



Wojskowa
Akademia
Techniczna

Wydział
Elektroniki



STRESZCZENIE

System automatycznego rozpoznawania mówcy oparty na analizie cepstralnej sygnału mowy i modelach mieszanin gaussowskich

Autor: por. mgr inż. Kamil Kamiński

Promotor: prof. dr hab. inż. Andrzej P. Dobrowolski

Celem niniejszej rozprawy było opracowanie systemu automatycznego rozpoznawania mówcy pozwalającego na skuteczną identyfikację głosów w utrudnionych warunkach rejestracji, zbliżonych do warunków transmisji telefonicznej. W rozprawie przedstawiono proces projektowania systemu oraz wyniki przeprowadzonych optymalizacji i testów z zastosowaniem baz głosów powstałych w Wojskowej Akademii Technicznej oraz baz komercyjnych.

Rozprawa składa się z czterech logicznych części prezentujących kolejne etapy realizowanych prac. W części pierwszej (rozdział 1) przedstawiona została tematyka biometrii głosu i determinujący ją sposób funkcjonowania ludzkiego narządu głosowego. Sformułowane zostały również potencjalne obszary zastosowań systemu rozpoznawania mówcy, a także stojące przed nim wyzwania eksploatacyjne wpływające na bezpieczeństwo realizowanych uwierzytelnień.

Kolejna część (rozdziały 2-6) przedstawia architekturę utworzonego systemu rozpoznawania mówcy opartego na analizie cepstralnej sygnału mowy i modelach mieszanin gaussowskich. W sposób chronologiczny omówiono wszystkie etapy przetwarzania sygnału mowy od akwizycji sygnału poprzez jego wstępne przetwarzanie, generację i selekcję cech osobniczych, utworzenie oszczędnych pamięciowo modeli głosów, aż do uzyskania decyzji o wyniku przeprowadzonej identyfikacji. Prezentowany system, w toku realizowanych badań, ulegał licznym modyfikacjom i optymalizacjom.

Trzecią część pracy (rozdziały 7-8) stanowią wyniki badań przeprowadzonych podczas projektowania i optymalizacji systemu. Pokazano w niej również wyniki identyfikacji mówców osiągnięte przez opisywany system w jego sfinalizowanej postaci, przeprowadzone w rozmaitych wariantach testowania. Zaprezentowano także aplikację implementującą opracowany system rozpoznawania mówcy w środowisku *Matlab*. Aplikacja ta umożliwia dodawanie nowych modeli głosów oraz wykonywanie testów identyfikacji mówcy.

Ostatnią część pracy stanowi podsumowanie, w którym na podstawie wcześniej zaprezentowanych badań udowodniona została teza niniejszej rozprawy, w brzmieniu: ***możliwe jest takie przetworzenie sygnału mowy, w wyniku którego można utworzyć oszczędne pamięciowo modele głosów pozwalające na identyfikację mówcy z wysoką skutecznością.*** W części tej postawiono także przewidywane dalsze obszary pracy rozwojowych prezentowanego systemu rozpoznawania mówcy.

Słowa kluczowe: rozpoznawanie mówcy, analiza cepstralna, modele mieszanin gaussowskich, algorytmy genetyczne, selekcja cech dystynktywnych

Kamil Kamiński
Warszawa, 21 czerwca 2018 r.