

Dr inż. Janusz Bobulski*

METODA IDENTYFIKACJI TWARZY W OPARCIU O FUZJE TRANSFORMACJI FALKOWEJ I UKRYTYCH MODELI MARKOWA

Problem identyfikacji osób jest obecnie czołowym zagadnieniem wielu ośrodków badawczo-rozwojowych. Zainteresowanie tą dziedziną wynika z potencjalnych możliwości praktycznego zastosowania nowych rozwiązań w identyfikacji osób w systemach wymagających autoryzacji dostępu osób uprawnionych do korzystania z potencjalnych zasobów. W pracy przedstawiono propozycje metody identyfikacji osób w oparciu o fuzję transformacji falkowej (ekstrakcja cech) oraz modele Markowa (proces decyzyjny). Oba wspomniane narzędzia matematyczne stosowane były dotychczas oddzielnie w rozpoznawaniu twarzy. Zaproponowana metoda przedstawia możliwości ich wspólnego zastosowania w metodzie identyfikacji twarzy.

1. WPROWADZENIE

Rosnąca popularność systemów rozpoznawania twarzy spowodowana jest specyfiką metody, która nie wymaga od użytkowników podejmowania akcji, nie jest inwazyjna. Identyfikowane osoby mogą nie wiedzieć, że w ich bezpośrednim otoczeniu znajduje się system, który właśnie je rozpoznaje. System rozpoznawania odcisków palców potrzebuje przyłożenia palca do czytnika linii papilarnych. Systemy rozpoznawania twarzy ograniczają się jedynie do pozyskania obrazu twarzy lub ich sekwencji, po czym w automatyczny sposób dokonują analizy obrazu i podejmują decyzję czy dana osoba figuruje w bazie danych, czy też nie [1].

Systemy rozpoznawania twarzy pod względem działania można podzielić na dwa typy: weryfikacji i identyfikacji. W przypadku weryfikacji system porównuje twarz użytkownika ze znanym mu wzorcem. Określenie identyczności następuje na podstawie miary podobieństwa dwóch obiektów (twarzy), porównując je jeden do jednego. W oparciu o weryfikację wspomniane systemy wykonują zadania autoryzacji

* Politechnika Częstochowska, Instytut Informatyki Teoretycznej i Stosowanej, 42-201 Częstochowa, ul. Dąbrowskiego 73, janusz.b@icis.pcz.pl

(autentyfikacji), czyli potwierdzenia tożsamości osoby na podstawie obrazu twarzy [Kuk03]. Systemy realizujące zadania identyfikacji porównują twarz użytkownika z twarzami zapisanymi w bazie, twarz, dla której miara podobieństwa jest największa uznawana jest za rozpoznaną. Systemy tego typu mają zastosowanie w zadaniach klasy „*access control*”, czyli kontroli dostępu do potencjalnych zasobów, np. ochronie pomieszczeń, dostęp do stanowisk komputerowych.

Potencjalne zastosowania metod identyfikacji osób z wykorzystaniem obrazu twarzy jest szerokie. Można je wykorzystać do uniemożliwienia nieautoryzowanych prób dostępu do bankomatów, komputerów osobistych, sieci komputerowych, telefonów komórkowych, budynków, a także w prewencji, ściganiu przestępców, identyfikacji osób zaginionych, kontroli paszportowych oraz wielu innych dziedzin życia.

W niniejszej pracy przedstawiono propozycje metody identyfikacji osób w oparciu o ukryte modele Markowa. Identyfikacja opierać się będzie na cyfrowym obrazie twarzy w skali szarości. Do ekstrakcji cech zastosowano transformację falkową, a proces decyzyjny wykorzystuje modele Markowa. Oba wspomniane narzędzia matematyczne stosowane były dotychczas oddzielnie w rozpoznawaniu twarzy. Zaproponowana metoda przedstawia możliwości ich wspólnego zastosowania w metodzie identyfikacji twarzy.

2.OPRACOWANIE SYSTEMU *FaMar* DO IDENTYFIKACJI OBRAZÓW TWARZY

2.1 PRZETWARZANIE WSTĘPNE

Pierwszym etapem w procesie uczenia jest wczytanie obrazu i jego obróbka wstępna. Dopuszczalne są odstępstwa w rozmiarze zdjęcia, gdyż po wczytaniu jest ono skalowane do rozmiaru wys.400 x szer.300 pikseli. Obraz poddawany jest filtracji przy pomocy filtru medianowego, w celu usunięcia zakłóceń w obrazie, które mogą być przyczyną błędów. Filtr medianowy przetwarza dany punkt obrazu jako medianę z wartości punktów go otaczających [2].

Drugim etapem jest normalizacja obrazu twarzy. Jest to niezbędny proces, za pomocą którego możliwe jest porównywanie obiektów, jakimi są w tym przypadku twarze. Założeniem normalizacji jest odpowiednie przeskalowanie obrazu twarzy tak, aby mogły one być porównywane. Pierwszą czynnością w procesie normalizacji jest wyznaczenie środków oczu. Dokonuje się tego z użyciem metody projekcji przez obliczenie pseudo-gradientu wartości pikseli całego obrazu, a następnie sumowaniu tych wartości dla każdego wiersza oraz kolumny [3].

Otrzymano w ten sposób dwa wektory odpowiednio sumy gradientów w poziomie i w pionie. Następnie znajdujemy wartości maksymalne tych wektorów, dla sumy

gradientów wierszy jedną wartość wyznaczającą linię oczu (współrzędna y -owe), a dla sumy gradientów kolumn dwie wartości (współrzędne x -owe). Rezultat tego procesu przedstawiono na Rys. 1.

Po wyznaczeniu środków oczu, zdjęcie twarzy skalowane jest tak, aby odległość między nimi wynosiła 60 pikseli. Dokonuje się tego poprzez wyznaczenie współczynnika skalowania f_s i przeskalowania względem niego całego obrazu. Dodatkowo odrzucana jest ta część obrazu, która leży poza obiektem, czyli nie będąca twarzą. Zapewnia to normalizację wszystkich twarzy, pozbycie się części wprowadzających zakłócenia oraz możliwość ich porównywania (Rys. 1).



Rys.1 Proces normalizacji obrazu twarzy

Fig. 1 The normalization process of the face image

Drugą częścią przetwarzania wstępnego jest podział twarzy na trzy obszary. Ze znormalizowanego obrazu twarzy wydzielane są obszary zawierające najwięcej informacji oraz gwarantujące wystarczającą ilość danych potrzebnych do procesu identyfikacji. Do tych obszarów należą rejon oczu, nosa i ust. Czoło i broda są odrzucane, ponieważ zawierają najmniej informacji oraz wprowadzają najwięcej zakłóceń. Podziału dokonuje się w oparciu o proporcje twarzy, czyli położenie jej elementów względem pozycji środków oczu.

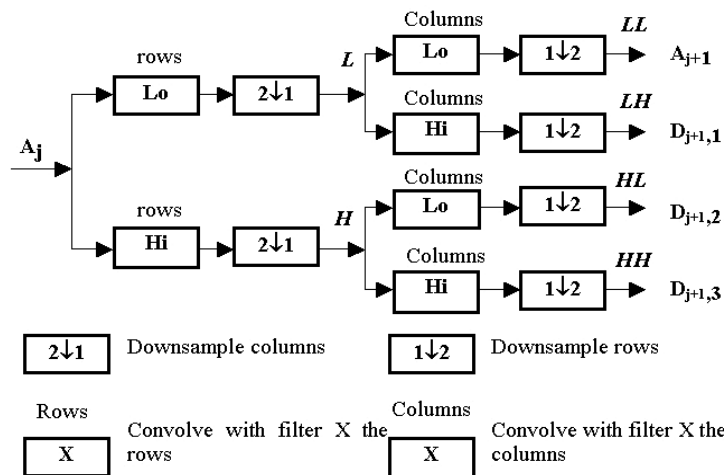


Rys.2 Wynik wydzielenia obszarów twarzy

Fig.2 The result of face area extract

2.2 EKSTRAKCJA CECH

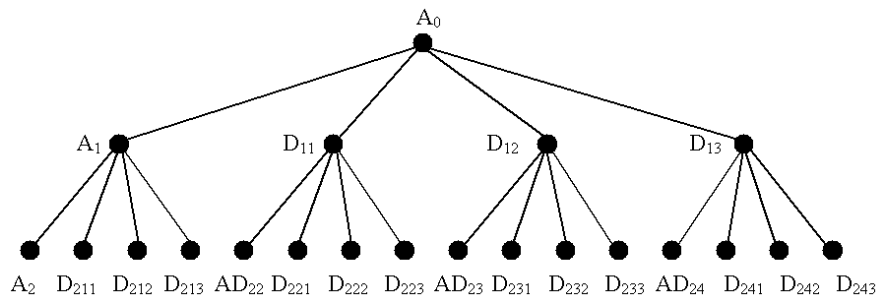
W procesie ekstrakcji cech została wykorzystana transformacja falkowa (TFL). Stosując 2D TFL obraz twarzy jest rozkładany przez filtry dolno- i górnoprzepustowe na cztery podobrazy. Wynikowe obrazy posiadają różną energię i tak podobraz LL (A_j) ma największą i jest pomniejszoną kopią oryginalnego obrazu. Natomiast pozostałe podobrazy LH , HL , i HH wydobywają zmieniające się składniki obrazu odpowiednio w kierunku poziomym (D_{j1}), pionowym (D_{j2}) oraz diagonalnym (D_{j3}) (Rys. 3) [4].



Rys.3 Schemat dwuwymiarowej transformacji falkowej pierwszego poziomu

Fig.3 Scheme of one-level two-dimensional Wavelet Transform

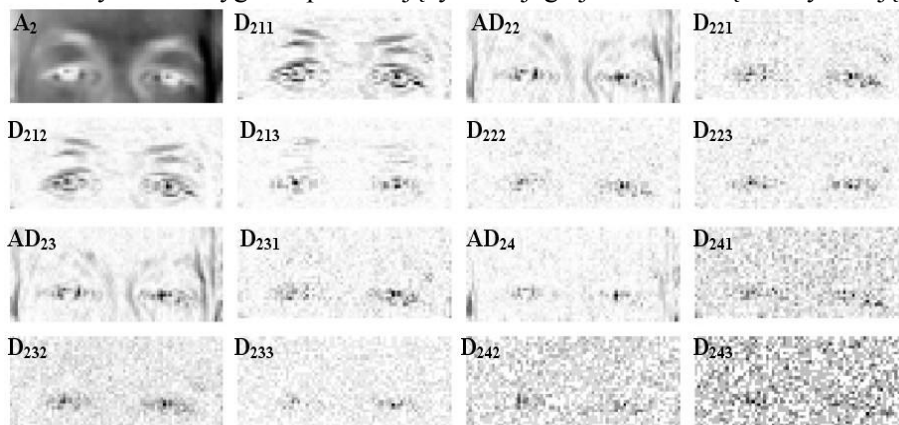
W proponowanej metodzie zastosowano dwuwymiarową transformację falkową drugiego poziomu. Po dokonaniu transformacji pierwszego poziomu obrazy wyjściowe stają się obrazami wejściowymi drugiego (Rys.4) [5]. Piksele w obrazach wynikowych przyjmują wartości od 0 do 255.



Rys. 4 Drzewo dekompozycji falkowej

Fig. 4 The wavelet decomposition tree

Gwarancją na pożądaną poziom skuteczności rozpoznawania jest odpowiedni dobór banku filtrów w transformacji falkowej. Wybór ten zależy od rodzaju przetwarzanej informacji i należy go dostosować do poszczególnych przypadków indywidualnie uwzględniając właściwości poszczególnych sygnałów [6]. Niewłaściwy dobór funkcji falkowej spowoduje problemy z analizą i identyfikacją badanego sygnału. Natomiast wybór odpowiedniego banku filtrów zagwarantuje selekcję indywidualnych cech sygnału pozwalających na jego jednoznaczną identyfikację.

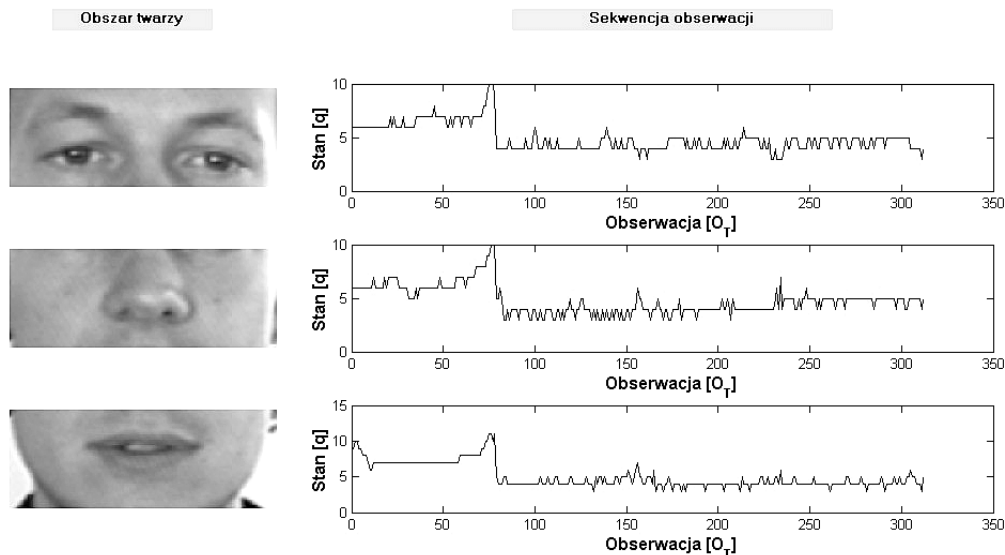


Rys. 5 Wynik transformacji falkowej drugiego poziomu obszaru oczu

Fig. 5 Example of level 2 of the wavelet decomposition of eyes image

Wyniki otrzymane z transformacji falkowej drugiego poziomu muszą być zakodowane do postaci akceptowanej przez model Markowa. Najprostszą metodą

kodowania informacji, jaka może być zastosowana jest odchylenie standardowe σ_X lub wartość średnia \bar{X} . W tym celu wylicza się odchylenie standardowe poszczególnych kolumn macierzy współczynników MW otrzymanych z transformaty falkowej drugiego poziomu obrazu reprezentującego poszczególne obszary twarzy.



Rys. 6 Widoki obrazów obszarów twarzy i odpowiadające im sekwencje obserwacji O_T

Fig.6 The images of face areas and corresponding to them sequences of observation O_T

2.3 UCZENIE

W procesie identyfikacji zastosowano Ukryte Modele Markowa, które są podwójnym procesem stochastycznym możliwym do zaobserwowania poprzez inny proces stochastyczny generujący sekwencje obserwacji O_T . Statystycznymi parametrami modelu są [7, 8]:

- Liczba stanów modelu N .
- Macierz prawdopodobieństwa przejść $A=\{a_{ij}\}$ $1 \leq i, j \leq N$, gdzie a_{ij} jest prawdopodobieństwem przejścia ze stanu i do stanu j , pod warunkiem $\sum_{j=1}^N a_{ij} = 1$.
- Macierz prawdopodobieństwa obserwacji $B=\{b_j(O_t)\}$, $1 \leq j \leq N$, $1 \leq t \leq T$, oznaczającą prawdopodobieństwo wygenerowania t -tej obserwacji w stanie j .
- Wektor prawdopodobieństw początkowych stanów $\pi = \{\pi_i\}$ $1 \leq i \leq N$.

W metodzie *FaMar* dla każdej części twarzy generowany jest jeden model. Na początku parametry modelu generowane są losowo, a następnie estymuje się je przy pomocy algorytmu Bauma-Welcha [9]:

1. Wygenerowanie losowe modelu λ_0 .
2. Wylczenie nowego λ w oparciu o λ_0 i obserwacje O .
3. Jeśli $\log P(O|\lambda) - \log P(O|\lambda_0) < DELTA$, to stop.
4. W przeciwnym razie, ustaw $\lambda_0 \leftarrow \lambda$ i przejdź do punktu 2.

Po wyznaczeniu parametrów modelu zapisywany jest on do bazy danych. Każda twarz reprezentowana jest przez trzy modele, dla każdego obszaru po jednym

$$\lambda^i = (A^i, B^i, \pi^i) \quad (1)$$

i –obszar twarzy: 1-oczy, 2-nos, 3-usta

Ponadto do bazy zapisujemy obraz twarzy, jednak nie jest on wykorzystywany do identyfikacji a jest jedynie używany do wyświetlenia informacji przy identyfikacji.

Doboru architektury i parametrów rodziny UMM dokonano eksperymentalnie. Zastosowano model ergodyczny, czyli taki, w którym dozwolone są przejścia między wszystkimi stanami, w dowolnej kolejności. Liczbę N stanów Q ustalono na 28, w związku z czym macierz przejść A ma wymiary 28×28 . Mniejsza liczba stanów jest niewystarczająca do poprawnej identyfikacji. Liczbę T obserwacji O ustanowiono na 14, co powoduje, że rozmiar macierzy obserwacji ma wymiar 28×14 . Wektor prawdopodobieństw początkowych π ma długość 28. Do zapisu informacji o jednej osobie potrzebne jest 36908 B (36 kB) [10].

2.4 IDENTYFIKACJA

Proces identyfikacji polega na wyliczeniu prawdopodobieństwa wygenerowania obserwacji O przez dany model λ pobrany z bazy danych, reprezentowany przez trzy parametry A, B, π . Wybierany jest ten model, dla którego prawdopodobieństwo jest największe. Zastosowano metodę maksymalnego prawdopodobieństwa [10].

W metodzie *FaMar*, twarz reprezentowana jest przez trzy modele λ . Prawdopodobieństwo wygenerowania obserwacji $P(O|\lambda)$ liczone jest oddzielnie dla każdego modelu reprezentującego jeden z obszarów twarzy. Następnie prawdopodobieństwa są sumowane – sumę oznaczono jako PF . Twarz, dla której wartość PF jest największa wybierana jest jako poprawna

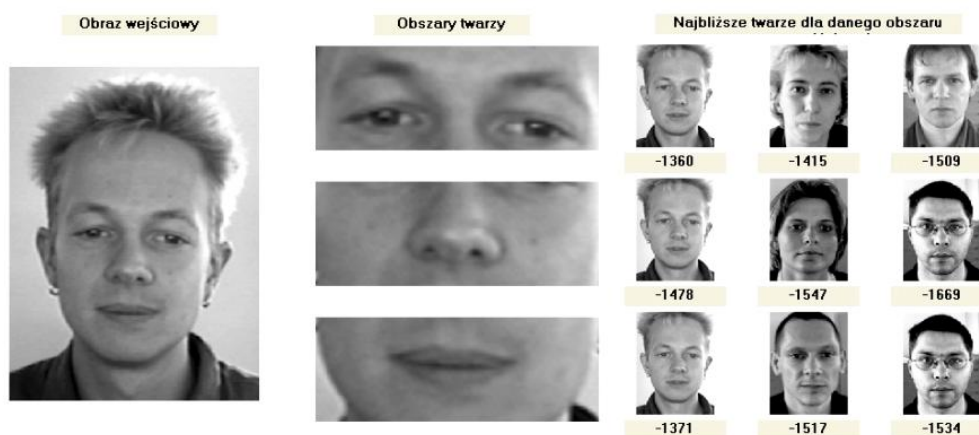
3. EKSPERYMENTALNY POMIAR POZIOMU BŁĘDÓW I CZASU IDENTYFIKACJI ZA POMOCĄ SYSTEMU *FaMar*

Tab. 1 Wyniki eksperymentu
Tab. 1 The results of experiment

Nazwa bazy	Liczba zestawów testowych [szt.]	Liczba części twarzy [szt.]	Liczba twarzy osób w zestawie [szt.]	Liczba twarzy błędnie rozpoznanych [szt.]	Poziom błędów [%]
<i>BioID</i>	10	3	24	3	12.51
<i>FaDab</i>	6	3	150	15	10.00
<i>FaDab</i>	6	1 (oczy)	150	12	8.00

Na podstawie przeprowadzonych eksperymentów można stwierdzić, że prezentowana metoda identyfikacji osób jest poprawna. Charakteryzuje się ona zadowalającą skutecznością zarówno przy identyfikacji w oparciu o trzy obszary twarzy (około 90%), jak i na podstawie obszaru oczu (około 92%).

Ponadto wyniki pozwalają wysnuć wniosek, iż metoda działa poprawnie w przypadku podejmowania decyzji tylko na podstawie obszaru oczu.



Rys. 7 Przykład wyniku wyliczonych logarytmów prawdopodobieństw wygenerowania obserwacji dla poszczególnych obrazów obszarów twarzy

Fig.7 The example result of calculation the probability of generate observations for each face area images

Tab. 2 Porównanie poziomów błędów identyfikacji
Tab.2 The comparison of error rates

<i>Metoda</i>	<i>Poziom błędów [%]</i>
Eigenface [2]	6.00
HMM [11]	16.00
<i>Wavelet-Face</i> [4]	6.00
FaMar	8.00

W porównaniu do istniejących metod identyfikacji twarzy skuteczność metody *FaMar* jest porównywalna z nimi. Poziom błędów metody *FaMar* wynosi 8%, podczas gdy poziom błędów istniejących metod waha się 3-16%.

Atutem metody *FaMar* jest jej złożoność obliczeniowa. Potrzebuje ona 2.5 s na nauczenie modeli reprezentujących twarz i 0.3 s na identyfikację jednego obrazu twarzy. W przypadku, gdy identyfikacja opiera się tylko na obszarze oczu czas ten skraca się trzykrotnie i wynosi 0.1s.

PODSUMOWANIE

Praca dotyczy opracowania metody *FaMar* identyfikacji użytkownika na podstawie obrazu frontalnego twarzy, w której wykorzystana jest fuzja transformacji falkowej i ukrytych modeli Markowa dla trzech obszarów twarzy; decyzja zostaje podjęta na podstawie maksymalizacji prawdopodobieństwa pojawienia się obserwacji. Przedstawiona metoda charakteryzuje się nowatorstwem w zakresie wykorzystania trzech obszarów twarzy (oczy, nos, usta) do identyfikacji i zbudowaniem dla każdego z nich niezależnego ukrytego modelu Markowa, które można wykorzystać oddzielnie lub razem, w zależności od zagadnienia.

W porównaniu do rozpatrywanych metod identyfikacji twarzy charakterystyki programu *FaMar* jest porównywalna z innymi. Poziom błędów rozpoznawania za pomocą systemu *FaMar* wynosi około 8%, podczas gdy poziom błędów istniejących metod waha się 3-16%.

LITERATURA

- [1] FROMHERZ T., STUCKI P., BICHSEL M., *A Survey of Face Recognition*, <http://www.ifi.unizh.ch/~fromherz>, 2002
- [2] WRÓBEL Z., KOPROWSKI R., *Przetwarzanie Obrazu w Programie MATLAB*, Wydawnictwo Uniwersytetu Śląskiego, Katowice, 2001
- [3] KUKHAREV G., KUZMINSKI A., *Techniki Biometryczne Część 1 – Metody Rozpoznawania Twarzy*, Politechnika Szczecińska Wydział Informatyki, 2003

- [4] CHIEN J.-T., WU CH.-CH., *Discriminant Waveletfaces and Nearest Feature Classifiers for Faces Recognition*, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 24 No. 12, December, 2002, 1644-1649
- [5] GARCIA C., ZIKOS G., TZIRITAS G., *Wavelet Packet Analysis for Face Recognition*, Image and Vision Computing 18, 2000, 289-297
- MISITI M., MISITI Y., OPPENHEIM G., POGGI J.-M., *Wavelet Toolbox User's Guide*, MathWorks, 1998
- [7] LI X., PARIZEAU M., PLAMONDON R., *Training Hidden Markov Models with Multiple Observations – A Combinatorial Method*, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 22 No.4, 2000, 371-377
- [8] KWONG S., CHAU K.F., TANG K.S., *Optimization of HMM Topology and its Model Parameters by Genetics Algorithms*, Pattern Recognition 34, 2001, 509-522
- [9] KANUNGO T., *Hidden Markov Model Tutorial*, www.cfar.umd.edu/~kanungo, 2003
- [10] BOBULSKI J., *Metoda Identyfikacji Użytkownika w Oparciu o Fuzje Transformacji Falkowej i Ukrytych Modeli Markowa*, Rozprawa Doktorska, Politechnika Częstochowska, 2004
- [11] SAMARIA F, YOUNG S., *HMM-based Architecture for Face Identification*, Image and Vision Computing, Vol. 12 No 8 October, 1994, 537-543

THE METHOD OF USER'S IDENTIFICATION USING THE FUSION OF THE WAVELET TRANSFORM AND HIDDEN MARKOV MODELS

A problem of persons' identification is one of the main questions of many research centres 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 work concerns creations of the new *method FaMar* 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 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.

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.