

doi:10.15199/48.2015.09.23

## Optymalizacja systemu automatycznego rozpoznawania mowy w warunkach zróżnicowanych torów akustycznych

**Streszczenie.** W referacie pokazano sposoby optymalizacji poszczególnych elementów systemu automatycznego rozpoznawania mowy w warunkach zróżnicowanych torów akustycznych. Dokonano tego za pomocą algorytmu genetycznego poprzez wyselekcjonowanie optymalnego zbioru cech silnie zależnych od głosu mówcy przy jednoczesnej minimalizacji wzajemnej korelacji cech oraz ich wrażliwości na zróżnicowanie toru akustycznego. Przedstawiono również wyniki optymalizacji wybranych elementów klasyfikatora, m.in. liczby rozkładów Gaussa.

**Abstract.** This paper presents optimization methods of an automatic speaker recognition system parameters for various acoustic paths. This idea has been performed using the genetic algorithm. The main goal of the proposed algorithm is to select the most important features, minimize correlation between the features and their sensitivity for various acoustic path. Additionally results of classifier optimization (for example number of Gaussian distributions), have been shown. (**Optimization of the automatic speaker recognition system for different acoustic paths**).

**Słowa kluczowe:** automatyczne rozpoznawanie mowy, algorytmy genetyczne, tor akustyczny, selekcja cech.

**Keywords:** automatic speaker recognition, genetic algorithms, acoustic path, feature selection.

### Wstęp

Dotychczasowe badania w obszarze systemów automatycznego rozpoznawania mówców – ASR (ang. *Automatic Speaker Recognition*) – realizowane były przez autorów w oparciu o bazy głosów nagrywane z wykorzystaniem jednakowych torów akustycznych. Stanowi to znaczące ograniczenie zastosowań ASR, utrudniające rozpoznawanie zarchiwizowanych nagrań głosowych zarejestrowanych na zróżnicowanym sprzęcie. Wspomniany problem jest tematem licznych badań naukowych na całym świecie [8, 11, 12], jednakże większość prowadzonych przez naukowców eksperymentów dotyczy wykazania różnic w skuteczności poprawnej identyfikacji mówców z wykorzystaniem różnych urządzeń, przy jednoczesnym zachowaniu zbieżności sprzętów nagrywających sygnał uczący i testowy lub wykazania wpływu różnych zewnętrznych zakłóceń na pracę systemu ASR. Natomiast w prezentowanym artykule, autorzy dokonali sprawdzenia działania oraz optymalizacji systemu w przypadku braku zbieżności urządzeń nagrywających wykorzystywanych do rejestracji sygnału uczącego i testowego. W tym celu dokonano ponownej selekcji cech z wykorzystaniem zmodyfikowanego względem wcześniejszych badań algorytmu genetycznego [6]. Funkcja przystosowania selekcjonuje cechy, które są silnie związane z reprezentowanym głosem i jednocześnie słabo zależne od toru akustycznego oraz od siebie wzajemnie.

### Baza głosów

Zastosowana baza głosów powstała w ramach pracy dyplomowej jednego z autorów. Podczas rejestracji nagrań zadaniem lektorów było dziesięciokrotne odczytanie jednominutowego tekstu o zróżnicowanym zabarwieniu emocjonalnym, za każdym razem rejestrowanego przez inne urządzenie nagrywające. W wypowiedzianym przez lektorów tekście można wyróżnić: dialog, tekst oficjalny, tekst śpiewany, tekst przygnębiający oraz tekst żartobliwy, co pozwala na minimalizację wpływu intonacji oraz treści wypowiedzi na skuteczność rozpoznawania mówców. Baza zawiera nagrania 50 lektorów (18 kobiet i 32 mężczyzn w wieku od 11 do 75 roku życia), nagrywanych z wykorzystaniem 10 odmiennych torów akustycznych, dając w sumie 500 rejestracji głosowych. Nagrania próbkowane były z szybkością 8 kS/s, przy rejestracji jednokanałowej z 16-bitową rozdzielczością amplitudową.

Zastosowana szybkość próbkowania uwarunkowana jest próbą sprawdzenia i dostosowania systemu ASR do pracy z niskojakościowymi nagraniami mówców często spotykanymi w kryminalistyce. W dalszych badaniach autorzy zamierzają sprawdzić skuteczność poprawnej identyfikacji mówców przy zastosowaniu ustandaryzowanych metod kodowania mowy stosowanych np. w telefonii [3].

Do rejestracji wykorzystano 3 mikrofony dynamiczne, mikrofon pojemnościowy, kierunkowy mikrofon biurkowy, mikrofon wewnętrzny komputera klasy *laptop*, 2 telefony komórkowe, kamerę internetową, dyktafon oraz zewnętrzny interfejs audio, który pośredniczył w podłączeniu części mikrofonów do komputera gromadzącego nagrania. Podczas realizacji sesji nagranych w pomieszczeniu panowała względna cisza, jednakże nie było ono w żaden sposób dodatkowo wyciszane. Ponadto w zarejestrowanych nagraniach słyszalny jest szum komputera oraz zakłócenia wprowadzane przez niektóre niskojakościowe urządzenia nagrywające.

### Wstępne przetwarzanie

W ramach przetwarzania wstępnego realizowana jest filtracja, normalizacja, segmentacja oraz trzyetapowa selekcja ramek, co pozwala na znaczące zmniejszenie wpływu konfiguracji sprzętu nagrywającego na skuteczność poprawnej identyfikacji mówcy. Filtracja sygnału w zaimplementowanym systemie ASR odbywa się przy pomocy filtru pasmowo-przepustowego o skończonej odpowiedzi impulsowej [4]. W następnej kolejności realizowana jest normalizacja sygnału względem jego najwyższej wartości, pozwalająca na prawidłowe wyskalowanie z zachowaniem odpowiednich relacji energetycznych wewnątrz nagrania. Następnie realizowany jest proces segmentacji z wykorzystaniem okna Hamminga.

Ostatnim elementem jest trój etapowa selekcja ramek. W etapie pierwszym usuwane są fragmenty cisy z nagrania poprzez wyeliminowanie ramek niespełniających wyznaczonego empirycznie kryterium mocy w ramce. W tym etapie odrzucane są wyłącznie długotrwałe ciche fragmenty sygnału tak, aby zbytnio nie dezintegrować nagrań głosowych. W drugim etapie za pomocą metody autokorelacyjnej dokonywany jest wybór tzw. *ramek dźwięcznych*, w których znajduje się większość informacji o tonie kraniowym mówcy. Próg dźwięczności ramki określany jest na podstawie wartości drugiego maksimum funkcji autokorelacji. W ostatnim etapie następuje

eliminacja ramek nadmiernie zasumionych, za pomocą oceny różnicy wartości częstotliwości podstawowej wyznaczonej metodą cepstralną i autokorelacyjną [4].

Wszystkie opisane elementy procesu wstępnego przetwarzania sygnału mowy zostały poddane wielokryterialnej optymalizacji, co pozwoliło na ściśle określenie zoptymalizowanych wartości poszczególnych stałych parametrów systemu [4].

### Generacja cech dystynktywnych

Kluczowym etapem przetwarzania systemów biometrycznych jest generacja cech dystynktywnych, gdyż powstałe w niej błędy są nierekompensowalne w dalszych etapach działania systemu. Sygnał mowy w postaci czasowej charakteryzuje się wysoką redundancją, dlatego konieczna jest jego transformacja do postaci częstotliwościowej umożliwiającej dalszą skuteczną analizę. W prezentowanej realizacji systemu wykorzystano tzw. *ważone cechy cepstralne* (1) oraz *cechy melcepstralne* (2) zdefiniowane następująco [1]:

$$(1) \quad CC(i) = \sum_{k=-dp}^{dp} \mathcal{F}^{-1} \left\{ \log \left( \left| \mathcal{F} \{s(t)\} \right| \right) \right\}$$

$$(2) \quad MFCC(i) = DCT \left\{ \log \left( \left| \mathcal{F} \{s(t)\} \cdot Mel(i)_{bank} \right| \right) \right\}$$

gdzie:  $i$  – indeks cechy ( $i = 1, \dots, 30$ ),  $dp$  – zakres sumowania w obrębie  $i$ -tego maksimum cepstrum,  $s(t)$  – ramka rzeczywistego sygnału mowy,  $Mel_{bank}$  – bank filtrów melowych.

W celu wyznaczenia ważonych cech cepstralnych należy widmo amplitudowe poddać logarytmowaniu, zamieniając związek multiplikatywny pomiędzy składową wolnozmienną wynikającą z artykulacji mowy i szybkozmienną charakteryzującą ton kraniowy, na związek addytywny. W dalszej kolejności przetworzony sygnał transformowany jest do dziedzinie tzw. pseudoczasu cepstralnego za pomocą odwrotnej transformacji Fouriera. W tej dziedzinie informacje o treści wypowiedzi usytuowane są w pobliżu zera, natomiast impulsy związane z tonem kraniowym w okolicach okresu tonu kraniowego i powtarzają co ten okres. W ostatnim etapie wyznaczania ważonych cech cepstralnych następuje wyznaczanie cepstrum przez bank filtrów sumacyjnych ułożonych wokół kolejnych maksimum cepstrum.

Podczas wyznaczania cech melcepstralnych widmo amplitudowe sygnału wymnażane jest przez filtry sumacyjne, których położenie i szerokości wynikają ze skali melowej. Zabieg ten pozwala na odwzorowanie nierównomierniej rozdzielczości częstotliwościowej ludzkiego słuchu. W następnej kolejności otrzymane cechy są logarytmowane oraz poddawane transformacji kosinusowej w celu ich dekorelacji.

### Selekcja cech

Prowadzone na świecie badania dowodzą, że nie zawsze maksymalna liczba deskryptorów pozwala na uzyskanie najlepszych wyników rozpoznania [2, 9, 13], gdyż niektóre z nich mogą nieść wyłącznie szum pomiarowy, który znacznie pogarsza wyniki identyfikacji lub też mogą być ze sobą silnie skorelowane, co może prowadzić do dominacji tych cech nad pozostałymi. W opisywanym systemie ASR do selekcji cech wykorzystano algorytm genetyczny. Selekcja optymalnego podzbioru cech za pomocą algorytmów genetycznych zaliczana jest do globalnych metod optymalizacji, ponieważ charakteryzuje się równoległym przetwarzaniem wielu punktów przestrzeni rozwiązań bez tendencji wpadania w lokalne ekstrema

optymalizacji [9]. Ponadto algorytm genetyczny uwzględnia synergię cech pozwalając uzyskać optymalny zbiór cech, jest jednak czasochłonny.

W pierwszym etapie działania algorytmu genetycznego następuje obliczenie informacji wzajemnej pomiędzy cechami w analizowanym zbiorze cech oraz pomiędzy cechami, a wektorem przynależności klasowej tych cech. W przypadku prezentowanej realizacji systemu ASR zastosowano dwa niezależne wektory przynależności klasowej. Pierwszy z nich określa przynależność wektorów cech do danego mówcy, natomiast drugi określa przynależność wektorów cech do danego urządzenia nagrywającego, na którym dokonywano rejestracji konkretnego nagrania. Dzięki informacji wzajemnej możliwe jest określenie stopnia przewidywania wektora przynależności klasowej lub innej cechy dystynktywnej wyłącznie na podstawie znajomości cechy rozpatrywanej. W celu wyznaczenia informacji wzajemnej w pierwszej kolejności należy określić stopień nieuporządkowania poszczególnych cech zwany entropią. Ponadto wyznaczana jest entropia łączna dwóch analizowanych zmiennych, która dla zmiennych zależnych jest zawsze mniejsza od sumy entropii tych dwóch zmiennych, natomiast dla zmiennych całkowicie niezależnych jest równa ich sumie. W ostatecznej postaci informacja wzajemna stanowi sumę entropii poszczególnych zmiennych pomniejszoną o ich entropię łączną. W rezultacie prowadzonych obliczeń otrzymywana jest macierz informacji wzajemnych pomiędzy wszystkimi cechami oraz pomiędzy cechami, a wektorami przynależności klasowej (zarówno przynależności do poszczególnych mówców, jak i przynależności do poszczególnych torów akustycznych). Obliczanie informacji wzajemnych jest czasochłonne i realizacja obliczeń dla każdej możliwej kombinacji cech w dużym zbiorze jest niewykonalna w racjonalnym czasie. Prezentowany algorytm genetyczny charakteryzuje się tym, że za każdym razem do oceny przystosowania populacji osobników wykorzystuje bazowe informacje wzajemne maksymalnego zbioru cech [7]. Początkowa populacja osobników wykorzystywana w algorytmie genetycznym jest generowana w sposób pseudolosowy, tworząc wektory losowo wybranych cech spośród ich maksymalnego zbioru. Długości wektorów cech (chromosomów) są jednakowe.

W następnej kolejności każdy  $k$ -ty zestaw cech (chromosom) w populacji jest oceniany na podstawie funkcji przystosowania  $S(k)$ :

$$(3) \quad S(k) = V(k)_M - V(k)_U - P(k)$$

gdzie:  $V(k)_M$  – sumaryczna informacja wzajemna pomiędzy cechami  $k$ -tego zestawu cech a wektorem przynależności klasowej do poszczególnych mówców,  $V(k)_U$  – sumaryczna informacja wzajemna pomiędzy cechami  $k$ -tego zestawu cech, a wektorem przynależności klasowej do poszczególnych urządzeń nagrywających,  $V(k)_M$  – sumaryczna informacja wzajemna pomiędzy poszczególnymi cechami występującymi w  $k$ -tym zestawie cech.

Selekcję przeprowadza się w sposób pseudolosowy jednakże faworyzując osobniki lepiej przystosowane. Wylosowane w efekcie chromosomy podlegają wielopunktowemu krzyżowaniu, tj. każda cecha dla nowego chromosomu dobierana jest pseudolosowo z dwóch cech występujących na danym indeksie w wektorach cech tzw. chromosomów rodzicielskich. Dzieje się tak, aż do momentu otrzymania nowego wektora o tej samej długości, co wektory rodzicielskie. Gdy cechy się duplikują następuje mutacja wektora poprzez podmianę powtórzonej cechy na inną niewystępującą w wektorze. W konsekwencji

opisanych operacji genetycznych tworzona jest nowa populacja, która również, jak poprzednia generacja, jest sprawdzana funkcją przystosowania. Procesy te są powtarzane, aż do spełnienia warunków stopu, czyli braku wystarczającego wzrostu funkcji przystosowania lub do osiągnięcia maksymalnej liczby pokoleń [6]. Ostatecznym wynikiem selekcji jest wektor cech z ostatniego pokolenia, wykazujący najwyższy poziom przystosowania.

### Klasyfikator GMM-UBM

W opisywanym systemie, jako klasyfikator wykorzystano liniową kombinację rozkładów Gaussa (*GMM* – ang. *Gaussian Mixture Models*), które korzystając z danych uczących pozwalają na generację oszczędnych pamięciowo modeli głosów [4]. Wartości początkowe podczas tworzenia modeli dobierane są pseudolosowo lub w sposób zdeterminowany przez algorytm *GMM-UBM* [3, 4]. Metoda polega na utworzeniu uniwersalnego modelu głosu (*UBM* – ang. *Universal Background Model*) na podstawie głosów wielu osób i zastosowaniu jego parametrów, tj. wartości oczekiwanych, macierzy kowariancji oraz wag rozkładów, jako danych inicjalizujących podczas tworzenia poszczególnych modeli głosów. Dzięki temu możliwe jest szybsze dopasowanie się generowanego modelu głosu mówcy do danych uczących. Ponadto takie podejście wpływa korzystnie na skuteczność rozpoznawania, co zostało udowodnione przez autorów we wcześniejszych badaniach [4].

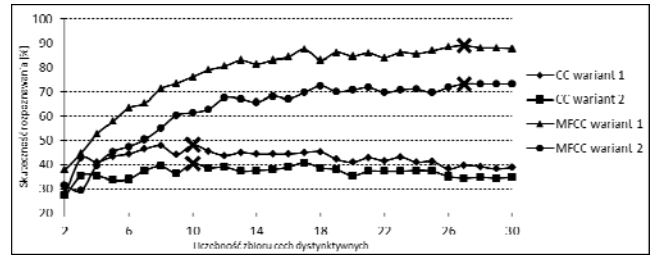
### Wyniki eksperymentów

Do optymalizacji systemu wykorzystano nagrania 40 z 50 dostępnych mówców, wykorzystując jedynie 8 z 10 dostępnych urządzeń. Pozostałe niezależne nagrania zostały użyte w ostatnim etapie sprawdzającym skuteczność systemu. Czas trwania segmentów uczących wynosił 25 s, natomiast segmentów testowych 5 s. Wykorzystywane segmenty testowe pochodziły z niezależnych od segmentów uczących fragmentów nagrania. Ponadto testy zostały przeprowadzone w dwóch wariantach. **Wariant pierwszy** wykorzystuje odcinki testowe i modele głosów pochodzące ze wszystkich 8 urządzeń dla 40 mówców. **Wariant drugi** natomiast nie uwzględnia wyników identyfikacji w przypadku zbieżności urządzenia wykorzystanego do tworzenia modeli głosów z urządzeniem wykorzystywanym do nagrania odcinka testowego.

### Selekcja cech dystyngtywnych

W trakcie badań zauważono zależność kolejnych cech cepstralnych od cechy pierwszej, dlatego przeprowadzono eksperyment normując cechy cepstralne (cechy o numerach od 2 do 30) względem cechy pierwszej. Pozwoliło to uzyskać 10% wzrost skuteczności identyfikacji mówcy za pomocą wyłącznie ważonych cech cepstralnych. Dokonano również próby analogicznej normalizacji cech melcepstralnych, jednakże w tym przypadku prowadziło to do wyraźnego pogorszenia wyników.

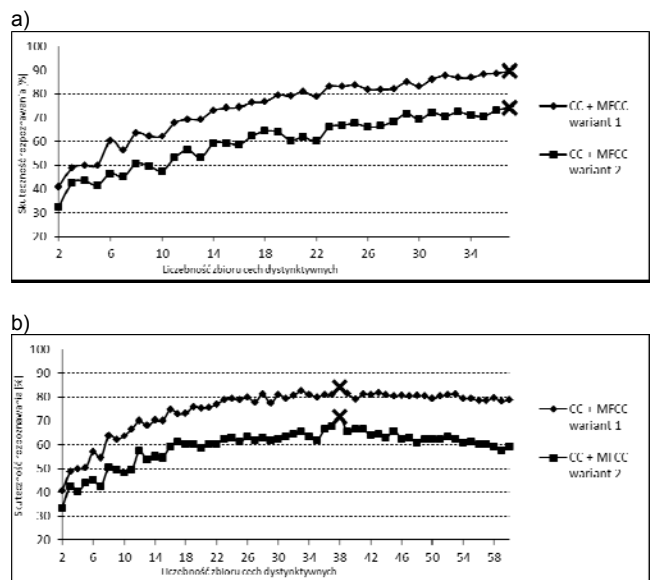
Zastosowany do selekcji cech algorytm genetyczny umożliwia wytypowanie optymalnych zbiorów cech o określonej długości. Dlatego niezbędne jest ustalenie optymalnej długości wektora cech. Na rysunku 1 przedstawione są wyniki skuteczności rozpoznawania mówców w funkcji długości wektora cech dla unormowanych ważonych cech cepstralnych (CC) oraz cech melcepstralnych (MFCC). Jak widać cechy melcepstralne znacznie lepiej reprezentują głos mówcy nawet w przypadku braku zbieżności urządzeń nagrywających.



Rys. 1. Wyniki poprawnej identyfikacji mówcy względem długości wektora cech wyselekcjonowanego przez algorytm genetyczny

### Fuzja cech

W kolejnym badaniu dokonano fuzji wyselekcjonowanych unormowanych cech cepstralnych i cech melcepstralnych z poprzedniego eksperymentu, a następnie przeprowadzono ich kolejną selekcję z wykorzystaniem algorytmu genetycznego (Rys. 2 a). Dokonano również selekcji cech na zbiorze wszystkich unormowanych ważonych cech cepstralnych i cech melcepstralnych (Rys. 2 b).

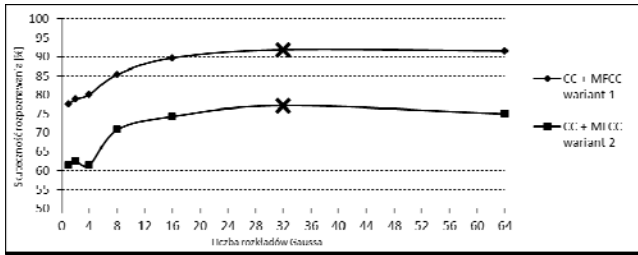


Rys. 2. Selekcja fuzji unormowanych ważonych cech cepstralnych oraz cech melcepstralnych

Jak widać wcześniejsza preselekcja cech w podzbiorach i ich fuzja pozwala na uzyskanie wyższej skuteczności identyfikacji niż selekcja cech w pełnym zbiorze. Wynika to z zastosowanej funkcji przystosowania, która faworyzuje zbiory cech o niskim poziomie wzajemnej korelacji, które niekoniecznie muszą być wysoce reprezentatywne dla poszczególnych głosów.

### Optymalizacja liczby rozkładów Gaussa stosowanych w modelach głosów

Wykorzystując otrzymany z poprzedniego eksperymentu zbiór 37 cech dystyngtywnych dokonano optymalizacji liczby rozkładów Gaussa wykorzystywanych podczas tworzenia modeli mieszanin Gaussowskich reprezentujących głosy poszczególnych mówców. Podobnie jak poprzednio, testowanie odbywało się w dwóch opisanych wcześniej wariantach.



Rys. 3. Optymalizacja liczby rozkładów Gaussa wykorzystywanych podczas tworzenia GMMs

### Test skuteczności systemu automatycznego rozpoznawania mowy

Niniejszy test został przeprowadzony dla zoptymalizowanych w poprzednich eksperymentach parametrów systemu, tj. wektora cech (10 CC + 27 MFCC) oraz 32 rozkładów Gaussa wykorzystywanych podczas tworzenia GMMs. Podczas testu wykorzystano głosy 10 mówców nagrywanych za pomocą dwóch niezależnych torach akustycznych (kamera internetowa i dyktafon). Zarówno głosy, jak i urządzenia nie brały udziału we wcześniejszej optymalizacji. Test przeprowadzono w dwóch wariantach. Pierwszy z nich zezwala na identyfikację mówcy w przypadku zbieżności urządzeń nagrywających wykorzystywanych do utworzenia sygnału uczącego i testowego, natomiast drugi bierze pod uwagę wyłącznie modele powstałe z nagrań pochodzących z niezależnego urządzenia względem sygnału testowego. Skuteczność poprawnej identyfikacji w przypadku zbieżności urządzeń wynosi 100%, zaś w przypadku przeciwnym zaledwie 30%. Ponadto dokonano próby inicjalizacji procesu tworzenia modeli głosów za pomocą uniwersalnego modelu głosu [11], co pozwoliło poprawić uzyskiwane wyniki dając w rezultacie 35% poprawnych identyfikacji mówców.

W celu sprawdzenia celowości stosowania opisanej w artykule selekcji cech przetestowano 22-elementowy zbiór cech dotychczas wykorzystywany przez autorów [6]. Dla przypadku zbieżności torów akustycznych, również otrzymano 100% skuteczność poprawnych identyfikacji, jednak dla braku zbieżności urządzeń wynik ten wyniósł zaledwie 15%.

### Podsumowanie

Przeprowadzone badania pozwoliły wyselekcjonować zbiór cech dystynktywnych o zmniejszonej względem wcześniejszych badań wrażliwości na zróżnicowanie toru akustycznego, dając tym samym wyższą skuteczność identyfikacji. Otrzymane względnie niskie wyniki świadczą o istotnym wpływie urządzeń nagrywających na właściwe działanie systemu ASR. Jednakże wybrane do testu urządzenia znacząco różniły się od urządzeń biorących

udział w procesie optymalizacji, co dodatkowo mogło utrudnić właściwą identyfikację. W dalszych badaniach autorzy zamierzają sprawdzić możliwość zastosowania kaskady klasyfikatorów integrującej wyniki niezależnych klasyfikacji cech CC i MFCC.

### LITERATURA

- [1] Dobrowolski A. P., Majda E., Cepstral analysis in the speakers recognition systems, *15th IEEE SPA Conference*, (2011), 85-90
- [2] Harrag A., Saigaa D., Boukharouba K., Drif M., Bouchelaghem A., GA-based feature subset selection: Application to Arabic speaker recognition system, *11th International Conference on Hybrid Intelligent Systems (HIS)*, IEEE Press, (2011), 382-387
- [3] Janicki, A., Staroszczyk, T., Klasyfikacja mówców oparta na modelowaniu GMM-UBM dla mowy o różnej jakości, *Krajowe Sympozjum Telekomunikacji i Teleinformatyki*, (2011)
- [4] Kamiński K., Majda E., Dobrowolski A. P., Automatic speaker recognition using Gaussian Mixture Models, *17th IEEE SPA Conference*, (2013), 220-225
- [5] Kamiński K., Dobrowolski A. P., E. Majda, Ocena funkcjonalności systemu rozpoznawania mówcy dla zdegradowanej jakości sygnału głosowego, *Przegląd Elektrotechniczny*, 90 (2014) nr 8, 164-167
- [6] Kamiński K., Dobrowolski A. P., Majda E., Selekcja cech osobniczych sygnału mowy z wykorzystaniem algorytmów genetycznych, *XXVIII Międzynarodowa Konferencja Naukowo-Techniczna „Ekomilitaris’14”*, (2014)
- [7] Ludwig O., Nunes U., Novel Maximum-Margin Training Algorithms for Supervised Neural Networks, *IEEE Transactions on Neural Networks*, 21 (2010) nr 6, 972-984
- [8] Ming J., Hazen T., Glass J., Reynolds D. A., Robust speaker recognition in noisy conditions, *IEEE Trans. Audio, Speech, Lang. Process.*, 15 (2007) nr 5, 1711-1723
- [9] Osowski S., Metody i narzędzia eksploracji danych, *BTC*, (2013)
- [10] Reynolds D. A., Quatieri T. F., Dunn R. B., Speaker Verification Using Adapted Gaussian Mixture Models, *Digital Signal Processing*, (2000) nr 10, 19-4
- [11] Shan Y. X., Liu J. Robust Speaker Recognition in Cross-Channel Condition, *2nd International Congress on Image and Signal Processing*, (2009), 1-5
- [12] Sturim D. E., Campbell W. M., et al., Robust Speaker Recognition with Cross-Channel Data: MIT/LL Results on the 2006 NIST SRE Auxiliary Microphone Task, *ICASSP*, (2007)
- [13] Zamalloa M., Bordel G., Rodriguez J.L., Penagarikano M., Feature selection based on genetic algorithms for speaker recognition, *IEEE Odyssey*, 1 (2006), 1-8

**Autorzy:** mgr inż. Kamil Kamiński, dr hab. inż. Andrzej P. Dobrowolski, dr inż. Ewelina Majda-Zdanczewicz, inż. Daniel Posiadała, Wojskowa Akademia Techniczna, Wydział Elektroniki, ul. gen. S. Kaliskiego 2, 00-908 Warszawa,  
E-mail: [kamil.kaminski@wat.edu.pl](mailto:kamil.kaminski@wat.edu.pl), [adobrowolski@wat.edu.pl](mailto:adobrowolski@wat.edu.pl), [emajda@wat.edu.pl](mailto:emajda@wat.edu.pl), [daniel.posiada1@gmail.com](mailto:daniel.posiada1@gmail.com)