



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI



SAMORZĄD WOJEWÓDZTWA
WIELKOPOLSKIEGO
WOJEWÓDZKI URZĄD PRACY
W POZNANIU

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Radosław Weychan **Politechnika Poznańska**

Stypendysta projektu pt. „Wsparcie stypendialne dla doktorantów na kierunkach uznanych za strategiczne z punktu widzenia rozwoju Wielkopolski”, Poddziałanie 8.2.2 Programu Operacyjnego Kapitał Ludzki

Rozpoznawanie mowy na podstawie transkodowanej mowy do interfejsów człowiek-maszyna

Rozpoznawanie mowy jest interesującym zagadnieniem w interfejsach człowiek-maszyna (ang. HMI, Human Machine Interfaces). Wykorzystywane jest przede wszystkim w systemach kontroli dostępu bezpośredniego do obiektów czy też maszyn lub elektroniki użytkowej, jednakże ze względu na coraz większy dostęp i wykorzystanie usług mobilnych pojawia się zapotrzebowanie na identyfikację zdalną. Takowa może wspomóc dostęp do usług bankowości, obsługę klientów w centrach usług telekomunikacyjnych, wspomóc identyfikację osób nadużywających m.in. telefony alarmowe. Jednakże ze względu na niższą jakość transmitowanego oraz kodowanego sygnału skuteczność identyfikacji jest mniejsza. Współczesne systemy identyfikacji dedykowane są najczęściej dla komputerów stacjonarnych lub urządzeń mobilnych z dostępem do Internetu. W odniesieniu do autonomicznych systemów wbudowanych, konieczne jest dobranie innej metodologii badań, która skupia się nie tylko na skuteczności, ale także na wymaganiach stawianych przez warstwę sprzętową, która musi zapewnić szybkie przetwarzanie oraz niski pobór energii. System taki ponadto musi spełniać rynkowe wymagania dotyczące ceny.

Celem przygotowywanej rozprawy doktorskiej jest opracowanie autonomicznego systemu automatycznego rozpoznawania mowy z sygnałów niskiej jakości celem poprawy skuteczności funkcjonowania platform ze zdalnym dostępem. W nawiązaniu przedstawionych problemów, autor proponuje implementację na stałoprzecinkowym procesorze sygnałowym, przy czym konieczne jest przeprowadzenie badań mających na celu odpowiedni dobór algorytmów ze względu na większy błąd obliczeń niż w przypadku platform zmiennoprzecinkowych.

Dodatkowo, w przypadku identyfikacji mówcy z rozmów telefonicznych, niezbędne jest opracowanie systemu segmentacji mówców, który umożliwiłby identyfikację tylko jednej ze stron konwersacji.

Przygotowywana rozprawa doktorska zawierać będzie przegląd aktualnie wykorzystywanych algorytmów rozpoznawania mówcy oraz ich analizę w odniesieniu do podejmowanego problemu. Zaprezentowane zostaną wyniki przeprowadzonych badań z wykorzystaniem anglo- i polskojęzycznych baz mówców, dotyczące rozpoznawania mówcy z krótkich wypowiedzi jakości telefonicznej (GSM). Przeanalizowane zostaną także metody poprawy skuteczności. Zaproponowana zostanie również dedykowana implementacja algorytmów dla systemu wbudowanego opartego o stałoprzecinkowy procesor sygnałowy oraz badania związane z doбором parametrów implementacji cechującej się najmniejszym błędem obliczeń. W odniesieniu do problemu segmentacji mówców, przedstawiona zostanie koncepcja autonomicznego systemu wbudowanego realizującego to zadanie w oparciu o algorytmy zapewniające niską latencję oraz niewrażliwe na konwersję z dziedziny analogowej do cyfrowej.

Przedstawione w rozprawie doktorskiej rozwiązania mogą przyczynić się do zwiększenia zainteresowaniem nowoczesnych systemów identyfikacji w wielu sektorach gospodarki obecnych w Wielkopolsce.