



Rodzaj pracy: magisterska

Dyplomant: mgr inż. Michał BOJSZA

Promotor: prof. dr hab. inż. Andrzej DOBROWOLSKI

OCENA WPŁYWU SZUMÓW I ZAKŁÓCEŃ NA SKUTECZNOŚĆ AUTOMATYCZNEGO SYSTEMU ROZPOZNAWANIA MÓWCY W WARUNKACH RÓŻNORODNOŚCI JĘZYKOWEJ

Wprowadzenie

Identyfikacja mówcy na podstawie głosu jest niezbędna wszędzie tam, gdzie niemożliwe jest użycie innego rodzaju weryfikacji, np. podczas korzystania z łączności radiowej czy telefonicznej. Stosowanie rękawiczek lub maseczek ochronnych również komplikuje weryfikację tożsamości przez odcisk palca lub rozpoznanie twarzy. W procesie rozpoznania głosu cały czas należy brać pod uwagę wiele czynników zakłócających, które wpływają na pracę urządzeń i systemów zajmujących się identyfikacją i weryfikacją mówcy. Autor skupia się na tym elemencie i ocenia w jakim stopniu potencjalne zakłócenia generowane przez środowisko i szумы wpływają na pracę systemu ARMiA – Automatyczne Rozpoznanie Mowy i Autoryzacja.

Baza Głosów i Baza Zakłócających Dźwięków Tła

Do badania wpływu szumów i zakłóceń na skuteczność automatycznego systemu rozpoznawania mówcy zostały stworzone dwie bazy danych. Pierwsza to wielojęzyczna baza głosów składająca się z 62 nagrań głosów (w maseczkach ochronnych i bez nich) należących do 31 użytkowników różniących się narodowością, płcią i wiekiem. Druga to baza dźwięków zakłócających, która zawiera różnego rodzaju szумы (szum biały i różowy generowane przez program MATLAB) i dźwięki tła, typu dworzec, lotnisko, start samolotu, gwar uliczny, wodospad, odgłosy zwierząt i natury, hałasy maszyn, pojazdów i in. (80 nagrań).

Oprogramowanie Interfejsu Użytkownika

Do oceny wpływu szumów i zakłóceń na działanie systemu rozpoznawania mówcy, stworzono interfejs użytkownika, weryfikujący skuteczność systemu. Interfejs ten składa się z 6 bloków. Pierwszy to „Rodzaj Zakłóceń” umożliwiający wybór typu szumu. Drugi to „Stopień Zaszumienia” odpowiedzialny za poziom zakłóceń jakie mają zostać dodane do sygnału mowy osoby badanej poprzez dobór wartości współczynnika SNR (stosunek mocy sygnału do mocy szumu). Trzeci blok, to „Wczytanie Danych” złożony z dwóch elementów tj. „wczytaj dane treningowe” dla zdefiniowania bazy głosów oraz „wczytania danych testowych” dla określenia grupy osób, które mają zostać poddane identyfikacji. Czwarty blok, to „Weryfikacja Systemu Identyfikacji Mówcy”, gdzie inicjuje się cały proces wstępnego przetwarzania sygnałów, ekstrakcji cech, tworzenia bazy modeli głosów, kwalifikacji i decyzji dla potrzeb identyfikacji mówcy. Piąty blok, to „Etap Przetwarzania Danych” złożony z dwóch wyświetlaczy, które zobrazowują progres przetwarzania danych. Jeden reprezentuje „Etap Ucznienia” i pokazuje progres w tworzeniu bazy modeli głosów w części treningowej. Drugi wyświetlacz reprezentuje „Etap Testowania” i pokazuje ilu mówców z grupy testowej zostało do tej pory zidentyfikowanych. Szósty blok, to „Wyniki”, przedstawiający na wyświetlaczu całość wyników tj. liczbę osób prawidłowo zidentyfikowanych przez system w odniesieniu do ogólnej liczby osób badanych, oraz skuteczność ogólną systemu wyrażoną w procentach.

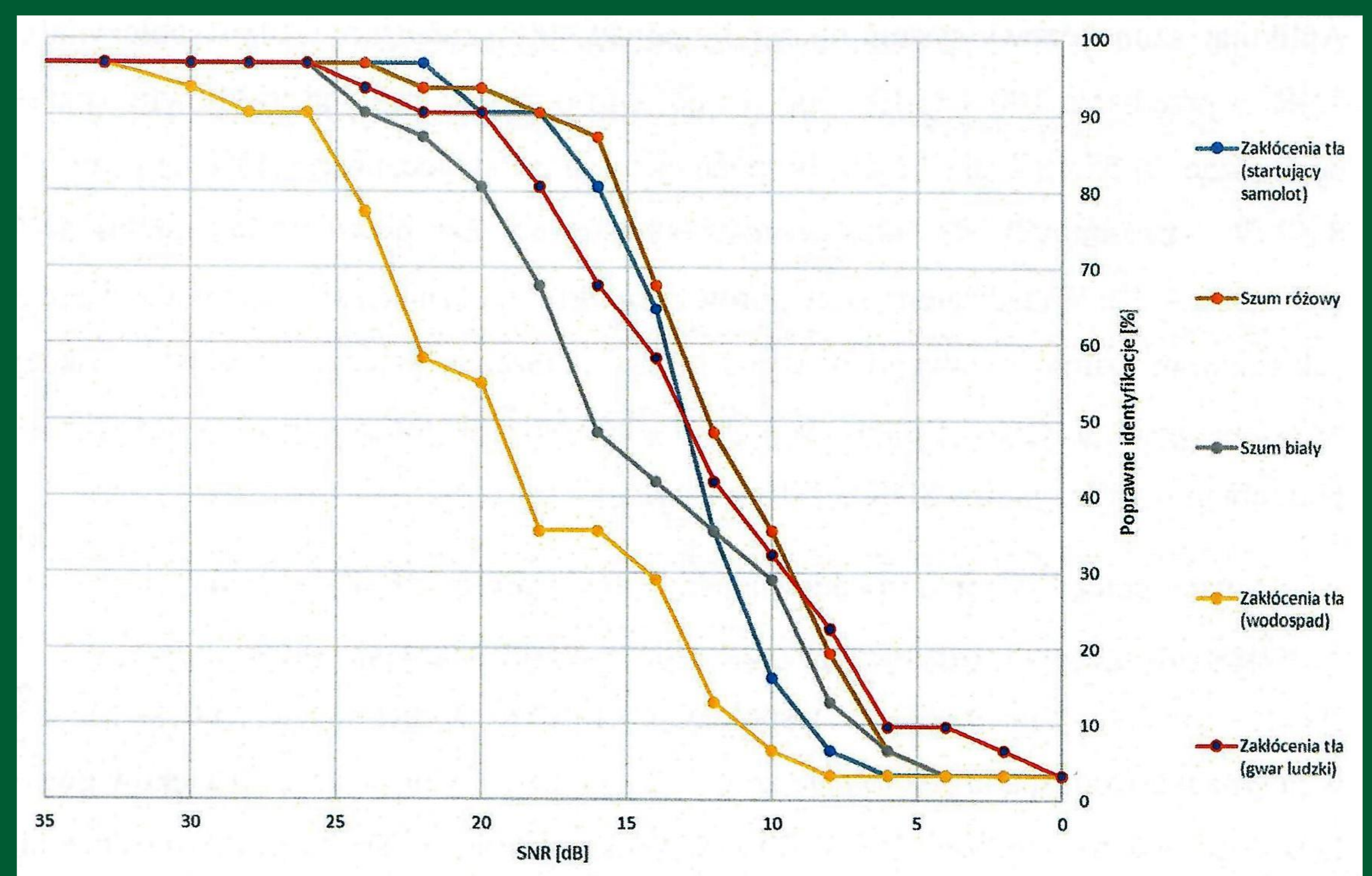
Weryfikacja cepstralnego systemu rozpoznawania mówcy

W oparciu o przygotowaną bazę głosów i bazę dźwięków zakłócających oraz interfejs użytkownika przeprowadzono badania oparte o cztery, kilkuwariantowe scenariusze, weryfikujące skuteczność systemu rozpoznawania mówcy oraz wpływu szumów i zakłóceń na pracę systemu ARMiA.

Wyniki Badań

Badania wykazały, że system ARMiA odznacza się wysoką odpornością na zaszumienie i sygnały zakłócające. System, niezależnie od zaaplikowanego zakłócenia na sygnał mowy testowej, jest w stanie utrzymać skuteczność identyfikowania mówcy na poziomie ponad 90% dla współczynnika SNR w przedziale 100 – 26 dB, a w przypadku szumów niskoczęstotliwościowych w przedziale 100 – 18 dB. Najbardziej negatywny wpływ na skuteczność systemu posiadają zakłócenia wysokotonowe tj. szum biały czy dźwięki zakłócające tła wodospadu. System jest niezwykle odporny na wszelkiego rodzaju szумы nisko tonowe tj. szum różowy. Charakterystyka skuteczności systemu ARMiA jest silnie zależna od rodzaju dźwięku użytego jako zakłócający. Dźwięki wysoko częstotliwościowe zbliżone do szumu białego (np. wodospad) mają bardziej negatywny wpływ na rozpoznanie mówcy niż te nisko częstotliwościowe.

Wyniki badań pokazały, że maseczki ochronne nieznacznie zmniejszają skuteczność systemu ARMiA oraz pogarszają efektywność systemu, gdy baza modeli głosów jest tworzona o mówców z maseczkami. Wynika to z faktu, że system ARMiA jest systemem opartym na biometrykach fizjologicznych, a nie behawioralnych. Ważne więc jest, aby tworzenie bazy modeli głosów odbywało się w jak najkorzystniejszych warunkach. Poniższy rysunek przedstawia procent poprawnych identyfikacji przy zmianie współczynnika SNR [dB] dla scenariusza 4.



Wnioski

Badanie potwierdziło, że systemy analizujące mowę są bardzo skuteczne w procesie identyfikacji użytkowników i z całą pewnością odznaczają się dużą odpornością na zakłócenia i szумы. Posiadają wiele atrybutów biometrii wymaganych dla dobrej pracy systemu biometrycznego. Ważnym elementem pracy był autorski interfejs użytkownika, wytwarzający zakłócenia i dodający je do testowanego sygnału mowy oraz decydujący o poprawności identyfikacji. Interfejs operował na wielonarodowej bazie mówców zgromadzonych na potrzeby tej pracy