

Roman Figura, Zdzisław Szymański
(Społeczna Wyższa Szkoła Przedsiębiorczości i Zarządzania w Łodzi)

Klasyfikacja dźwięków języka polskiego przy zmianie tempa mowy

Mowa jest dźwiękową formą języka naturalnego, który w tradycyjnym ujęciu jest kodem służącym do porozumiewania się pomiędzy ludźmi. Rozwój cywilizacji oraz nauk cybernetycznych spowodował nowe możliwości wykorzystania języka w branży telekomunikacyjnej, bankowej i administracji oraz wielu innych. Jednak mimo ogromnej mocy obliczeniowej, komputery wciąż mają problemy z rozpoznawaniem mowy. Dotyczy on naturalnej złożoności sygnału mowy, podobieństwa fonetycznego, problemów segmentacji na podstawowe jednostki językowe oraz dopasowania czasowego. Analizie poddajemy najczęściej fonemy lub wyrazy na podstawie określonych cech akustycznych lub fonetycznych.

Celem pracy jest opracowanie algorytmu zmiany tempa mowy przy zachowaniu jej struktury oraz trwałości. Ponieważ niemożliwym jest zebranie takiej ilości nagrań, która zawierałaby wystarczająco dużo realizacji wszystkich dźwięków, w referacie przedstawiono koncepcję wyróżnienia kilku klas elementów i przyporządkowaniu znalezionych w sygnale dźwięków do jednej z nich. Przeprowadzone przez nas badania nad zależnością długości dyftongów języka polskiego od zmiany tempa mówienia wykazują, że różne dźwięki przy zmianie tempa mowy mogą zachować taką samą wewnętrzną strukturę. Z podobnym podejściem spotkaliśmy się w pracach innych badaczy analizujących inne języki np. język angielski, język ukraiński. Oni również w swoich opracowaniach wskazują, że dźwięki jednej grupy zachowują się podobnie. Uwzględniając sugestie autorytetów w tej dziedzinie oraz wyniki przeprowadzonych do tej pory badań opracowaliśmy nowy podział fonemów języka polskiego dla zadań zmiany tempa który jest przedstawiony w naszym referacie.