

Streszczenie rozprawy doktorskiej

Analiza przestrzennych modeli akustycznych głosek dentalizowanych w diagnostyce sygnatyizmu

Author: mgr inż. Michał Kręcichwost

Promotor: dr hab. inż. Paweł Badura, prof. PŚ

Promotor pomocniczy: dr inż. Joanna Czajkowska

Niniejsza rozprawa doktorska podejmuje temat zastosowania komputerowego przetwarzania sygnału mowy do wspomagania diagnostyki logopedycznej. Prowadzone badania dotyczą analizy i klasyfikacji normatywnych i nienormatywnych realizacji fonemów dentalizowanych przez dzieci w wieku przedszkolnym. Badania oparto na opracowanej bazie mowy dzieci 5 i 6-letnich, zawierającej dane akustyczne oraz opis logopedyczny przygotowany przez specjalistów terapii mowy.

W ramach pracy zaprojektowano, wykonano i przetestowano urządzenie do akwizycji 15-kanałowego sygnału mowy. Opracowano metodykę przetwarzania wielokanałowej reprezentacji sygnału, opartą na uczeniu głębokim. W skład metodyki wchodziła synchronizacja oraz przetwarzanie wstępne sygnałów dedykowane dla głosek dentalizowanych. Agregację danych wykonano przy użyciu technik przetwarzania sygnału przestrzennego. Wyodrębnianie cech bazowych oparte zostało na dekompozycji widma sygnału z wykorzystaniem dedykowanego zespołu filtrów. Zaproponowano własny zakres częstotliwości, odnosząc się do zakresów znanych z literatury i wykazując zasadność rozszerzenia pasma w kierunku wyższych częstotliwości i linearyzacji skali zespołu filtrów. Dane zebrano do trójwymiarowej struktury obrazów barwnych, nazwanej wolumenem akustycznym. Dla celów analizy zaproponowano autorską, trójgałęziową architekturę konwulucyjnej sieci neuronowej. Rolą każdej z gałęzi było poszukiwanie wzorców w oddzielnych wymiarach: czasowym, częstotliwościowym oraz mieszanym. Zależnie od sposobu interpretacji danych z 15 kanałów akustycznych opracowano i przebadano różne modele. Przeprowadzono analizę statystyczną zależności pomiędzy cechami akustycznymi i artykulacyjnymi w mowie normatywnej i patologicznej głosek /s/ i /ʃ/ dla wariantu 5-kanałowego z uwzględnieniem kanałów bocznych.

Przeprowadzone eksperymenty z wykorzystaniem różnych wariantów proponowanej struktury sieci pozwalają stwierdzić, iż wykorzystanie technik uczenia głębokiego umożliwia rozpoznawanie różnych realizacji analizowanych głosek. Sieci konwulucyjne poddano dodatkowo analizie wrażliwościowej w celu zbadania wpływu ważniejszych parametrów i ustawień na działanie modeli. Wyniki badań eksperymentalnych potwierdzają słuszność zastosowania zaproponowanych cech akustycznych, pasma i skali zespołu filtrów trójkątnych, konfiguracji i procedur przetwarzania sygnałów oraz modeli głębokiego uczenia w rozwiązaniu problemu detekcji różnych realizacji badanych sybilantów.

Słowa kluczowe: komputerowe wspomaganie diagnostyki logopedycznej, sygnatyizm, sygnał mowy, modele akustyczne, konwulucyjne sieci neuronowe