PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

# 3D fast wavelet network model-assisted 3D face recognition

Salwa Said, Olfa Jemai, Mourad Zaied, Chokri Ben Amar

Salwa Said, Olfa Jemai, Mourad Zaied, Chokri Ben Amar, "3D fast wavelet network model-assisted 3D face recognition," Proc. SPIE 9875, Eighth International Conference on Machine Vision (ICMV 2015), 98750E (8 December 2015); doi: 10.1117/12.2228368

Event: Eighth International Conference on Machine Vision, 2015, Barcelona, Spain

# 3D Fast Wavelet Network Model-Assisted 3D Face Recognition

Salwa Said, Olfa Jemai, Mourad Zaied and Chokri Ben Amar
REsearch Groups in Intelligent Machines (REGIM-Lab)
University of Sfax, National Engineering School of Sfax BP 1173, 3038 Sfax, Tunisia
salwa.said@ieee.org, olfa.jemai@ieee.org, mourad.zaied@ieee.org, Chokri.benamar@ieee.org

## ABSTRACT

In last years, the emergence of 3D shape in face recognition is due to its robustness to pose and illumination changes. These attractive benefits are not all the challenges to achieve satisfactory recognition rate. Other challenges such as facial expressions and computing time of matching algorithms remain to be explored. In this context, we propose our 3D face recognition approach using 3D wavelet networks. Our approach contains two stages: learning stage and recognition stage. For the training we propose a novel algorithm based on 3D fast wavelet transform. From 3D coordinates of the face (x,y,z), we proceed to voxelization to get a 3D volume which will be decomposed by 3D fast wavelet transform and modeled after that with a wavelet network, then their associated weights are considered as vector features to represent each training face . For the recognition stage, an unknown identity face is projected on all the training WN to obtain a new vector features after every projection. A similarity score is computed between the old and the obtained vector features. To show the efficiency of our approach, experimental results were performed on all the FRGC v.2 benchmark.

**Keywords:** 3D face recognition, Fast wavelet transform, Wavelet network.

## 1. INTRODUCTION

During the last decade, the high security has become a major problem for government institutions (military, nuclear, airport, ...). The identification by face recognition provides a number of solutions to this problem. In addition, face recognition is very important in all Human Machine Interface and remote monitoring systems. For security systems, recognition can locate a possible intruder automatically without that a person appears in front of a sensor. Despite advancements made in recent years, robust 2D recognition techniques to lighting conditions, posture changes or occlusion, are far from being developed. With the emergence and development of 3D acquisition techniques, 3D face recognition is a promising alternative to overcome these problems. The major advantage of these approaches is that the 3D model preserves all information about the geometry of the face, which provides a true representation of that. 3D face recognition techniques are divided into two classes: Surface based approaches and Model based approaches. Surface based approaches use the geometry of the face surface. These approaches are divided into global and local methods: global methods use the whole face in the recognition process (generally the pixel information is used) and local methods subdivide the face into regions (generally the eyes, nose and mouth regions) or also extract the feature points on the face. For local methods, Suikerbuik [1] proposed to use the Gaussian curvatures to find five characteristic points on the 3D model. Sometimes locate feature points with a maximum error of 4mm. In [2], the authors propose an hybrid technique using local geometry with Gaussian-Hermite moments and global surface with 3D mesh. These data will be represented by a single vector using PCA approach. In [3] the authors propose a 3D face recognition system using two different acquisition devices of 3D data: a 3D laser scanner and a structured light. Eight feature points in the face (geometrically invariant) are extracted and then used to calculate a feature vector containing the distances and angles between these points. In the recognition step, the authors implemented two different algorithms: the first based depth (depth-based DP) and the second used the SVM (feature-based SVM). Global methods use the whole face as input for the recognition system, so we have to align two 3D surfaces representing the two faces to be matched. The algorithm used in the alignment is generally ICP (Iterative Closest Point), which its objective is to optimize the matching and transformations associated. An approach based on ICP algorithm is presented by Cook et al [4]. The ICP algorithm is used to make the correspondence between 3D surfaces in order to correct errors due to non-rigid type of faces. Indeed, the faces are compared using a statistical model, which is the Gaussian Mixture Models (GMM). Model based approaches build face models from 3D points, that they use later for recognition. Blanz et al. [5] have proposed a method based on a 3D morphable face model. 3D points of the generated face models are represented by their cylindrical coordinates defined relative to a vertical axis. For each reference face, coordinates and texture values of all vertices are grouped to form two

Eighth International Conference on Machine Vision (ICMV 2015), edited by Antanas Verikas, Petia Radeva, Dmitry Nikolaev, Proc. of SPIE Vol. 9875, 98750E · © 2015 SPIE
CCC code: 0277-786X/15/$18 · doi: 10.1117/12.2228368

vectors: a shape vector and texture vector. After creating a generic model, the next step is to adjust the 2D image from shape and texture parameters. The image synthesis allows to make the new positions projected vertices of the 3D model using the lighting and color extracted. After creating a generic model, the next step is to adjust this model with the 2D image using shape and texture parameters. The image synthesis allows to make the new positions projected vertices of the 3D model using the lighting and color extracted. Finally, recognition step is performed by measuring the Mahalanobis distance between shape and texture parameters of the gallery models and the generic model. This approach was evaluated on two free databases: CMU-PIE [6] and FERET [7]. Recognition rate of 95% and 95.9% were obtained respectively over the CMU-PIE and FERET data. In [8] an automatic method for locating feature points on the 3D model using a stereo camera. These points are located on 2D images by Active Shape Model(ASM), and then they are transformed to 3D model by the stereo sensor algorithm. A similar approach was proposed by Ansari and Abdul Mottaleb [9] based on the stereo image and CANDIDE-3 model [10]. Feature points extraction is made (around the eyes, nose and mouth) on 2D images and then they are converted as 3D feature points. Finally, the 3D model is obtained by transforming the CANDIDE-3 generic face to match the extracted points. In this paper we propose a 3D face recognition approach based on 3D fast wavelet transform. The rest of the paper is organized as follows: section 2 outlines the proposed approach for 3D face recognition. Section 3 presents our experimental results and section 4 closes with a conclusion and discussion on possible enhancements.

## 2. OVERVIEW OF THE PROPOSED APPROACH

We propose in this paper a complete solution of 3D face recognition using 3D wavelet networks. The choice of the wavelet networks was based on its promoting results in the domains of image classification [11]–[15], speech recognition [16]–[18], images copies detection [19], content based image recognition [20] [21], hand gestures recognition [22]–[25], 2D face recognition [26]–[28], drowsy driver detection [29], [30].

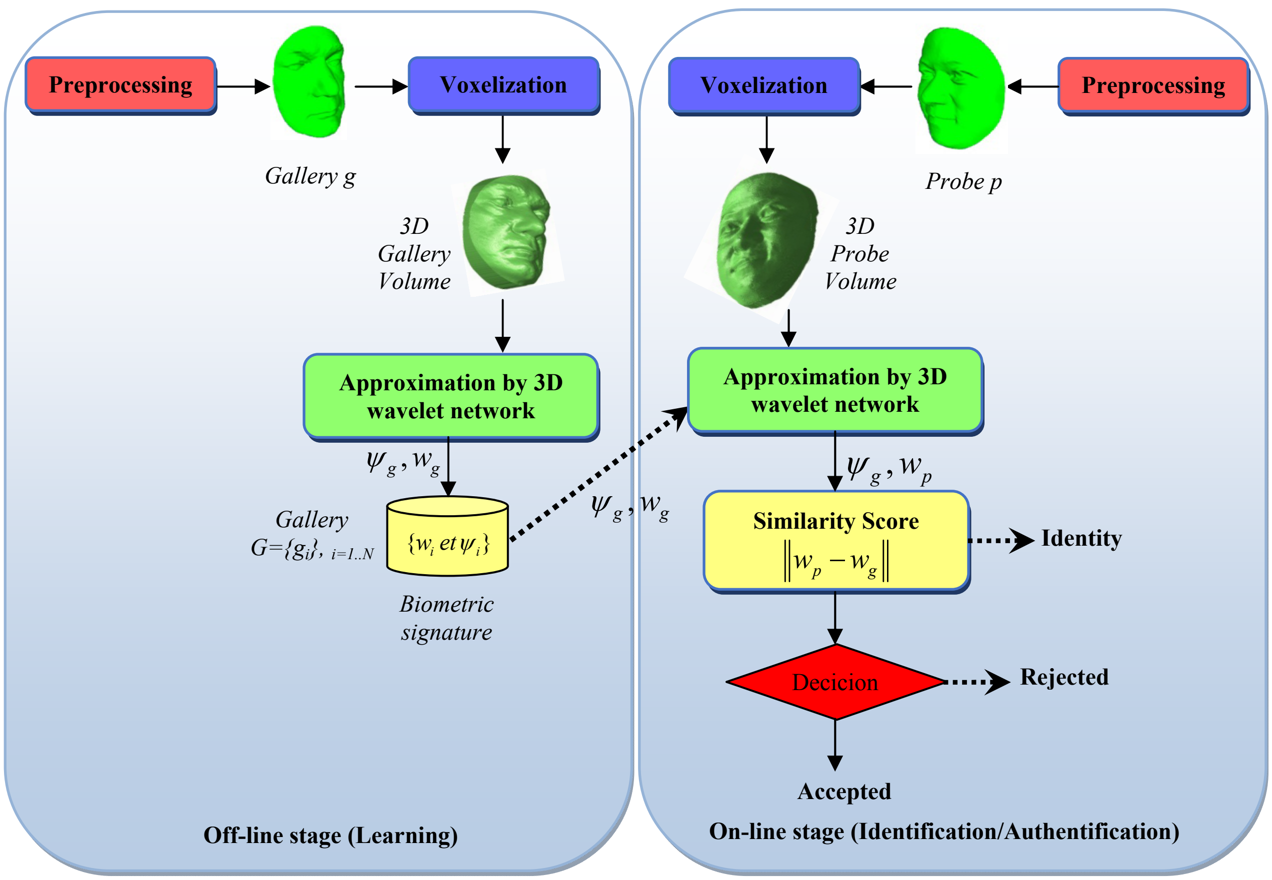


Figure 1. Overview of the proposed approach

Face recognition approach takes place mainly in two stages: learning or training stage (off-line) and recognition stage (online). The learning phase, aims to collect biometric signature on individuals to identify, and then save them in a learning database. Recognition phase, depending on the context of the application, it can operate in verification mode or in identification mode. For our approach, after acquisition, every 3D scan of the face will be preprocessed to extract only the useful part of the face. Then, every 3D facial mask will be voxelized to get a 3D volume of the face. This step is common for the gallery images and the probe ones. This volume is decomposed using 3D fast wavelet transform (3DFWT), biometric signatures (wavelets and their weights) are then stored in a database. The similarity score is then calculated between weights of gallery image wg and weights of probe image wp (Figure 1).

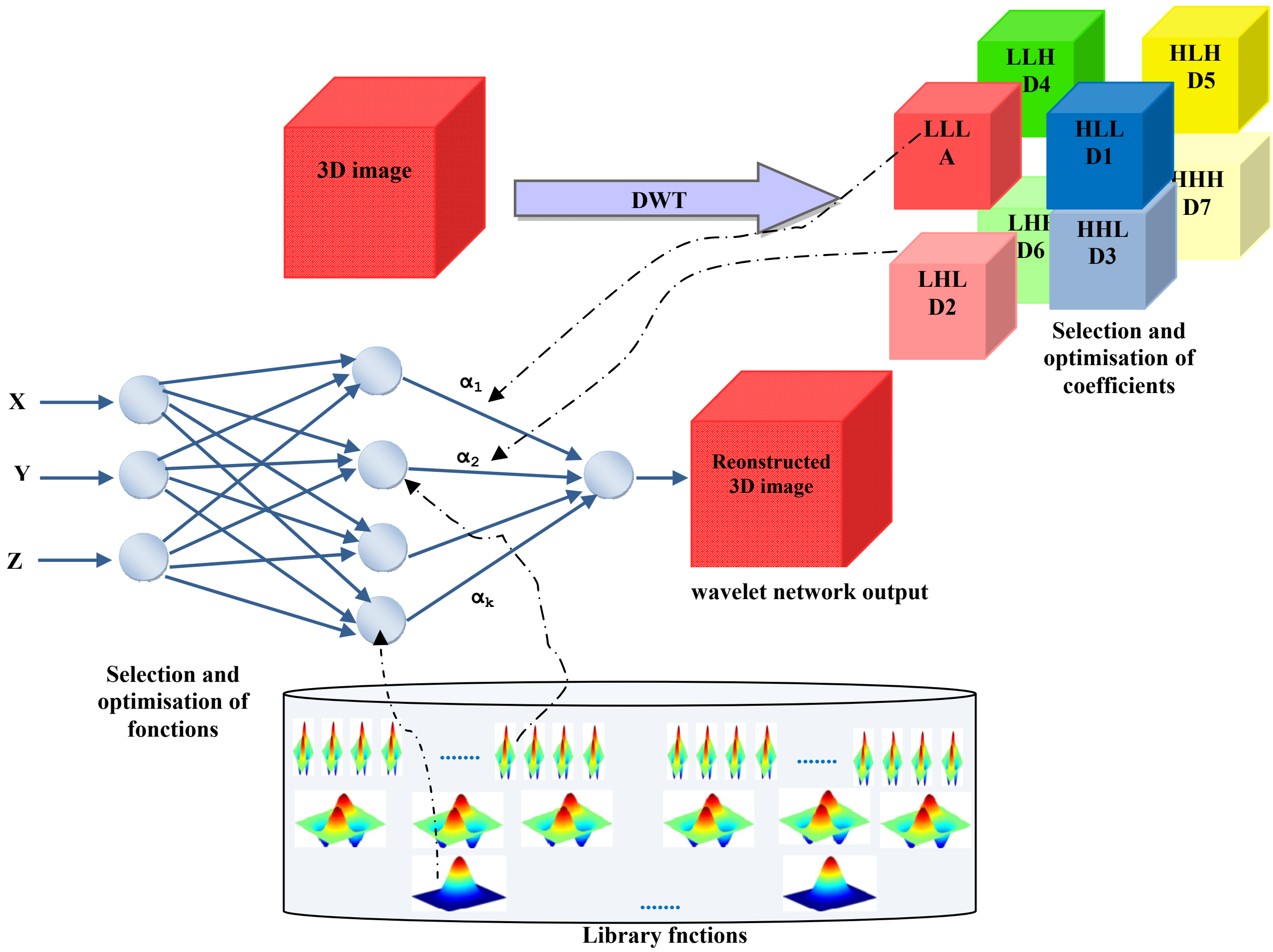


Figure 2. 3D learning process

## 2.1 Learning stage using 3DFWT

The main objective of this stage is to approximate the gallery images in order to get a biometric signature composed by wavelet parameters and their weights. We start this stage by the construction of 3D candidates wavelets of the library to be used in the network. Then, each 3D volume will be decomposed by 3DFWT, in order to get approximation and details coefficients. After that we compute the contribution of each wavelet, and add to the network the function that contributes more in the reconstruction of the image. An error, between the original image and the reconstructed one, will be defined as a stop learning condition. we present in the next section our proposed learning algorithm (figure 2).

### 2.1.1 Proposed learning algorithm

**Step 1:** Begin the learning stage by preparing the library of candidate wavelet and scaling functions to be used as transfer function in the wavelet network, obtained from a dyadic sampling of a mother wavelet that cover the 3D image to approximate. We nominate those functions $g_l$ with l = 1...M x N x K, and (M,N,K) the size of 3D image

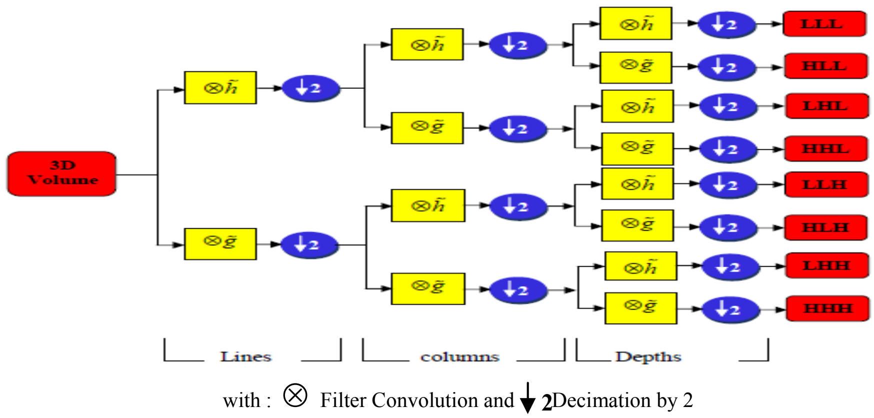


Figure 3. Fast decomposition of a 3D image using filter bank.

**Step 2:** Apply a series of 3DFWT to the volume v using dual set of filters $(\tilde{h}, \tilde{g})$ [31] (figure 3) to calculate the coefficients $\alpha_l$ corresponding to functions $g_l$ of the library. Set a volume r equal to the volume v, so r is defined as:

$$r = \sum_{l=1}^{M \times N \times K} \alpha_l g_l \quad (1)$$

**Step 3:** Compute the contribution $\alpha_l\, g_l$ of all activation functions of the library to reconstruct the volume r.

**Step 4:** Choose from the library the function $g_k$ (k the number of selected functions) that contributes more in the reconstruction of r, then optimize their parameters (dilatation, position and rotation) to better approximate r and improve the accuracy of the output network. This optimized function $g'_k$ does not belong anymore to the library, so we calculate the corresponding weight $\alpha'_k$ .

**Step 5:** The function $g'_k$ will be added to the hidden layer of the network that approximate v and set its corresponding weight to $\alpha'_k$ . The output of the network is $\tilde{v} = \tilde{v} + \alpha'_k g'_k$.

**Step 6:** Optimize the wavelet network architecture: the function $g'_k$ not necessarily formed a basis with functions already added to the hidden layer. In this case $g'_k$ formes a frame with those functions, a process to reduce the hidden nodes number and to update the weights is introduced [28].

**Step 7:** Calculate the residual volume r $= v - \tilde{v}$ and return the step 4 if the stop learning condition ( an error E, between the volume v and the output of the network) is not reached.

## 2.2 Recognition

For recognition stage, after preprocessing and voxelization of the probe image, the resulting volume is decomposed using 3DFWT. To compute the similarity score, we takes each ($\boldsymbol{\alpha}_g$,$w_g$) stored in the learning database, find the positions of the wavelets used and extract the corresponding weights for training image $w_g$ and test image $w_p$, then we compare the two weights vectors. Finally, the gallery image that gives the minimum distance between these vectors, corresponds to the search person.

# 3. EXPERIMENTAL RESULTS

In this section, we present the results given by our 3D face recognition approach using wavelet networks. Experiments was done on all the base FRGC v.2., it contains 4007 neutral and non neutral faces of 466 different persons.

Table 1: Description of the FRGC V.2. Base.

| | |
|---|---|
| Total Images number | 4007 |
| Classes number | 466 |
| Images number per class | 1 to 21 |

To better evaluate the performance of our 3D face recognition approach, we have divided the database into 466 classes corresponding to 466 peoples in the FRGC base. Each class contains 1 up to 21 images, with one neutral image and the others with facial expressions. Description of the FRGC base is given in table 1. In our experiments we found that several criterions can influence the performance of our approach:

### 3.1 Base distribution

For our experiments, we used the following protocol: for each class we have taken a subset of images to the learning stage and the rest for recognition stage. At first we took the half of images for learning and the other half for testing, our approach gives a recognition rate around 81%. As a second test, we chose 2/3 of images for the learning stage and 1/3 for the test stage and we noticed that the recognition rate increased to 89%, while keeping the same resolution for 3D images and the same classification criterion using k = 1 for the k-nearest neighbors (1-KNN).

Table 2: Evaluation According to base Distribution

| Training Subset | Test Subset | Recognition Rate |
|---|---|---|
| 1/2 | 1/2 | 81% |
| 2/3 | 1/3 | 89% |

Table 2 summarizes the rates obtained by our approach by adjusting the values of this criterion. We notice that in the second case with (2/3,1/3) the performance is better and this is because we have in the training set more samples to learn with different facial expressions, so more chance to recognize a test face.

### 3.2 Classification with KNN

The k nearest neighbors was chosen as a second criterion in the recognition stage. First, a similarity score is computed between the probe image and all the gallery images as shown in figure 1. Then, sort these distances in decreasing order and finally, look for k minimum distances and see their classes. Case k = 1, our approach gives 81% as recognition rate. If we takek = 3 our approach reach a rate of 93%.

Table 3 : Evaluation According to the Value of K for Knn.

| k-NN | Recognition Rate |
|---|---|
| 1 | 81% |
| 2 | 90% |
| 3 | 93% |

Experimental results presented in table 3, were obtained using the same resolution for 3D images and (1/2,1/2) as base distribution criterion.

### 3.3 3D image resolution

The last criterion is the resolution of 3D images used in the learning and the recognition stages. The image quality depend on the number of voxels used.

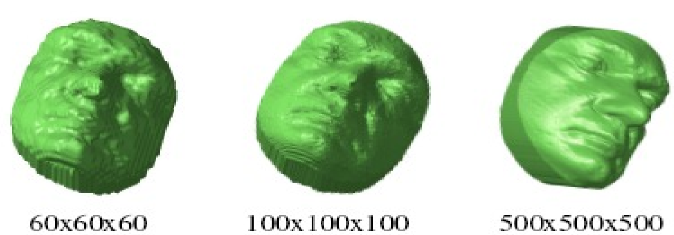


Figure 4. Examples of 3D faces by varying the resolution.

Figure 4 shows un example of 3D images obtained by varying the voxel number. We tested different resolutions and we conclude that our approach gives better performances if we increase the resolution. It reached a rate around 81% for a resolution about (60x60x60), and 73% for a resolution about (40x40x40), using the same experimental protocol (1-NN for classification and (1/2,1/2) base distribution). Table 4 shows the results obtained for this criterion.

Table 4: Evaluation According to 3D Image Resolution.

| 3D resolution | Recognition rate |
|---|---|
| 40x40x40 | 73% |
| 60x60x60 | 81% |

## 4. CONCLUSION

In this paper we have presented a new 3D face recognition approach based on 3D wavelet and 3D fast wavelet transform (3DFWT). After preprocessing, every gallery or probe face requires a voxelization step to get a 3D volume that will be then approximated by 3D wavelet networks. For the recognition stage, a similarity score is calculated between the two weight vectors of the gallery and the test face. We tested this approach on all the FRGC database, and we got recognition rates from 82% up to 93% according to precise values for the evaluation criterion selected in our experiments.

## ACKNOWLEDGMENT

The authors would like to acknowledge the financial support of this work by grants from General Direction of Scientific Research (DGRST), Tunisia, under the ARUB program.

## REFERENCES

[1] C.A.M. Suikerbuik, J.W.H. Tangelder, H.A.M. Daanen, A.J.K. Oudenhuijzen. Automatic feature detection in 3D human body scans. The conference SAE Digital Human Modelling for Design and Engineering, (2004).

[2] C. Xu, Y. Wang, T. Tan, L. Quan. Automatic 3D face recognition combining global geometric features with local shape variation information. International Conference on Automated Face and Gesture Recognition,308-313, (2004).

[3] A. S. Gawali and R. R Deshmukh. 3D Face Recognition Using Geodesic Facial Curves to Handle Expression, Occlusion and Pose Variations. International Journal of Computer Science and Information Technologies, Vol. 5 (3), 4284-4287, (2014).

[4] J.Cook, V.Chandran, S.Sridharan, C. Fookes. Face recognition from 3D data using iterative closest point algorithm and Gaussian mixture models. The 2nd International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT), Thessaloniki, Greece, 6-9 September, (2004).

[5] V. Blanz, T. Vetter. Face recognition based on fitting a 3D morphable model. IEEE Transactions on Pattern Analysis and Machine Intelligence vol 25, 1063-1074, (2003).

[6] T. Sim, S. Baker, M. Bsat. The CMU pose illumination and expression (PIE) database. 5th International Conference on Automatic Face and Gesture Recognition, pp.5358, Washington, D.C., USA, May, (2002).

[7] P. J. Phillips, H. Moon, S. A. Rizvi, P. J. Rauss. The FERET Evaluation Methodology for Face-Recognition Algorithms. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol 22(10), 1090-1104, Octobre (2000).

[8] S. Jaiswal, S. S. Bhadauria, R. S. Jadon. 3D face recognition and modeling system. Journal of global research in computer science, Vol. 2, No 7, 30-37, July (2011).

[9] A. Ansari, M. Abdel-Mottaleb. 3D Face modeling using two views and a generic face model with application to 3D face recognition. IEEE Conference on Advanced Video and Signal Based Surveillance, 37-44, (2003).

[10] W. Zeng, M. Yang, and Z. Cui. Ultra-Low Bit Rate Facial Coding Hybrid Model . Journal of Image and Graphics, Vol. 3, No. 1, 25-29, June (2015).

[11] O. Jemai, M. Zaied, C. Ben Amar, A. M. Alimi. FBWN: an architecture of Fast Beta Wavelet Networks for Image Classification, IEEE World Congress on Computational Intelligence (IEEE WCCI 2010), the International Joint Conference on Neural Networks (IJCNN 2010), 1953-1960, July, 18-23, CCIB, Barcelona, Spain, (2010).

[12] O. Jemai, M. Zaied, C. Ben Amar and A. M. Alimi, Pyramidal Hybrid Approach : Wavelet Network with OLS Algorithm Based-Image Classification, IJWMIP: International Journal of Wavelets, Multiresolution and Information Processing, Vol.9(1)(World scientific), 111-130, (2011).

[13] O. Jemai, M. Zaied, C. Ben Amar and A.M Alimi, Fast Learning algorithm of wavelet network based on Fast Wavelet Transform, Int. J. Pattern Recognition and Artificial Intelligence, Vol. 25(8), 1297-1319,( 2011).

[14] T. Bouchrika, O. Jemai, M. Zaied, C. Ben Amar, A New Supervised Image Classifier Architecture Based on Multiresolution Wavelet Network Including a Fuzzy Decision Support System, In Fifth International Conference on Information, Intelligence, Systems and Applications (IISA), 88 - 91, Juillet 07-09, Chania, Grece, (2014).

[15] O. Jemai,T. Bouchrika, M. Zaied and C. Ben Amar, Supervised Wavelet Network based Fuzzy-Logic Classifier Performance on the UCI Databases, In 6th International Conference on Soft Computing and Pattern Recognition (SOCPAR), 128-133, DOI:10.1109/SOCPAR.2014.7007993, August 11-13, Tunis, Tunisia, (2014).

[16] O. Jemai, R. Ejbeli, M. Zaied and C. Ben Amar, A speech recognition system based on hybrid wavelet network including a fuzzy decision support system, Proc. SPIE 9445, Seventh International Conference on Machine Vision (ICMV 2014), 944503 (February 12, 2015); doi:10.1117/12.2180554; http://dx.doi.org/10.1117/12.2180554.

[17] R. Ejbali, Y. Benayed, M. Zaied, A.M. Alimi, Wavelet Networks for phonemes Recognition, International Conference on Systems and Information Processing, Guelma-Algeria May 02-04 (2009).

[18] R. Ejbali, M. Zaied and C. Ben Amar, Wavelet network for recognition system of Arabic word, International Journal of Speech Technology, Vol.13(3),Springer-Verlag, New York, 163-174, (2010).

[19] A. El Adel, M. Zaied and C. Ben Amar., Learning wavelet networks based on Multiresolution analysis: Application to images copy detection, in 2011 Int. Conf. on Communications, Computing and Control Applications, Tunisia, March 3-5 (2011).

[20] B. Guedri, M. Zaied and C. Ben Amar, Indexing and images retrieval by content, IEEE 2011 International Conference on High Performance Computing & Simulation, Istanbul Turkey, pp. 369-375, July 4-8 (2011).

[21] S. Othman, O. Jemai, M. Zaied and C. Ben Amar, Medical image retrieval using hybrid wavelet network classifier, In 6th International Conference on Soft Computing and Pattern Recognition (SOCPAR) proceedings, 221-225, DOI: 10.1109/SOCPAR. 2014.7008009, August 11-13, 2014, Tunis, Tunisia, (2014).

[22] T. Bouchrika, M. Zaied, O. Jemai and C. Ben Amar, Ordering computers by hand gestures recognition based on wavelet networks, 2nd International Conference on Communications, Computing and Control Applications proceedings, Décembre, 06-08, Marseilles, France, DOI: 10.1109/CCCA.2012.6417911, pp. 1-6, (2012).

[23] T. Bouchrika, M. Zaied, O. Jemai and C. Ben Amar, Neural solutions to interact with computers by hand gesture recognition, MTAP: Multimedia Tools and Applications, Vol.72(3), 2949-2975, (2014).

[24] T. Bouchrika, O. Jemai, M. Zaied, C. Ben Amar, A new hand posture recognizer based on hybrid wavelet network including a fuzzy decision support system, 15th International Conference on Intelligent Data Engineering and Automated Learning (IDEAL), Vol. 8669, 183-190, Salamanca, Espagne, Septembre. (2014).

[25] T. Bouchrika, O. Jemai, M. Zaied, C. Ben Amar, Cascaded Hybrid Wavelet Network for Hand Gestures Recognition, In IEEE International Conference on Systems, Man, and Cybernetics (SMC), DOI: 10.1109/SMC.2014.6974104, 1360-1365, Octobre 05-08, San Diego, Californie, USA, (2014).

[26] M. Zaied , S. Said , O. Jemai and C. Ben Amar, A novel approach for face recognition based on fast learning algorithm and wavelet network theory, IJWMIP: International Journal of Wavelets, Multiresolution and Information Processing, Vol. 9(6),(World scientific), 923-945, (2011).

[27] M. Zaied, O. Jemai, C. Ben Amar. Training of the Beta wavelet networks by the frames theory: Application to face recognition. International Workshops on Image Processing Theory, Tools and Applications, IPTA, 165-170, Tunisia November (2008).

[28] S. Said, B. Ben Amor, M. Zaied, C. Ben Amar and M. Daoudi, Fast and efficient 3D face recognition using wavelet networks, 16th IEEE International Conference on Image Processing (ICIP). 4153-4156, Egypt, November (2009).

[29] O. Jemai, I. Teyeb, T. Bouchrika and C. Ben Amar, A Novel Approach for Drowsy Driver Detection Using Eyes Recognition System Based on Wavelet Network, IJES: International Journal of Recent Contributions from Engineering, Science & IT, 1 (1), 46-52, (2013).

[30] I. Teyeb, O. Jemai, M. Zaied and C. Ben Amar, A Novel Approach for Drowsy Driver Detection Using Head Posture Estimation and Eyes Recognition System Based on Wavelet Network, In The Fifth International Conference on Information, Intelligence, Systems and Applications (IISA), 379 - 384, July 07-09, Chania, Greece, (2014).

[31] M. Zaied, C. Ben Amar, A.M. Alimi. Beta Wavelet Networks for Face Recognition. Journal of Decision Systems, Lavoisier 2005 Edition, Vol. 14, 109-122, (2005).