

**REVIEW OF SELECTED TECHNIQUES AND
MEASURES AFFECTING THE RELIABILITY OF FACE
IMAGE ACQUISITION SYSTEMS**

**PRZEGLĄD WYBRANYCH TECHNIK I ŚRODKÓW
WPLYWAJĄCYCH NA NIEZAWODNOŚĆ DZIAŁANIA
SYSTEMÓW AKWIZYCJI OBRAZU TWARZY**

Michał Wiśnios, Tadeusz Dąbrowski

Wojskowa Akademia Techniczna, Wydział Elektroniki
e-mail: mwisnios@wat.edu.pl, tdabrowski@wat.edu.pl

***Abstract:** The theme of the publication is to review techniques and measures affecting the reliability of the process of image acquisition face. This issue is closely linked to the biometric identification based on facial image. The focus here in particular on factors external / environmental, not taking into account the quality of this same algorithms. Tests were carried out both on a database built on the model of the face as well as on the basis of the actual images. The object of the research being predestined is not only for stationary applications, but can also be used in mobile objects.*

***Keywords:** facial image, biometrics face, biometric identification*

***Streszczenie:** Tematem opracowania jest przegląd technik i środków wpływających na niezawodność procesu akwizycji obrazu twarzy. Problematyka ta ściśle związana jest z identyfikacją biometryczną opartą o obraz twarzy. Skupiono się tu w szczególności na czynnikach zewnętrznych - środowiskowych, bez uwzględniania jakości działania samych algorytmów. Badania zostały przeprowadzone zarówno na bazie danych zbudowanej w oparciu o model (fantom) twarzy jak i w oparciu o obrazy rzeczywiste. System będący obiektem badań predystynowany jest nie tylko do zastosowań stacjonarnych, ale może być również wykorzystany w obiektach mobilnych.*

***Słowa kluczowe:** obraz twarzy, biometria twarzy identyfikacja biometryczna*

1. Wstęp

Własności charakterystyczne twarzy są danymi biometrycznymi, które z jednej strony stanowią najstarszą ze znanych i najbardziej naturalną z metod rozpoznawania człowieka, natomiast z innego punktu widzenia stwarzającą największą trudność w automatyzacji procesu rozpoznawania. Przyczyn, które mają wpływ na taki stan rzeczy jest bardzo wiele. Przede wszystkim obraz twarzy, z założenia, powinien umożliwiać bezinwazyjną akwizycję i nie powinien wymagać od osoby będącej obiektem tego procesu współpracy z systemem. Założenie to jest głównym powodem występowania wspomnianych trudności, gdyż twarz charakteryzuje się naturalną zmiennością. Cecha ta wynika nie tylko z procesu starzenia, ale również z naturalnych stanów emocjonalnych, chorobowych i zmęzeniowych. Dodatkowy wpływ ma tu również mimika i grymasy wywołane np. mową podczas akwizycji obrazu. Istotną przeszkodą w pozyskiwaniu wiarygodnego obrazu są również czynniki maskujące prawdziwy obraz twarzy takie jak np. makijaż, okulary, broda czy wąsy. Ważną grupą utrudnień stojących na drodze automatycznego pozyskiwania obrazu twarzy są warunki obserwacji panujące podczas akwizycji. Poza kątem obserwacji [7] należy tu uwzględnić również warunki oświetlenia odpowiedzialne na przykład za cienie i refleksy [1, 2, 10].

Pomimo szeregu problemów, stojących na drodze rozwoju tej techniki biometrycznej, prace w tym obszarze są bardzo zaawansowane ze względu na jej znaczenie w procesie monitorowania miejsc publicznych. Jednym z miejsc predystynowanych do zastosowania ww. technik identyfikacji jest zdalna identyfikacja osób np. z zastosowaniem zdalnie sterowanych bezzałogowych obiektów latających. Przy obecnym poziomie techniki automatyczne rozpoznawanie człowieka na podstawie obrazu twarzy stanowi tanią i wygodną metodę rozpoznawania biometrycznego, głównie tam gdzie uzyskanie wysokiego współczynnika wiarygodności rozpoznawania nie jest krytyczne.

Systemy automatycznego rozpoznawania twarzy zasadę swojego działania opierają na intuicyjnych procedurach stosowanych w takim procesie przez człowieka. Systemy te początkowo skupiały się na szczegółowej analizie poszczególnych charakterystycznych elementów osobniczych takich jak kolor oczu, kształt nosa czy też warg [11]. Przyczyną takiego podejścia był fakt, iż moce obliczeniowe komputerów były do niedawna stosunkowo niskie, natomiast biometryka ta charakteryzuje się tym, iż dedykowana jest do zastosowań na bardzo obszernych bazach danych. Aktualne badania dążą w kierunku syntetycznej, holistycznej analizy całej twarzy. Z badań wynika, że istnieją psychologiczne uwarunkowania łatwości rozpoznawania twarzy [1]. Dodatkowo potwierdzone zostało, że wystarczy jedynie uwypuklenie jednego charakterystycznego elementu twarzy np. na karykaturze, aby bez trudności rozpoznać daną osobę. De facto w psychologii widzenia istnieje zjawisko bezwarunkowej analizy poszczególnych elementów twarzy pozwalające bez problemów rozpoznawać twarze poddane manipulacjom takim jak np. odwracanie elementów.



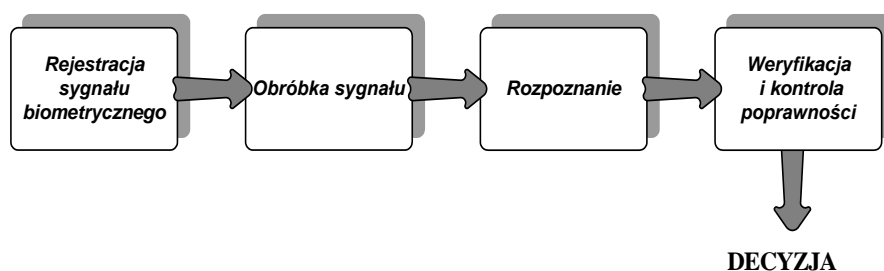
Rys. 1. Ilustracja manipulacji obrazem twarzy Napoleona polegająca na odwróconej orientacji oczu i ust [2]

Obecnie najszerzej rozwijane są metody rozpoznawania twarzy bazujące na statystycznej bądź probabilistycznej technice analizy. Zakładają one, że podstawa subiektywności twarzy jest rezultatem szeregu czynników, nie zawsze możliwych do zdefiniowania.

2. Charakterystyka problemu

W przebiegu procesu działania systemu rozpoznawania twarzy można wydzielić cztery podstawowe etapy:

1. Pozyskanie obrazu z urządzenia wejściowego (np. kamery cyfrowej);
2. Wstępne przetworzenie otrzymanego obrazu (tj.: wykrycie twarzy na obrazie, normalizacja obrazu, procedury przyspieszające rozpoznawanie, upraszczające obliczenia redukujące przestrzeń cech);
3. Rozpoznanie (np. metodą metryki L0);
4. Weryfikacja i kontrola poprawności (np. potwierdzenie wyników wcześniejszego rozpoznania innym algorytmem rozpoznawania).



Rys. 2. Schemat systemu identyfikacji biometrycznej

W trakcie procesu pozyskania obrazu z urządzenia wejściowego system dokonuje akwizycji obrazów twarzy. Akwizycji tej można dokonywać w zakresie widzialnym i w podczerwieni [6].

Pierwszy z zakresów jest szczególnie interesujący ze względu na dojrzałość i zaawansowanie techniki pozyskiwania takiego obrazu. Zarówno kamery analogowe jak i cyfrowe posiadają szereg wad i zalet, jednakże tendencja rozwojowa zmierza w kierunku cyfryzacji technik akwizycji obrazu - szczególnie obrazu twarzy. Podstawową zaletą takiego podejścia są szerokie możliwości kompresji i przetwarzania tak otrzymanego obrazu. Gwałtowny rozwój technik akwizycji obrazu twarzy nastąpił, gdy rozwinęła się technologia kamer cyfrowych IP. Wcześniejsze prace dotyczyły głównie analiz wykorzystujących przetworzony na postać cyfrową obraz uzyskany za pomocą kamer analogowych.

Głównym problemem w procesie akwizycji obrazu twarzy w zakresie widzialnym jest zapewnienie odpowiednich warunków otoczenia m.in. odpowiedniego oświetlenia. W związku z tym, iż proces ten odbywa się najczęściej automatycznie, wskazane jest, aby warunki podczas akwizycji obrazu twarzy były jak najbardziej zbliżone do tych panujących w trakcie tworzenia bazy danych. Część parametrów może zostać skorygowana w późniejszych etapach działania systemu. Zabiegi tego typu są przeprowadzane głównie software'owo. Pożądanym rozwiązaniem jest opracowanie metod niepodatnych na zmienne warunki akwizycji.

Istnieją również systemy akwizycji obrazu twarzy typu „off-line” pozwalające operatorowi takiego systemu na korygowanie poszczególnych parametrów obrazu, które nie spełniają założonych warunków. Rozwiązanie to ma korzystny wpływ na skuteczność procesu rozpoznawania, jednakże powoduje znaczne spowolnienie działania takiego systemu.

Obecnie dobrze rozwinięte są techniki rozpoznawania obrazu twarzy osoby znajdującej się centralnie - przodem do kamery. Zaawansowane są także badania mające na celu rozpoznawanie człowieka na podstawie trójwymiarowego obrazu twarzy - 3D [5].

Aby systemy tego typu działały skutecznie konieczne jest rozwiązanie następujących problemów:

- poprawne zlokalizowanie obrazu zawierającego twarz jednej osoby tj. wyodrębnienie jej z tła, z otoczenia innych twarzy lub elementów przypominających ludzką twarz kształtem i kolorem - np. z powodu uniesionej dłoni. Należy także wykluczyć możliwość zlokalizowania twarzy ludzkiej na rysunku. System musi umieć rozróżnić struktury podobne do twarzy ludzkiej i odrzucić je nie powodując przy tym straty obiektów będących twarzami[4];
- wybranie tych części obrazu twarzy, które zawierają jak największą ilość informacji;
- eliminacja wpływu obrotu, skali i przesunięcia twarzy;
- eliminacja wpływu czynników zmiennych (np. mimiki, uczesania, okularów itp.);
- eliminacja wpływu jakości oświetlenia.

3. Czynniki wpływające na niezawodność działania systemu akwizycji obrazu twarzy

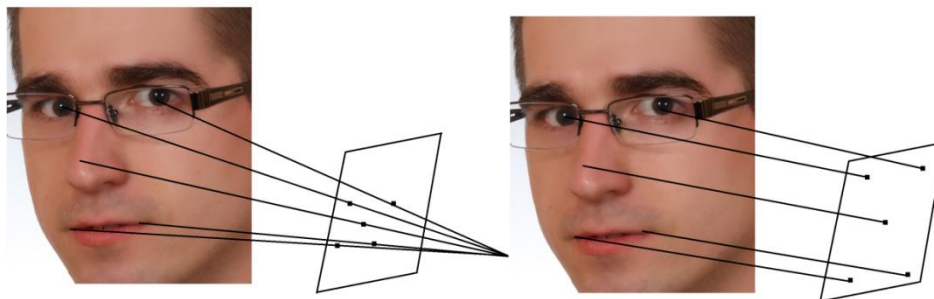
O niezawodności sprzętowej systemu akwizycji obrazu twarzy decyduje bezawaryjna praca następujących jego elementów: modułu akwizycji źródłowego obrazu (jest to na ogół obiektyw kamery), urządzeń przełączających i przetwarzających obraz (m.in. multiplekserów), urządzeń do zapisu informacji, urządzeń służących do ekspozycji obrazu (m.in. monitorów, wyświetlaczy wielkoformatowych), zasilaczy, oświetlaczy oraz przewodów i złączy.

Skuteczność systemu akwizycji obrazu twarzy istotnie zależy od właściwości urządzeń rejestrujących tj. kamer. W celu poprawnego doboru kamery i jej obiektywu, należy zwrócić szczególnie uwagę na następujące parametry tych elementów:

- czułość kamery i liczba aperturowa obiektywu (przy najmniej korzystnych warunkach oświetlenia);
- ogniskowa obiektywu (według wymagań pola widzenia);
- rozdzielczość kamery i obiektywu (jakość obrazu);
- wybrana kombinacja obiektywu i kamery;
- pole obrazowe obiektywu - powinno być równe lub większe od pola przetwornika obrazu w kamerze.

Do zadań polegających na akwizycji obrazu twarzy najodpowiedniejsze wydają się kamery wykonane w technologii IP. Dzięki zgodności z otwartymi standardami IP kamery sieciowe mogą łączyć się z siecią IP dowolnego rodzaju, w tym także z Internetem, w celu umożliwienia oglądania i rejestrowania obrazu z dowolnego miejsca na świecie. Umożliwiają one prawidłową komunikację z serwerem oraz posiadają możliwość sprzętowej obróbki sygnału, co w znacznym stopniu ogranicza liczbę przesyłanych danych.

Do akwizycji obrazu twarzy z niewielkiej odległości najlepsze są kamery megapikselowe/HDTV. Na szczególną uwagę zasługują tu kamery posiadające rozdzielczość rzędu kilku megapikseli, gdyż liczba danych przechwytywanych z takiej kamery jest jeszcze możliwa do stosunkowo szybkiej obróbki. Rozdzielczość tych kamer jest co najmniej trzy razy większa od rozdzielczości kamer analogowych. Łącząc sieciowe kamery megapikselowe z wybranymi kamerami bez technologii megapikselowej, ale przystosowanymi do innych potrzeb - np. charakteryzujących się silnym zoomem optycznym, wyjątkową czułością na światło lub niskimi kosztami - możemy uzyskać efektywny, niezawodny i ekonomiczny system akwizycji obrazu twarzy. Natomiast do monitoringu twarzy w tłumie najlepiej nadają się kamery o możliwie wysokiej rozdzielczości. Umieszczenie takich kamer w znacznej odległości powoduje pozbycie się efektu ustawienia twarzy pod niewłaściwym kątem.



Rys. 4. Rzutowanie przestrzeni 3-wymiarowej na 2-wymiarowy obraz:
Lewa strona - rzut zbieżny; prawa strona - rzut równoległy

Zagadnienie rzutowania obrazu z przestrzeni 3D na płaszczyznę 2D (rys.4) występuje głównie podczas analizy obrazu twarzy. Zagadnienie rzutowania zbieżnego ma miejsce wówczas, gdy rozmiar obserwowanego obiektu i jego odległość od kamery są porównywalnego rzędu. W miarę odsuwania takiego obiektu od kamery na znaczną odległość, wystarczające jest stosowanie modelu rzutu równoległego [8].

4. Miary oceny jakości akwizycji obrazu twarzy

Ocena jakości działania metody akwizycji obrazu twarzy wymaga wcześniejszego zdefiniowania metodyki oceny jakości przetworzonego obrazu. Z pomocą przychodzi tu zmodyfikowana metodyka obiektywnej oceny jakości obrazów medycznych [12]. Do szczególnie oczekiwanych cech miary oceny obiektywnej należy zaliczyć wysoki poziom korelacji z subiektywną oceną jakości obrazu oraz dużą podatność na analizę obliczeniową. Rozpatrując miary skalarne możemy stwierdzić, iż otrzymanie odpowiedniej korelacji z oceną subiektywną jest szczególnie trudne. Jednak miary te pozwalają na przejrzystą interpretację i analizę porównawczą.

Zakładając, że oryginalny obraz cyfrowy, wielopoziomowy, ze skalą szarości o szerokości M i wysokości N jest opisany funkcją jasności:

$$0 < f(m,n), 0 \leq m < M, 0 \leq n < N$$

to można wyznaczyć dla niego takie użyteczne miary jak [3]:

- maksymalna różnica (szczytowy błąd bezwzględny) PAE (ang. *Peak Absolute Error*):

$$MD = \max_{m,n} \{|f(m,n) - \tilde{f}(m,n)|\} \quad (1)$$

- błąd średniokwadratowy (ang. *Mean Square Error*):

$$MSE = \frac{1}{MN} \sum_{m,n} [f(m,n) - \tilde{f}(m,n)]^2 \quad (2)$$

- szczytowy stosunek sygnału do szumu (ang. *Peak Signal to Noise Ratio*):

$$PSNR = 10 \log \frac{MN [\max_{m,n} \{f(m,n)\}]^2}{\sum_{m,n} [f(m,n) - \tilde{f}(m,n)]^2} \quad (3)$$

gdzie:

$\max_{m,n} \{f(m,n)\}$ to największa możliwa, nie zawsze rzeczywista, wartość funkcji jasności, np. 255 danych 8-bitowych;

- średnia różnica:

$$AD = \frac{1}{MN} \sum_{m,n} |f(m,n) - \tilde{f}(m,n)| \quad (4)$$

- jakość korelacyjna (ang. *Correlation Quality*):

$$CQ = \frac{\sum_{m,n} f(m,n) - \tilde{f}(m,n)}{\sum_{m,n} f(m,n)} \quad (5)$$

- wierność obrazu:

$$IF = 1 - \frac{\sum_{m,n} [f(m,n) - \tilde{f}(m,n)]^2}{\sum_{m,n} [f(m,n)]^2} \quad (6)$$

- miara chi-kwadrat (ang. *chi-square measure*):

$$\chi^2 = \frac{1}{MN} \sum_{m,n} \frac{[f(m,n) - \tilde{f}(m,n)]^2}{f(m,n)} \quad (7)$$

Ograniczeniem dla powyższych miar jest fakt, iż stosuje się je do obrazów monochromatycznych. Natomiast w przypadku obrazów kolorowych wartość miary zniekształcenia odnosi się często do określonej składowej luminancji. Istnieje jednak możliwość określenia tej dla wszystkich składowych przestrzeni kolorów lub wyznaczenia nowej miary, jako odległości euklidesowej w tej przestrzeni.

Najbardziej naturalną miarą jakości obrazu twarzy wydaje się być, liczona w pikselach, rozdzielczość takiego obrazu. Większa liczba mniejszych pikseli powoduje dodanie szczegółów i wyostrenie krawędzi. Analiza obrazów twarzy wskazuje na to, iż zmiana rozdzielczości obrazu nie jest liniowo proporcjonalna do rozmiaru przestrzeni pamięci zajmowanej przez to zdjęcie. W przykładzie podanym w tabeli 1 czterokrotne zmniejszenie rozdzielczości obrazu JPG z 2 do 0,5 Mpx powoduje 10-krotne zmniejszenie zajmowanej pamięci.

Tabela 1. Ilustracja wpływu zmiany rozdzielczości obrazu twarzy na wielkość pliku JPG

Rozdzielczość pozioma [px]	Rozdzielczość w pionie [px]	Rozdzielczość [Mpx]	Rozmiar [KB]
1152	1728	1,99	616
1024	1536	1,57	124
800	1200	0,96	91
640	960	0,61	68
576	864	0,50	66

5. Wyniki

Jednoznaczne określenie parametrów niezawodnego procesu identyfikacji osób na podstawie obrazu twarzy jest zagadnieniem jak dotychczas nierozwiązanym. Na podstawie przeprowadzonych badań można wywnioskować, iż wartościami zbliżonymi do optymalnych w tej kwestii są wymienione poniżej wartości:

- praktycznie wymaganą liczbą pikseli na metr, w procesie identyfikacji twarzy, jest 490. Oznacza to na przykład, że jeśli konieczna jest dokładna identyfikacja osób przechodzących przez obszar 2 na 2 metry to kamera powinna pracować z rozdzielczością 1050x1050, czyli niewiele ponad 1 megapiksel;
- średnie natężenie oświetlenia padającego na twarz wynosi przynajmniej 100lx, przy czym szczególnie istotnym jest w tym przypadku również współczynnik równomierności oświetlenia [13];
- optymalna odległość sztucznego źródła światła od identyfikowanej twarzy powinna wynosić ok 2 m (w przypadku punktowych źródeł światła). Wartość ta związana jest nie tylko z zapewnieniem równomierności oświetlenia, ale również z komfortem psychofizycznym prowadzenia takiego procesu;
- maksymalna wartość kąta akwizycji obrazu twarzy dla obrotu twarzy w pionie wynosi 35°, natomiast dla obrotu twarzy w poziomie wynosi do 25°. Dla osi obrotu prostopadłej do powierzchni twarzy kąt ten wynosi do ok 20°. Niestety wykonanie badań polegających na sumowaniu obrotów w dwóch lub trzech osiach okazało się procesem zbyt czasochłonnym (przy pomiarze ze skokiem tylko co 10° rozpatryć trzeba dodatkowo ok 300 przypadków dla każdego badanego obrazu twarzy). Wyniki te uwarunkowane są głównie przez algorytm automatycznej detekcji obrazu twarzy [14];
- na podstawie przeprowadzonych dotychczas badań można stwierdzić, iż nie ma konieczności aktualizowania bazy danych częściej niż raz na 2 lata, biorąc pod uwagę wyłącznie czynniki starzeniowe obrazu twarzy.

5. Podsumowanie

W opracowaniu zaprezentowano obszar zagadnień związanych z systemami akwizycji obrazu twarzy, a w szczególności z czynnikami wpływającymi na jakość i niezawodność działania takich systemów. Problematyka ta poza klasycznym podejściem do zagadnienia niezawodności systemowej jest skupiona głównie wokół niezawodnego działania samych algorytmów detekcji twarzy. Odpowiednia optymalizacja algorytmów oraz obrazu przekazywanego z detektora ma kluczowe znaczenie dla niezawodnego działania systemów tego typu w czasie rzeczywistym. Zaprezentowana tu zależność rozmiaru statycznego pliku z obrazem twarzy od zastosowanej rozdzielczości stanowi wyłącznie namiastkę tej kwestii w stosunku do streamingu video. Przedstawione na łamach artykułu wyniki mogą stanowić punkt odniesienia nie tylko w zastosowaniu do systemów stacjonarnych, ale również zostać wykorzystane przy definiowaniu parametrów wersji mobilnej np. jedna z funkcjonalności bezzałogowych obiektów latających. Opisane na łamach artykułu kwestie stanowią przegląd i podsumowanie spojrzenia na tę problematykę, wynikającego nie tylko z przeglądu literaturowego i rozwiązań technicznych, ale także z własnych badań autorów. Opracowanie to pozwala na optymalne ukierunkowanie kolejnych prac związanych z poruszoną tematyką. Co więcej może stanowić punkt odniesienia dla osób zajmujących się problematyką pojedynczych elementów systemu identyfikacji osób na podstawie obrazu twarzy. Systemy tego typu mogą również zostać zastosowane do zabezpieczania komunikacji w bezprzewodowym systemie sterowania[15].

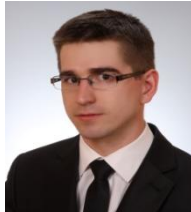
6. Literatura

- [1] P. Sinha, B. Balas, Y. Ostrowsky, R. Russell: Face Recognition by Humans: 19 Results All Computer Vision Researchers Should Know About. Proceedings of the IEEE, Vol.94, No.11, 2006.
- [2] Ślot K.: Wybrane zagadnienia biometrii, WKŁ Warszawa 2008.
- [3] A. Przelaskowski: Miary jakości rozdział w "Multimedia - Algorytmy i Standardy kompresji" W. Skarbek, Akademicka Oficyna Wydawnicza PLJ, 1998.
- [4] Wiśnios M., Dąbrowski T., Bednarek M., Badania weryfikacyjne metody rozpoznawania twarzy, Biuletyn Wojskowej Akademii Technicznej, nr 4/2013, 205-218.
- [5] A. Mian, M. Bennamoun, R. Owens: Automatic 3D Face Detection, Normalization and Recognition.
- [6] Wong Wai Kit; Hui, Joe How; Desa, Jalil Bin Md; Ishak, Nur Izzati Nadiah Binti; Sulaiman, Azlan Bin; Yante Binti Mohd Nor: Face detection in thermal imaging using head curve geometry, Image and Signal Processing (CISP), 2012 5th International Congress on.
- [7] Xiangxin Zhu, D. Ramanan: Face detection, pose estimation, and landmark localization in the wild, Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on Biometrics Compendium, IEEE.

- [8] W. Kasprzak: Rozpoznawanie obrazów i sygnałów mowy, OW PW Warszawa 2009.
- [9] A. Bobcow, M Dąbkowski: Wprowadzenie do analizy obrazu w procesie detekcji i identyfikacji twarzy, XVII Seminarium Zastosowanie komputerów w nauce i technice'2007.
- [10] Xiangxin Zhu, Ramanan D., *Face detection, pose estimation, and landmark localization in the wild*, Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on Biometrics Compendium, IEEE.
- [11] A. Samal, P.A. Iyengar *Automatic recognition and analysis of human faces and facial expressions a survey*, Pattern Recognition, January 1992, Volume 25, Issue 1.
- [12] Betts B.J., Li J., Cosman P.C., Gray R.M. et al. : Image quality in digital mammography. Revision of final report to the Army Medical Research and Materiel Command, Compression and Classification of Digital Mammograms for Storage, Transmission, and Computer Aided Screening, September 1998.
- [13] Wiśnios M., Dąbrowski T., Bednarek M., Analiza wpływu czynników zakłócających proces identyfikacji osób w oparciu o obraz twarzy, X Szkoła-Konferencja MWK 2014 Waplewo 2014.
- [14] Viola P., Jones M., Rapid object detection using a boosted cascade of simple features, Computer Vision and Pattern Recognition, 2001.
- [15] Bednarek M., Dąbrowski T., Wiśnios M.: Bezpieczeństwo komunikacji w rozproszonym systemie sterowania, Przegląd Elektrotechniczny, nr 9/2013, 72-74.



Dr hab. inż. Tadeusz Dąbrowski prof. WAT. Dyrektor Instytutu Systemów Elektronicznych Wydziału Elektroniki WAT. Zainteresowania naukowe koncentruje w obszarze teorii eksploatacji - a w tym głównie na zagadnieniach diagnostyki technicznej obiektów i systemów antropotechnicznych.



Mgr inż. Michał Wiśnios (ur. 1986), absolwent Wydziału Elektroniki Wojskowej Akademii Technicznej (2010). Obecnie doktorant i asystent n-d w Instytucie Systemów Elektronicznych WAT. Zajmuje się metodami wiarygodnego rozpoznawania osób na podstawie cech biometrycznych, a w szczególności identyfikacją na podstawie obrazu twarzy.