.

Dubey, M. and L. Singh, 2016. Automatic Emotion Recognition Using Facial Expression: A Review. International Research Journal of Engineering and Technology (IRJET), 3(2): 488-492.

Gosavi, V., 2015. Review Paper on - Emotion Recognition. International Journal of Latest Trends in Engineering and Technology (IJLTET), 5(1): 213-220.

Goyal, N and M. Aggarwal, 2012. Implementation of Image Compression Using Fast Wavelet Transform Using Haar and Daubechies Wavelets. IOSR Journal of Engineering (IOSRJEN), 2(11): 18-24.

Kandemir, M and G. Ozmen, 2013. Facial Expression Classification with Haar Features, Geometric Features and Cubic Bézier Curves. IU-JEEE, 13(2): 1667-1673.

Michael, J.L., A. Shigeru, K. Miyuki and G. Jiro, 1998. Coding Facial Expressions with Gabor Wavelets. 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp: 200-205.

Novakovic, J., M. Minic and A. Veljovic, 2011. Classification Accuracy of Neural Networks with PCA in Emotion Recognition. Theory and Applications of Mathematics & Computer Science, 3(1): 11-16.

Rani, J and K. Garg, 2014. Emotion Detection Using Facial Expressions: A Review. International Journal of Advanced Research in Computer Science and Software Engineering, 4(4): 465-467.

Rojas, J.A., A.R. Rivera and O. Chae, 2012. Facial Expression Recognition Based on Local Sign Directional Pattern. IEEE, pp: 2613-2616.

Ruivo, J.P.P., T. Negreiros, M.R. Barretto and B. Tinen, 2016. A Model for Facial Emotion Inference Based on Planar Dynamic Emotional Surfaces. (IJARAI) International Journal of Advanced Research in Artificial Intelligence, 5(6): 45-54.

Saini, R and N. Rana, 2014. Facial Expression Recognition Techniques, Database & Classifiers. International Journal of Advances in Computer Science and Communication Engineering (IJACSCE), 2(2): 15-20.

Sim, S and J. Cho, 2016. Expression Recognition Based on Artificial Neural Network Using Error Back propagation Learning Algorithm. International Journal of Applied Engineering Research, 11(2): 820-823.

Singh, N and V. Laxmi, 2014. Review: Audio Noise Reduction Using DWT Technique with Hard and Soft Threshold. International Journal of Software and Web Sciences (IJSWS), 10(1): 28-31.

Singh, R and S. Kushwaha, 2016. Analysis and Recognition of Facial Feature Expressions. International Journal of Advanced Research in Computer Science and Software Engineering, 6(4): 315-329.

Thai, L.H., N.D.T. Nguyen and T.S. Hai, 2011. A Facial Expression Classification System Integrating Canny, Principal Component Analysis and Artificial Neural Network. International Journal of Machine Learning and Computing, 1(4): 388-393.

Verma, V and L.K. Sharma, 2012. A Comprehensive Survey on Human Facial Expression Detection. International Journal of Image Processing (IJIP), 7(2): 171-182.