HMM and WT fusion for face identification

Janusz Bobulski

Czestochowa University of Technology, Institute of Computernad Information Science, Dabrowskiego Street 73, 42-200 Czestochowa, Poland
januszb@icis.pcz.pl

This paper describes the original *FaMar* method of user's identification. The method bases on the fusion of Wavelet Transform (WT) and Hidden Markov Models (HMM), which is used for three parts of the face (eyes, nose, and mouth) separately.

Key words: Identification of human, face recognition, biometrics, Hidden Markov Model, Wavelet Transform

1 Introduction

A problem of persons' identification is a leading issue of many research centres. The interest of this domain results from the potential possibilities applying the new approach to person's identification in systems that require access authorizations to resources [1, 2]. Research on the face recognition systems has lasted for over twenty years. However, there is still no 100% effective method, which could be used to access authorizations. In recent years, considerable progress has been made on the problem of face detection and face recognition [3] and efficient algorithms have been created especially for stable conditions such as: small variations in lighting, facial expression and pose. These methods can be roughly divided into two different groups: geometrical features matching and template matching. In the first case, some geometrical measures about distinctive facial features such as eyes, mouth, nose and chin are extracted [4]. In the second case, the face image, represented as a two-dimensional array of intensity values, is compared to a single or several templates representing a whole face. The earliest methods for template matching are correlation-based, thus computationally very expensive and require a great amount of storage, and for a few years, the Principal Components Analysis (PCA) method is successfully used in order to perform dimensionality reduction [5, 6]. Other popular methods are using Wavelet Transform (WT) [7] or Hidden Markov Models (HMM) [8]. Analysis of the existing solution revealed their defects, which caused their weak effectiveness. The disadvantages of these methods are as follow:

- In case of the new user's registration, process of learning and addition his facial image to a database require repeated learning of whole system.
- They work with whole face image.
- 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 WT and HMM are used for three parts of face (eyes, nose and mouth); the decision is made on the basis of the sum maximalisation of likelihood of generating of models observation [9].

2 The proposed method

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

1. Pre-processing: normalization and face division on three parts.
2. Features extraction: WT of 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 face.
   Testing: making a decision - maximum likelihood of model.

Pre-processing

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

2.1 Features extraction

WT is used to features extraction. Using 2D WT (Fig. 2) [10], the face image is decomposed into four subimages via the high-pass and low-pass filtering. The image is decomposed along column direction into subimages 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 subimages become $m/2$ x n matrices. At second step the images H and L are decomposed along row vector direction and respectively produce the high and low frequency band

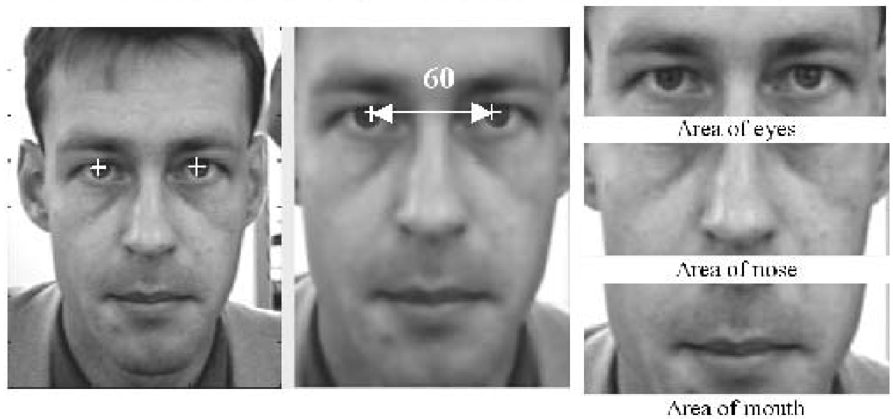

Fig. 1. Pre-processing of the face image

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 subimage *LL* (*A1*) possesses high energy, and is a smallest copy of original images (*A0*). The remaining subimages *LH*, *HL*, and *HH* respecticly extract the changing components in horizontal (*D11*), vertical (*D12*), and diagonal (*D13*) direction [7]. Wavelet Transform of second level (Fig. 3) is used to features extraction in propose technique. After first level wavelet decomposition, the output images become input images of 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 the 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 [11].

2.2 Training

HMM is used to the identification process. A 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 $O = \{O_1, \dots, O_T\}$ 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 [10].

- The number of states of the model, N
- The transition probabilities of the underlying Markov chain, $A = \{a_{ij}\} 1 \leq i, j \leq N$ where a_{ij} is the probability of transition from state i to state j subject to the constraint

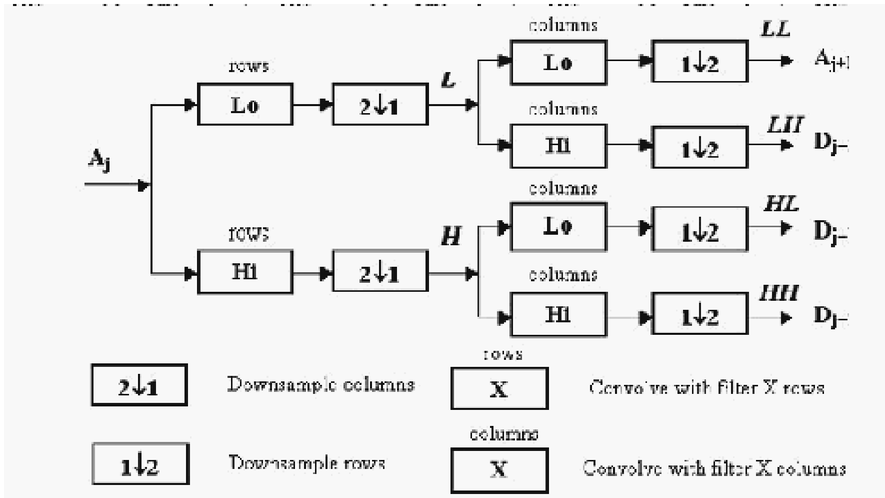

Fig. 2. One-level two-dimensional wavelet transform

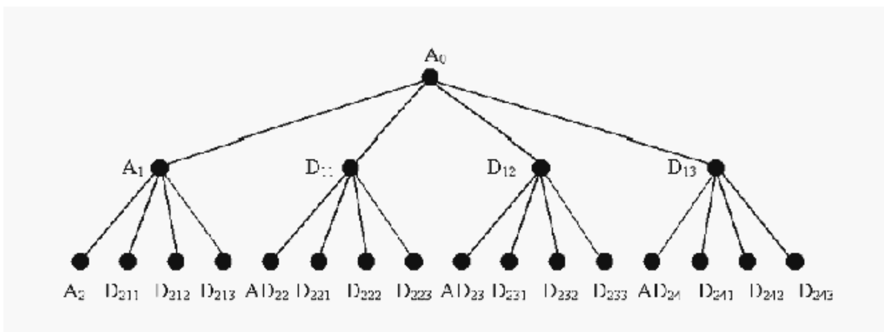

Fig. 3. The wavelet decomposition tree

- 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, $\Pi = \{\pi_i\}, 1 \leq i \leq N$.

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 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-

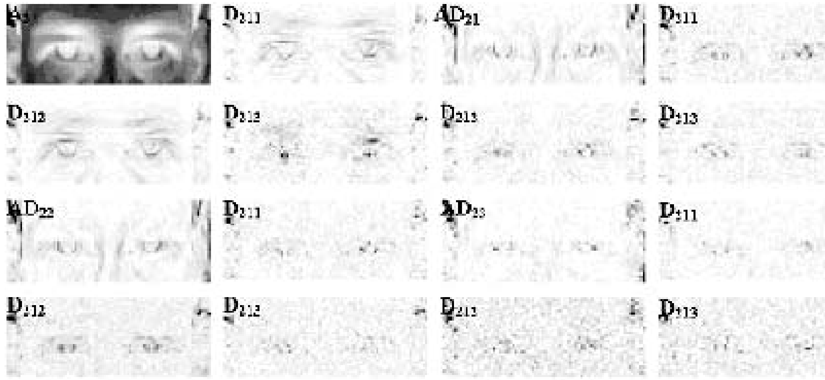

Fig. 4. Example of level 2 of the wavelet decomposition of image

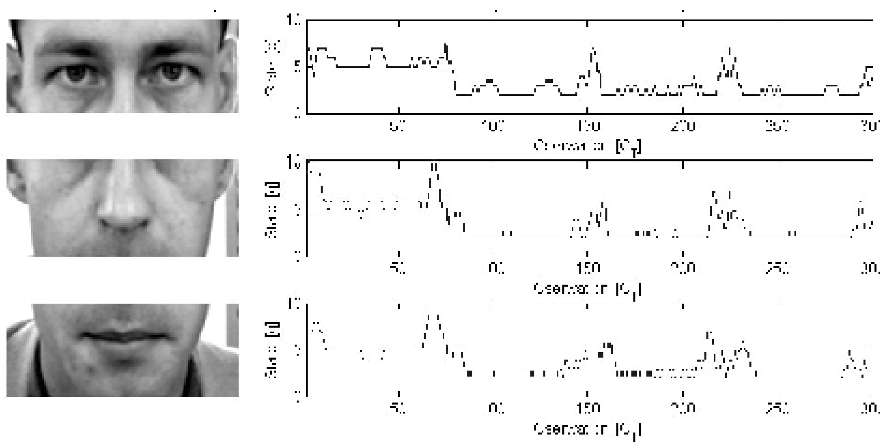

Fig. 5. Parts of face and correspond them sequences of observation

backward algorithm. The forward algorithm calculates the coefficient $a_t(i)$ (probability of observing the partial sequence $(o_1, \ldots, o_t)$ such that state q_t is i). The backward algorithm calculates the coefficient$b_t(i)$ (probability of observing the partial sequence $(o_{t+1}, \ldots, o_T)$ 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 $\log P(O \mid \lambda) - \log P(O \mid \lambda_0) < DELTA$ stop
4. Else set $\lambda_0 \rightarrow \lambda$ and goto step 2.

The parameters of new model λ, based on λ_0 and observation O, are estimated from equation of Baum-Welch algorithm [13], and then are recorded to the database.

2.3 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, 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 [13].

$$P(O \mid \lambda) = \sum_q P(O \mid q, \lambda) P(q \mid \lambda) \tag{2}$$

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

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

3 Experimenting

The following constrains are imposed to guarantee the 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 face database in which there are 24 subjects [14]. The second and the third experiment were carried out using my own face database FaDab in which there are 150 subjects [15]. 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. [16].

Table 1. The results of experiment

Face database	Number of face parts	Number of persons	Error rate [%]
BioID	3	24	12,51
FADAB	3	150	10,00
FADAB	1 (eyes)	150	8,00

4 Conclusion

The new method of face identification and the effective identification system were presented. 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.
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 also.
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.

References

1. Wu Ch.J., Huang J.S. (1990) Human face profile recognition by computer. Pattern recognition, vol. 23, No 3/4: 255–259.
2. Blaszczyk M. (1993) Automatyczne rozpoznawanie twarzy. Rozprawa doktorska, Politechnika lska, Gliwice.
3. Garcia C., Zikos G., Tziritas G. (2000) Wavelet pocket analysis for face recognition. Image and Vision Computing 18: 289–297.
4. R. Brunelli, T. Poggio (1993) Face Recognition: Features versus Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(6): 1042–1052.
5. Kirby M., Sirovich L. (1990) Application of the Karhunen-Loeve Procedure and the Characterization of Human Faces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(1): 103–108.
6. Belhumeur P., Hespanha J., Kriegman D., (1997) Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7): 711–720.
7. Chien J.T., Wu Ch.Ch. (2002) Discriminant Waveletface and Nearest Feature Classifiers for Face Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(12): 1644–1649.
8. Samaria F., Young S. (1994) HMM-based Architecture for Face Identification, Image and Vision Computing, Vol. 12 No 8.
9. Bobulski J. (2003) The Face Recognition Technique Using Wavelet-face, Hidden Markov Model and Maximum Likelihood Principle, ACS'2003, Miedzyzdroje Poland.
10. Misiti M., Misiti Y., Oppenheim G., Poggi J.-M. (2000) Wavelet Toolbox User's Guide, MatLab 5.0, The MathWorks.
11. Kozuani A.Z. He f. Sammut K. (1997) Wavelet Packet Face Representation and Recognition. Proc. IEEE Conf. Systems, Man, and Cybernetics: 1614–1619.

12. Rabiner L. R. (1989) A tutorial on hidden Markov models and selected application in speech recognition. Proc. IEEE 77: 257–285.
13. Kanungo T. Hidden Markov Model Tutorial, http://www.cfar.umd.edu/ kanungo.
14. Face Database BioID, http://www.humanscan.com/bioid.htm.
15. Face Database FaDab, http://icis.pcz.pl/ januszb/baza.htm
16. Bobulski J. (2004) The Method of User's Identification Using the Fusion of the Wavelet Transformand Hidden Markov Models, PhD Thesis, Czestochowa University of Technology.