The Method of User's Identification Using the Fusion of Wavelet Transform and Hidden Markov Models

Janusz Bobulski

Częstochowa University of Technology, Institute of Computer and Information Sciences,
Dąbrowskiego 73, 42–200 Częstochowa, Poland,
janusz.bobulski@icis.pcz.pl

The work concerns creation of the original *FaMar* method of user's identification on the basis of the frontal facial image, in which the fusion of Wavelet Transformation (WT) and Hidden Markov Models (HMM) are used for the three parts of face (eyes, nose, mouth); the decision is made on the basis of the sum maximalisation of likelihood of generating of the models observation.

1. Introduction

A problem of persons' identification is one of the main questions of many research centers at present. Interest of this discipline is a result of potential possibilities of practical application of new possibilities in persons' identification in the systems demanding authorizations of persons' access entitled to use potential resources. The most popular method of face identification is Principal Component Analysis (PCA) [1, 2]. Other popular methods are using Wavelet Transform [4] or Hidden Markov Models [6]. Analysis of the existing solutions revealed their defects, which caused their weak effectiveness. The disadvantages of these methods are as follows:

- In case of the new user's registration, process of learning and addition his/her facial image to a database, require repeated learning of the whole system.
- The identification error rate is 3-6%.
- They work with whole face.
- They are computationally very expensive.

The work concerns creation of the original *FaMar* method of user's identification on the basis of the frontal facial image, in which the fusion of the Wavelet Transformation (WT) and Hidden Markov Models (HMM) are used for three parts of face (eyes, nose, mouth); the decision is made on the basis of the sum maximalisation of likelihood of generating of the models observation [7].

2. The Proposed Method

The proposed method is a combination of two mathematical tools, Wavelet Transform (WT) and Hidden Markov Model (HMM). Both were mainly used for speech recognition. Here, WT is used for features extraction, and HMM for identification. This system works in two modes, learning and testing. These modes differ from each other. The algorithm of this method consists of four main parts:

1. Pre-processing: normalization and face division into three parts
2. Features extraction: WT of the face image
3. Training: generating and learning HMM for each part of the face
   Testing: testing models from the database
4. Training: saving to database the learned models of the face
   Testing: making a decision - maximum likelihood of the model

2.1. Pre-processing

The normalization consists of fixing the centers of the eyes, and then respective scaling of face so that the distance between them equals 60 pixels. The second part of this process is division of the normalized face into three parts: the area of eyes, nose, and mouth (Fig.1).

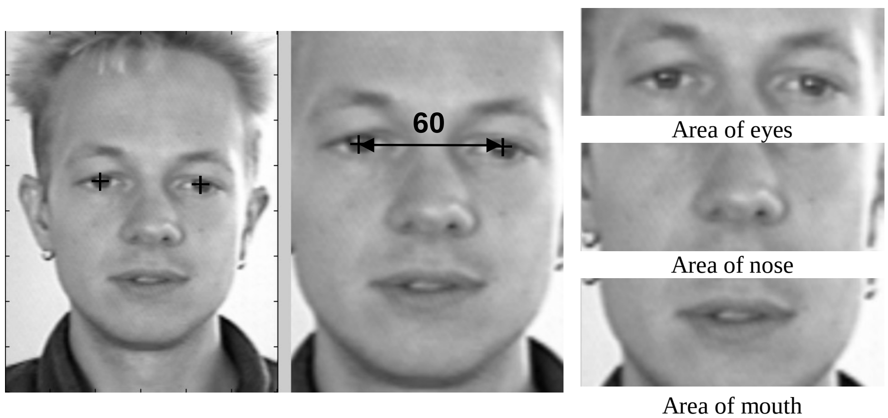

Fig. 1. Pre-processing of the face image

2.2. Features extraction

WT is used for features extraction [5]. Using 2D WT (Fig. 2) [11], the face image is decomposed into four sub images via the high-pass and low-pass filtering. The image is decomposed along column direction into sub images to high-pass frequency band H and low-pass frequency band L. Assuming that the input image is a matrix of m x n pixels, the resulting sub images become $m/2$ x n matrices. At the second step the images H and L are decomposed along row vector direction and respectively produce the high and low frequency band HH and HL for H, and LH and LL for L. The four output images become the matrices of $m/2$ x $n/2$ pixels. Low frequency sub image LL (A_1) possesses high energy, and is a smaller copy of the original images (A_0). The remaining sub images LH, HL, and HH respectively extract the changing components in horizontal (D_{11}), vertical (D_{12}), and diagonal (D_{13}) direction [3].

Wavelet Transform of the second level (Fig. 3) is used for features extraction in the propose technique. After first level wavelet decomposition, the output images become the input images of the second level decomposition. The results of two-level 2D WT are coded in this way, so that they can be applied in HMM (Fig. 5). One of the simplest methods of reduction and information coding is calculating of standard deviation or mean value. Each part of the face is transformed separately by discrete wavelet transform (Fig. 4). The bank filters' selection is an important thing in this transformation. It guarantees a good recognition rate. More information about it can be found in [8].

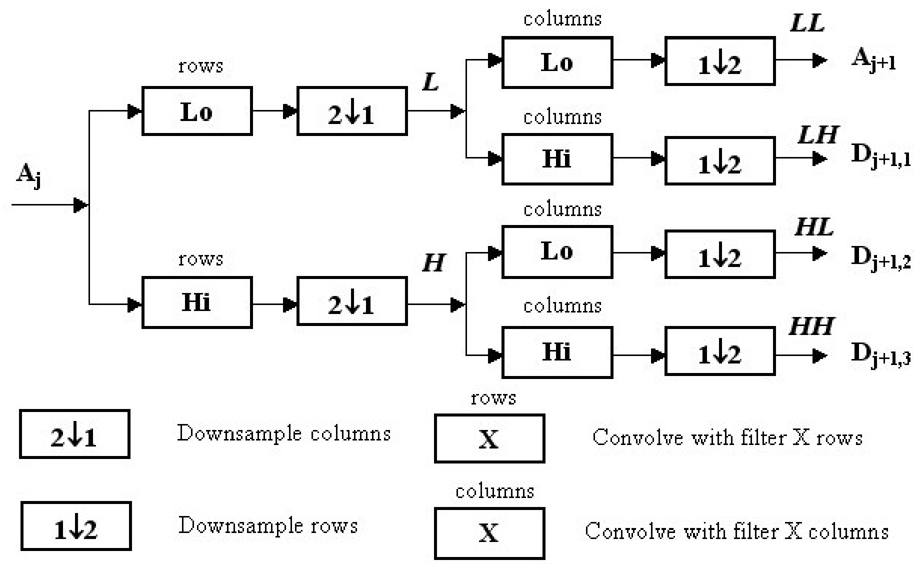

Fig. 2. One-level two-dimensional wavelet transform

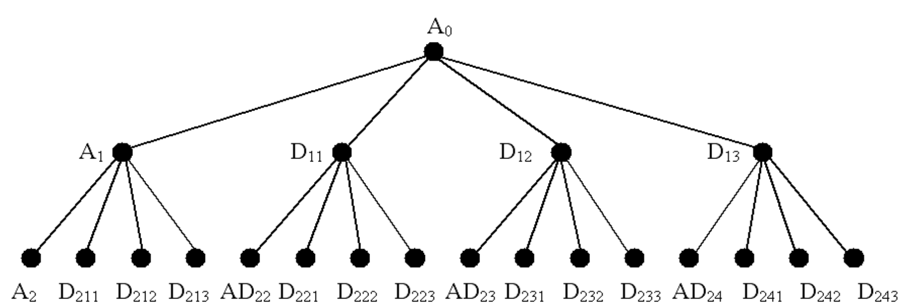

Fig. 3. The wavelet decomposition tree

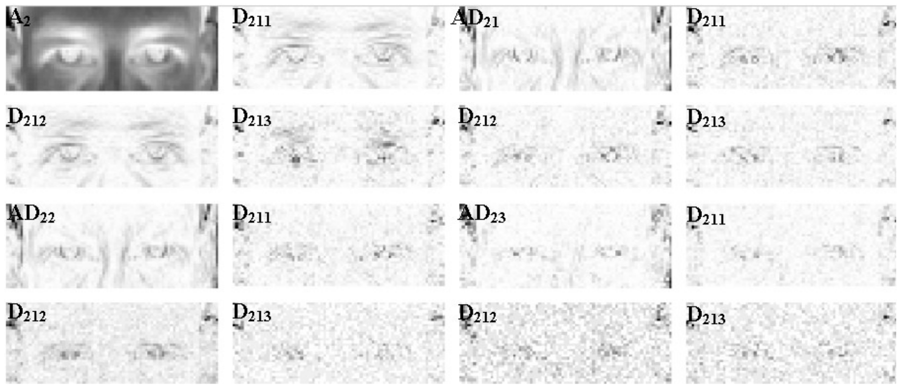

Fig. 4. Example of level 2 of the wavelet decomposition of image of eyes area

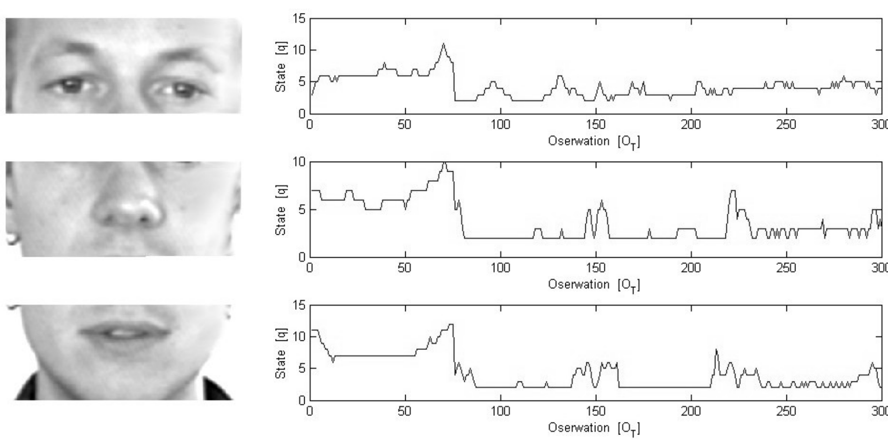

Fig. 5. of face and correspond them sequences of observation

2.3. Training

HMM is used for the identification process. HMM is a double stochastic process with underlying stochastic process that is not observable (hidden), but can be observed through another set of stochastic processes that produce a sequence of observation. Let O=$\{O1,,OT\}$ be the sequence of observation of feature vectors, where T is the total number of feature vectors in the sequence. The statistical parameters of the model may be defined as follows [9].

- The number of states of the model, N
- The transition probabilities of the underlying Markov chain, A=$\{aij\}$ $1 \leq i,j \leq N$ where a_{ij} is the probability of transition from state i to state j subject to the constraint $\sum_{j=1}^{N} a_{ij} = 1$
- The observation probabilities, $B = \{b_j(O_T)\}$, $1 \leq j \leq N$, $1 \leq t \leq T$ which represents the probability of the t^{th} observation conditioned on the j^{th} state.
- The initial probability vector, $\prod = \{\pi_i\} 1 \leq i \leq N$.

Hence, the HMM requires three probability measures to be defined, A, B, π and the notation:

$$\lambda = (A, B, \pi) \tag{1}$$

is often used to indicate the set of parameters of the model.

In the proposed method, one model is made for each part of the face. The parameters of the model are generated at random at the beginning. Then they are estimated with Baum-Welch algorithm, which is based on the forward-backward algorithm. The forward algorithm calculates the coefficient $\alpha_t(i)$ (probability of observing the partial sequence $(o_1, , o_t)$ such that state q_t is i). The backward algorithm calculates the coefficient $\beta_t(i)$ (probability of observing the partial sequence $(o_{t+1}, , o_T)$ such that state q_t is i). The Baum-Welch algorithm, which computes the λ, can be described as follows [10].

1. Let initial model be λ_0
2. Compute new λ based on λ_0 and observation O
3. If $log(P(O|\lambda) - log(P(O)|\lambda)_0 < DELTA$ stop
4. Else set $\lambda_0 \leftarrow \lambda$ and go to step 2.

The parameters of new model $\lambda(1)$, based on λ_0 and observation O, are estimated from equation of Baum-Welch algorithm [10], and then are recorded to the database.

2.4. Testing

The testing process consists of computing the probability of observation generating by the models saved in database and choosing this model for which the likelihood is maximum. In the proposed method, probabilities are calculated separately for each of the three models representing parts of the face, and then they are added. The face, for which the sum of probability is maximum, is chosen as the correct face. The probability of generating sequences of observations is computed from the equations 2-4 [10].

$$P(O|\lambda) = \sum_{q} P(O|q,\lambda)P(q|\lambda) \tag{2}$$

$$P(O|q,\lambda) = \prod_{i=1}^{T} P(o_t|q_t,\lambda) = b_{q_1}(o_1)b_{q_2}(o_2)\dots b_{q_T}(o_T) \tag{3}$$

$$P(q|\lambda) = \pi_{q_1} a_{q_1 q_2} a_{q_2 q_3} \dots a_{q_{T-1} q_T} \tag{4}$$

3. Experimenting

The following constrains are imposed to guarantee reliability of face identification: a frontal face view is located in the centre of an input image; eyes are open and mouth is closed; the face should not be covered with shadow, the rotate angle of an input image must be less than ten degree. The results of three experiments are presented here. The first one was carried out on the basis of the *BioID* [12] face database, in which there are 24 subjects, one image for one person. The second and the third experiment were carried out using my own face database *FaDab* [13] in which there are 150 subjects, one image per one person. The first two experiments were carried out on the basis of the three parts of a face, whilst the third one only on the area of the eyes, which means that the face was represented by one model. The results of experiments are shown in Tab. 1.

Table 1. The results of experiment

Name of face database	Number of face parts [pieces]	Total images in database [pieces]	Number of Misclassified [pieces]	Error rate [%]
BioID	3	24	1	4,16
FaDab	3	150	5	3,33
FaDab	1 (eyes)	150	3	2,00

The second part of experiment compare *FaMar* method with the other face identification methods: PCA, HMM and Wavelet-face.

4. Conclusion

On the basis of experimental research it was stated the area of eyes contains the most useful information for the persons' identification, and it could be successfully applied in specific methods of identification (e.g. detection). The method is characterized by following novelties:

1. The usage of the three areas of the face for identification and creating for each of them one independent HMM (which it is possible to use separately or together). This procedure gives possibility to short calculation request and permit obtaining a recognition rate as good as in modern method.

Table 2. The comparison of error rate *FaMar* method with other methods

Name of method	Name of face database	Total images in database [pieces]	Number of Misclassified [pieces]	Error rate [%]
PCA	BioID	24	2	8,34
HMM	BioID	24	4	16,67
Wavlet-face	BioID	24	2	8,34
FaMar	BioID	24	1	4,17
PCA	BioID	150	13	8,66
HMM	BioID	150	24	16,00
Wavelet-face	BioID	150	9	6,00
FaMar	BioID	150	5	3,33

2. The transition from 2D pictures to 1D-WT of the facial areas. This procedure permits to obtain the recognition rate as good as in modern method and gives possibility to short calculation request too.
3. The fusion of WT and HMM (see p.1 and p.2) with using the assumption of maximalization of the likelihood's sum of generating of the observation.

The advantages of the method:

1. In case of the new user's registration, the process of learning and addition his/her facial image to a database, does not require repeated learning of the whole system. After addition a new facial image, any traditional recognition system must be learned again.
2. Using the *BioID* (Germany) and *FaDab* (made by the author) bases on facial images, the *FaMar* system level of identification errors equals about 3,5 % in contrary of 3-6 % for the reviewed modern recognition systems.
3. The user authentication time equals about 0.3 second (the three mentioned HMMs). Accordingly, the user authentication time equals about 0.1 second in case of a usage any area of the facial image.

5. Acknowledgments

I would like to thank professor Leonid Kompanets for valuable advises and knowledge, which he shares with me. This work is a part of the grant sponsored by KBN, No 4T11C00425

References

[1] Kirby M., Sirovich L.: Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces, IEEE Transaction on Pattern Analysis and Machine Intelligence, 12(1) (1990).

[2] Kukharev G., Kuzminski A.: Techniki Biometryczne Cz 1 - Metody Rozpoznawania Twarzy, Politechnika Szczeciska Wydzia Informatyki (2003).

[3] Garcia C., Zikos G., Tziritas G.: Wavelet pocket analysis for face recognition, Image and Vision Computing 18 (2000) 289-297.
[4] Chien J.T. Wu Ch.Ch.: Discriminant Waveletface and Nearest Feature Classifiers for Face Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 24(12) (2002) 1644-1649.
[5] Bobulski J.: Ekstrakcja cech twarzy za pomoc transformaty falkowej, Informatyka Teoretyczna i Stosowana 3(4) (2003) 127-134.
[6] Samaria F, Young S.: HMM-based Architecture for Face Identification, Image and Vision Computing, Vol. 12 No 8 October (1994).
[7] Bobulski J.: The Face Recognition Technique Using Wavelet-face, Hidden Markov Model, and Maximum Likelihood Principle, ACS'2003, Midzyzdroje (2003).
[8] Kozuani A.Z. He f. Sammut K.: Wavelet Packet Face Representation and Recognition, Proc. IEEE Conf. Systems, Man, and Cybernetics (1997) 1614-1619.
[9] Rabiner L. R.: A tutorial on hidden Markov models and selected application in speech recognition, Proc. IEEE 77 (1989) 257-285.
[10] Kanungo T.: Hidden Markov Model Tutorial, www.cfar.umd.edu/ kanungo.
[11] Wavelet Toolbox User's Guide, MatLab 5.0, The MathWorks.
[12] Face Database BioID, http://www.humanscan.com/bioid.htm.
[13] Face Database FaDab, http://icis.pcz.pl/ januszb/baza.htm.