Robust face recognition using wavelets and neural networks

Yesenia Sánchez Loredo
Polytechnic University of Victoria
Victoria, Tamaulipas, México.
1129017@upv.edu.mx

Rubén Machucho Cadena
Polytechnic University of Victoria
Victoria, Tamaulipas, México.
rmachuchoc@upv.edu.mx

***Abstract*— For humans recognize a person looking at his face is a trivial and daily task that brain performs in seconds and almost ever with high precision. Nevertheless, even when this task seems easy, face recognition remains as a challenge when implemented in a computer. In this article, a new methodology for accomplish two of the three stages involved in the design of an automatic face recognition system is presented: for feature extraction stage the discrete wavelet transform is used, and for the recognition stage a personalized multilayer neural network is proposed. Experiments were performed using 300 images from the public database of faces Faces94. The results show a recognition rate of 95%. This demonstrates that the proposed approach is a feasible and effective solution to the face recognition problem.**

***Keywords*— Face recognition; neural networks; wavelets**

I. INTRODUCTION

Face recognition is a research topic that in last years has been in the spotlight, due to this technology allow to identify a person with more accuracy than traditional systems, which are based in passwords, PIN's magnetic cards, or a combination of these. The disadvantage is that this elements can be transferred, damaged, lost, forgotten or stolen [1,2]. At this point, biometric systems appear like a robust option for recognize people facing the problems mentioned.

Among biometric treats that could be used for recognize a person, the face is a interesting characteristic to be studied due its properties of low intrusivity, naturalness, easy use and operability even without the voluntary cooperation of the person that we are trying to identify [3, 4].

Generally, a face recognition system consists of the following stages [5]: image acquisition, detection and segmentation of face, and finally, recognition or classification step.

In this stages, the feature extraction module is responsible for transforming the input data into a set of representative values and useful as possible for further processing in subsequent stages of the system [6, 7].

Among the techniques that have been presented to extract features, template-based methods are easy to implement but do not represent the overall facial structure; Color segmentation based methods use a color model to detect faces and morphological operations for detect features, so that if a different color model is used and variations in lighting are present this will result in a poor performance. On the other hand, appearance based methods represent optimal feature points that can represent the overall facial structure, but with the disadvantage of being computationally expensive.

The next stage in the flow of a recognition system is the classification, which aims to assign a label a test pattern [8].

In the case of face recognition system, the aim is to get the ID of the person registered on the system database and against which the test image has a higher degree of similarity.

With the goal of carrying out the two steps mentioned above, this work proposes the use of discrete wavelet transform as a robust feature extraction tool, due to its invariance to lighting issues, facial expression, scale and pose, and the design of a custom multilayer network for classification step.

The paper is organized as follows: section 2 presents essential concepts of discrete wavelet transform and its application in 2D signals, section 3 addresses topics related to artificial neural networks. Section 4 presents the proposed methodology and Section 5 shows the results using a set of test images of the Faces94 public database. Finally, in Section 6 we give the conclusions about the work done.

II. DISCRETE WAVELET TRANSFORM

The implementation of the discrete wavelet transform (DWT) is based on the use of filter banks for analysis and reconstruction of signals [9].

The DWT uses an orthonormal wavelet family; ie, the wavelets are orthogonal and normalized to have unit energy [10]:

$$\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k) \tag{1}$$

Where j and k are integers that scale and dilate the mother wavelet to generate discrete wavelet family.

The practical application of the DWT is performed implementing filter banks, whose main idea is to separate the signal into different frequency bands [11]. An example of a signal *X(n)*, which is decomposed into approximations *aj(n)* and details *dj(n)* when we apply over this the high-pass filters *Hj(n)* and low-pass filters *Gj(n)* is shown in Fig. 1, where *n* is an integer number and $j = \{1,2,3 \ldots k\}$ is the desired decomposition level. The union of the two filter outputs contains the same frequency content that the original input signal, however, the amount of data is duplicated. Therefore, decimation by a factor of two, represented as ↓2 is applied to the analysis filter bank output. The filter bank can be expanded to an arbitrary level depending on the desired resolution.

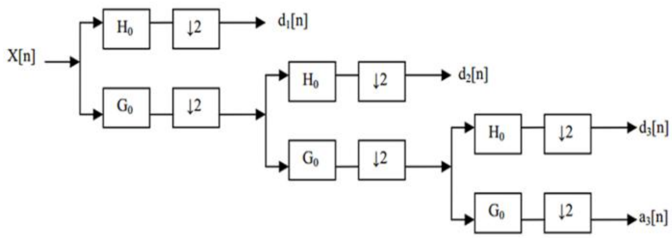

Fig. 1. Wavelet decomposition tree to 3 levels.

A. *Two dimensional DWT*

The two dimensional wavelet transform can be applied to images, resulting in a matrix of coefficients known as wavelet coefficients. Its implementation is performed by applying the transform to the rows and columns of the image, which are seen as one-dimensional signals; this generates four types of coefficients or sub matrices: approximation matrix (LL), horizontal details matrix (LH), vertical details matrix (HL) and diagonal details matrix (HH) [12].

III. Artificial Neural Networks

Neural networks are a robust and widely technique used in image segmentation and classification problems. Its structure is essentially a collection of interconnected neurons that are executed in parallel to perform the task of learning [13, 14]. Its application is given in areas such as: interpreting visual scenes, speech recognition, control strategies, etc..

Among the major architectures, the multilayer perceptron is one of the most widely used, because it allow to solves real problems, in addition to its easy use and application. Regarding the training, it can be carried out using various algorithms, where the *backpropagation* technique is the most widely used, but this has the disadvantage of being very slow for its application to real problems. Because of this, there have been some optimizations to the original *backpropagation* algorithm, among which the Levenberg-Marquardt method stresses, since this achieves a faster convergence to a result, in contrast to any other existing optimization [15].

IV. Proposed Method

Fig. 4 shows the block diagram of the method proposed. As can be seen, before the step of feature extraction is desirable to apply some pre-processing techniques, with the goal to improve quality of the images and achieve higher recognition rates. The activities carried out at this stage were:

- Histogram equalization. This activity allowed to maximize the contrast of the images used in experiments, with the aim of improving their quality and get more accurate results at the stage of face detection and segmentation. In Fig. 2 some examples of images to which we applied histogram equalization are shown.

Fig. 2. Histogram equalization over the images used in experiments.

- Detection and segmentation of the face. We use the detection algorithm proposed by Viola-Jones, due to their ability to detect effectively and in real time objects of interest within an image with high accuracy [16]. In Fig. 3 some examples of images on which we applied the algorithm are shown.

Fig. 3. Detection and segmentation of the face using Viola-Jones algorithm.

- Normalizing face size. This activity was carried out using the bilinear interpolation method, in order to adjust the face previously detected and segmented to a predefined size.

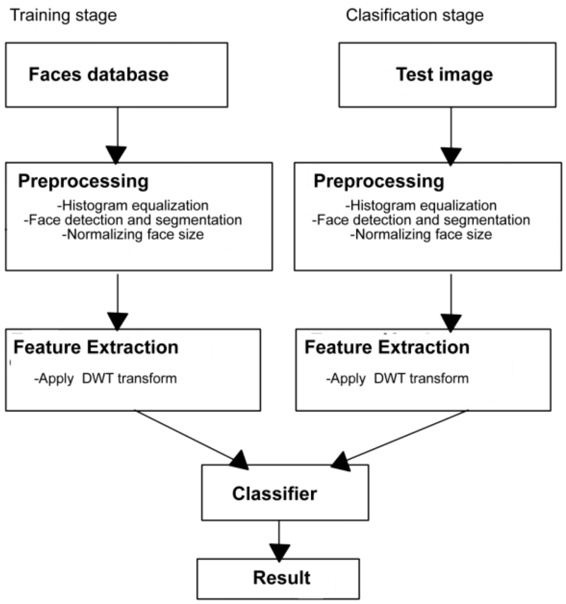

Fig. 4. Block diagram of the proposed methodology.

A. *Feature extraction*

Below is the algorithm proposed in the feature extraction stage:

1. Take the picture *I* resulting from the previous activity and set the number of decomposition levels to J.

2. Let *F* be the wavelet filter used for the decomposition.

3. Apply the discrete wavelet transform to *I*, using the low-pass and high-pass filters obtained from F, as many times as directed by J.

4. Take the approximation coefficients and divide it into 4 sets of equal size, discarding the detail coefficients.

5. If the number of features in the sets exceeds the amount of 25, proceed to perform sampling on each set.

B. *Sampling technique*

Below, is an example that shows how to apply the proposed sampling method: Setting the size of the face detected at 80 x 80 pixels and applying the wavelet transform using two levels of decomposition, the size of approximation coefficients region will be 20 x 20 pixels. The coefficients in this region will be sorted and grouped into four regions as shown in Fig. 5

1	2	**3**	4	**5**	6	**7**	8	**9**	10	**11**	12	**13**	14	**15**	16	**17**	18	**19**	20
21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
41	42	**43**	44	**45**	46	**47**	48	**49**	50	**51**	52	**53**	54	**55**	56	**57**	58	**59**	60
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
81	82	**83**	84	**85**	86	**87**	88	**89**	90	**91**	92	**93**	94	**95**	96	**97**	98	**99**	100
101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
121	122	**123**	124	**125**	126	**127**	128	**129**	130	**131**	132	**133**	134	**135**	136	**137**	138	**139**	140
141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160
161	162	**163**	164	**165**	166	**167**	168	**169**	170	**171**	172	**173**	174	**175**	176	**177**	178	**179**	180
181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200
201	202	**203**	204	**205**	206	**207**	208	**209**	210	**211**	212	**213**	214	**215**	216	**217**	218	**219**	220
221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240
241	242	**243**	244	**245**	246	**247**	248	**249**	250	**251**	252	**253**	254	**255**	256	**257**	258	**259**	260
261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280
281	282	**283**	284	**285**	286	**287**	288	**289**	290	**291**	292	**293**	294	**295**	296	**297**	298	**299**	300
301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320
321	322	**323**	324	**325**	326	**327**	328	**329**	330	**331**	332	**333**	334	**335**	336	**337**	338	**339**	340
341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360
361	362	**363**	364	**365**	366	**367**	368	**369**	370	**371**	372	**373**	374	**375**	376	**377**	378	**379**	380
381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400

Fig. 5. Proposed sampling method for sort and divide the approximations coefficients obtained after application of the transform wavelet.

The proposed sampling technique involves taking only the coefficients shaded with deeper color than the color of the background region. This should be carried out for each of the four regions. Finally, in the current example we can see that of the total of 100 data constituting each region, after applying the sampling, only 25 will remain.

That is, of the total of 400 characteristics obtained with the wavelet transform, only 100 will be considered for submission to the classifier. Fig. 6 shows the final feature vector.

1	3	5	7	9	11	13	15	17	19
41	43	45	47	49	51	53	55	57	59
81	83	85	87	89	91	93	95	97	99
121	123	125	127	129	131	133	135	137	139
161	163	165	167	169	171	173	175	177	179
201	203	205	207	209	211	213	215	217	219
241	243	245	247	249	251	253	255	257	259
281	283	285	287	289	291	293	295	297	299
321	323	325	327	329	331	333	335	337	339
361	363	365	367	369	371	373	375	377	379

Fig. 6. Feature vector to represent each face after applying the sampling.

C. *Classifier design*

For classification step a personalized multilayer neural network is proposed. Its architecture is shown in Fig. 7.

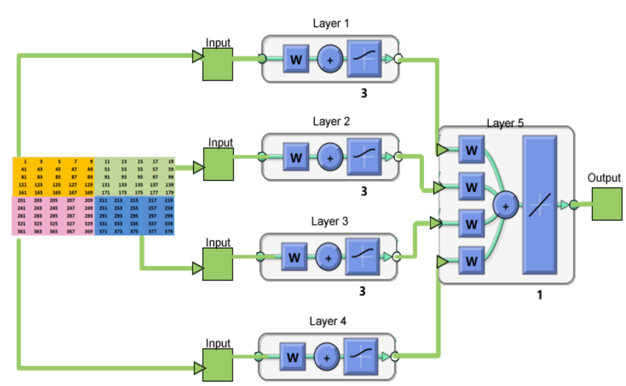

Fig. 7. Proposed personalized neural network architecture for classification stage

The network has 5 layers. The layers 1, 2, 3 and 4 are inputs layers, and their responsibility is to receive the input feature vector. The layer 5 gives the id of the picture in the database with which a test image is more similar.

The configuration network parameters are listed below:

- Layers 1, 2, 3 and 4 use *logsig* activation function and three neurons in each of these layers.
- Layer 5 receives the output of layers 1, 2, 3 and 4, and provides the final output of the network. This layer uses a *purelin* activation function.
- The training method is Levenberg-Marquardt.
- The network error is set to 0.01.

The main contribution of this network architecture is that, compared to a fully-connected multilayer network, the proposed design provides less weights to adjust, which directly achieves a reduction of the computational cost required for the training of the network since the amount of free parameters (weights) was decreased.

V. Implementation and Results

The proposed methodology was implemented in a desktop computer with an Intel Core i7 2.4 GHz, 12 GB of RAM and operating system Windows 8, 64-bit. The development platform used was MATLAB.

The tests were performed by taking 300 images of Faces94 database corresponding to 15 different people, each person has 20 shots; where 16 samples were used for training and the remaining 4 to measure the generalization capability of the network.

Experiments using different decomposition levels, wavelet basis and neurons in layers 1, 2, 3 and 4 were performed. The results obtained are shown in Table I, where we can observe that the best results were obtained using the third level of decomposition, the wavelet Bior 1.3 and 4 neurons in layers 1, 2 , 3 and 4 of the neural network.

TABLE I. Results using different levels of decomposition and bases wavelets

Decomposition Level	**Wavelet**	**Number of neurons in layers 1, 2 3 and 4**	**Recognition Rate using**	
			Training patterns	***Test patterns***
2	Daub 4	3	100%	87%
		4	100%	92%
	Bior 1.3	3	100%	77%
		4	100%	80%
	Coif 5	3	100%	72%
		4	100%	77%
3	Daub 4	3	100%	86%
		4	100%	90%
	Bior 1.3	3	100%	85%
		4	100%	95%
	Coif 5	3	100%	89%
		4	100%	90%

VI. Conclusions

We presented a methodology for achieve automatic face recognition, based on the use of discrete wavelet transform and artificial neural networks.

The conducted research proved that the combination of the wavelet Bior 1.3 and the third decomposition level gives a recognition rate of 95%, which demonstrated that use of the wavelet transform provides excellent tool for image decomposition and textures description.

Results also allow to validate that the proposed neural network architecture provides recognition rates comparable to those obtained with the use of fully-connected multilayer networks; typically used in face recognition research, but with the advantage that the number of weights to adjust in the proposed network is lower, which reduces the computing burden necessary for training the net.

Finally, it is worth mentioning that results obtained with the proposed methodology enable testing with other faces databases, and the use of various classifiers with the purpose of comparing the recognition rates offered by each of these.

References

[1] S. A. Jassim, Wavelet–Based Face Recognition Schemes, Face Recognition, Milos Oravec (Ed.), ISBN: 978-953-307-060-5, InTech, 2010.

[2] M. Ruiz-Echartea, Sistema de identificación automática de huella dactilar latente, Tesis de Maestría, Universidad Politécnica de Victoria, 2011.

[3] A. Jain and S. Li, Handbook of Face Recognition, Springer-Verlag, New York, Inc., Secaucus, NJ, USA, 2005.

[4] D. González, Sistemas de reconocimiento facial, DIntel, vol. 2, no. 1, pp. 118-119, 2009.

[5] A. Jain, A. Ross, and K. Nandakumar, Introduction to biometrics, Springer-Link : Bucher, Springer US, 2011.

[6] J. Leiva-Murillo, Extracción de características mediante criterios basados en teoría de la información, Universidad Carlos III de Madrid, 2007.

[7] L. Introna and H.Nissenbaum, Facial recognition technology: A survey of policy and implementation issues, Reporte Técnico, Report of the Center for Catastrophe Preparedness and Response, New York University, 2009.

[8] A. Jain, R. P. W. Duin, and J. Mao, Statistical pattern recognition: a review, Pattern Analysis and Machine Intelligence, IEEE Transactions, vol. 22, no. 1, pp. 4-37, 2000.

[9] R. Merry, Wavelet Theory and Applications. A literature study, 2005.

[10] R. A. González-González, Algoritmo basado en wavelets aplicado a la detección de incendios forestales, Universidad de las Américas Puebla, 2010.

[11] B. García-Meléndez, E. Mancilla-Ambrona, and R. Montes-Fraile, Optimización de la transformada wavelet discreta (DWT). Universidad Complutense de Madrid, 2005.

[12] Jagadeesh, H. S.; Babu, K. Suresh and Raja, K. B., Face recognition based on DWT and DBC, Computer Communication and Informatics (ICCCI), 2012 International Conference on , vol., no., pp.1,6, 10-12 Jan. 2012.

[13] J. Collantes-Duarte, Predicción con redes neuronales: Comparación con las metodologías de Box y Jenkins, Universidad de los Andes, 2001.

[14] J. M. Valls, Redes de neuronas. Perceptrón multicapa. Universidad Carlos III de Madrid, 2007.

[15] M. Hagan, H. Demuth, and M. Beale, Neural Network Design. Electrical Engineering Series, Pws Pub., 1996.

[16] P. Viola and M. J. Jones, Robust real-time face detection, Int. J. Comput.Vision, vol. 57, pp. 137-154, May 2004.