

Lingwistyka matematyczna w Katedrze Elektroniki AGH

Lingwistyka jaka jest każdy widzi

Lingwistyka matematyczna:

- identyfikacja rozmówcy
- przetwarzanie języka naturalnego
- przetwarzanie mowy
- rozpoznawanie mowy
- synteza mowy

Katedra Elektroniki

Zespół Przetwarzania Sygnałów

Akademia Górniczo-Hutnicza im. S. Staszica
w Krakowie

Wydział Informatyki, Elektroniki i Telekomunikacji

prof. Mariusz Ziółko

dr inż. Jakub Gałka - adiunkci - dr inż. Bartosz Ziółko

mgr inż. Dawid Skurzok - asystenci - mgr inż. Mariusz Mąsior

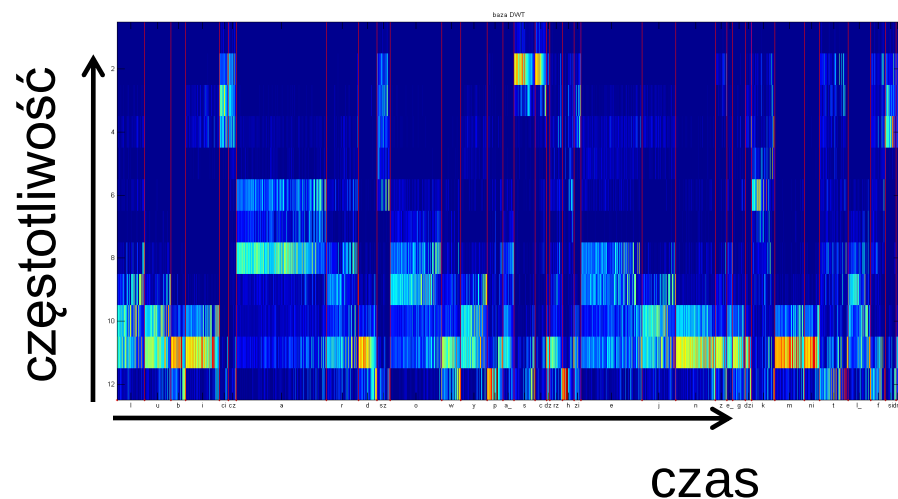
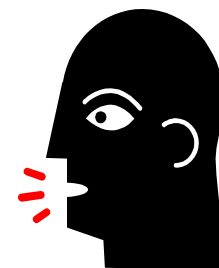
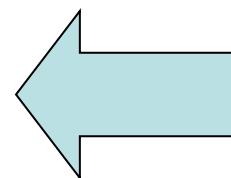
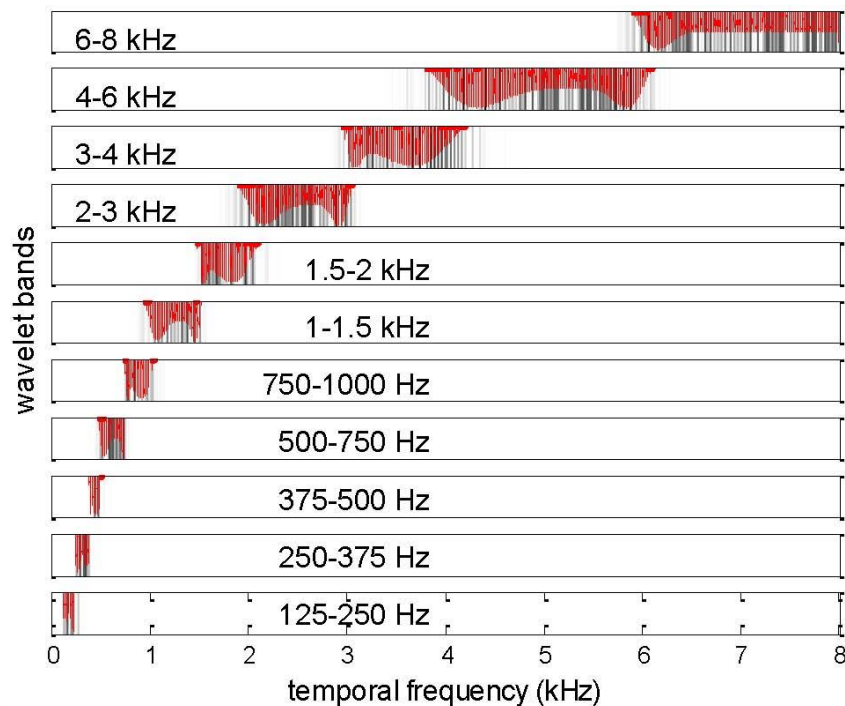
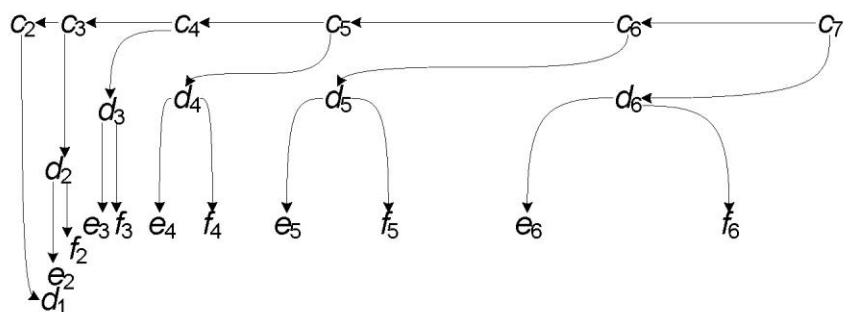
mgr inż. Tomasz Jadczyk - doktoranci - mgr inż. Magdalena Igras

mgr inż. Rafał Samborski - doktoranci - mgr inż. Jan Wicijowski

+ pracownicy i studenci pomagający realizować granty +

<http://www.dsp.agh.edu.pl>
<http://rozpoznawaniemowy.blogspot.com/>

Psychoakustyczna falkowa analiza mowy



Zbiory tekstów języka polskiego

Źródło	MBajty	Mil. słów	Różnych słów	Różnych dwójek	Różnych trójek
Rzeczpospolita	879	104	856 349	18 115 373	43 414 592
Wikipedia	754	97	2 623 358	31 139 080	61 865 543
Literatura	490	68	1 151 043	23 830 490	50 794 854
Transkrypcje	325	32	381 166	6 848 729	16 283 781
Literatura 2	6500	949	6 162 530	153 152 158	441 284 743
Literatura 3	285	181	1 229 331	36 297 382	93 751 340

W Literatura 2 :

- Słowa występujące więcej niż 10 razy to 951 844
- Dwójki słów występujące więcej niż 10 razy to 6 426 198
- Trójki słów występujące więcej niż 10 razy to 5 166 537

„Aplikacje technologii mowy w systemach bezpieczeństwa publicznego”

Realizowany w okresie od 8.IX.2008 do
7.IX.2011

POLSKA PLATFORMA
BEZPIECZEŃSTWA WEWNĘTRZNEGO



Wyniki projektu

Sąd Apelacyjny w Krakowie

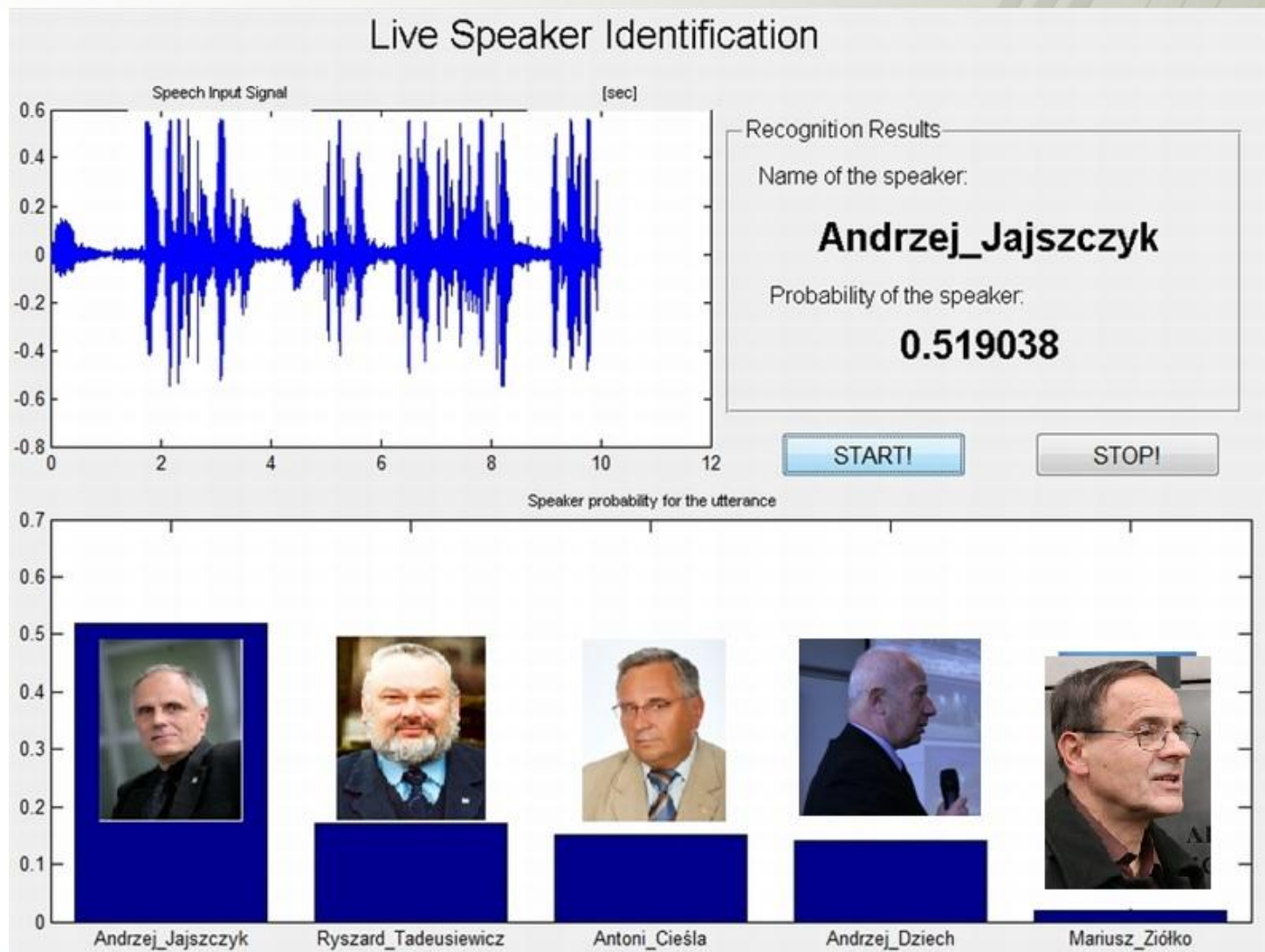
Głosowe sterowanie programem do zarządzania bazą danych akt i dokumentów sądowych



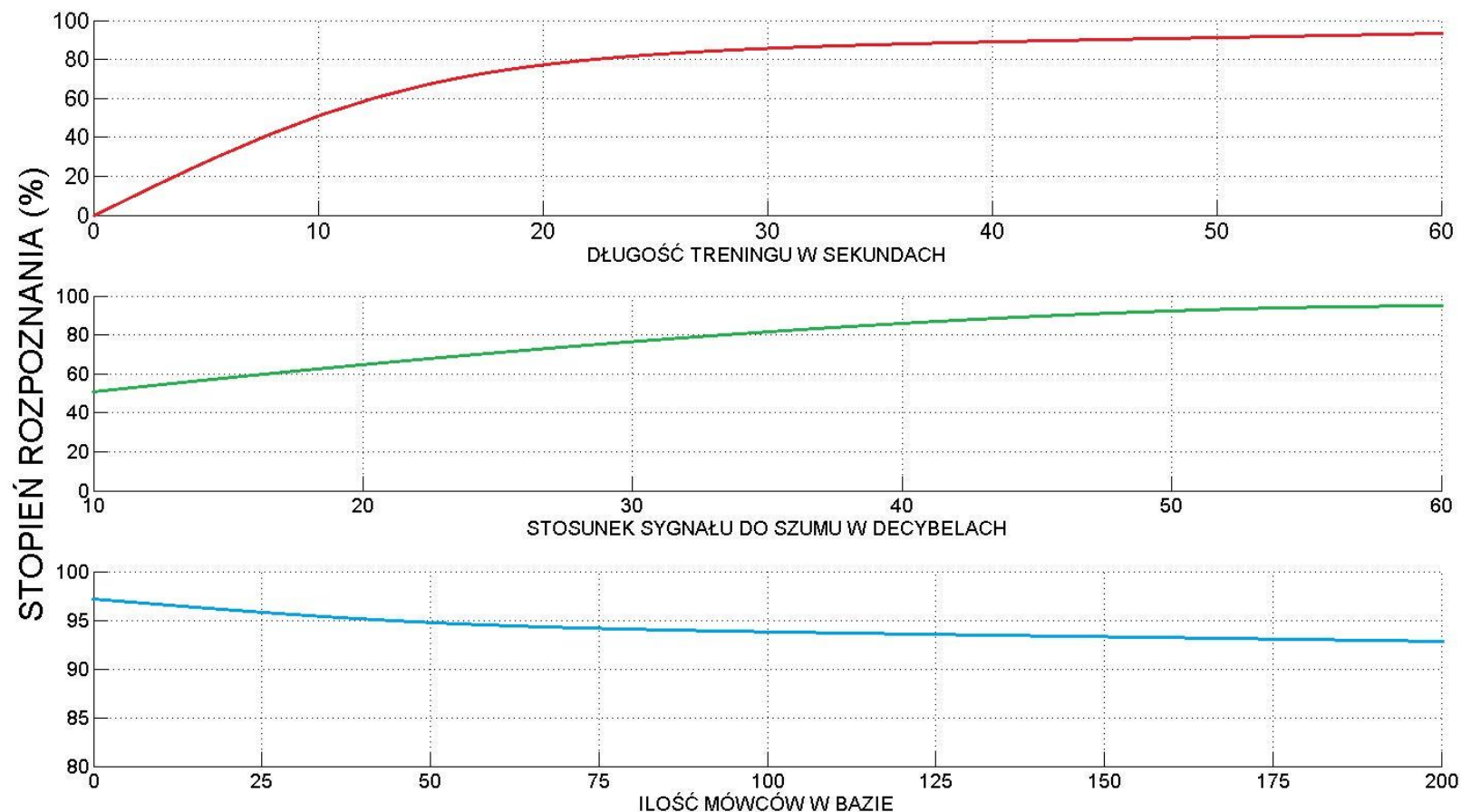
Policja

Głosowe sterowanie stanowiskiem koordynatora działań operacyjnych Komendy Wojewódzkiej Policji w Krakowie

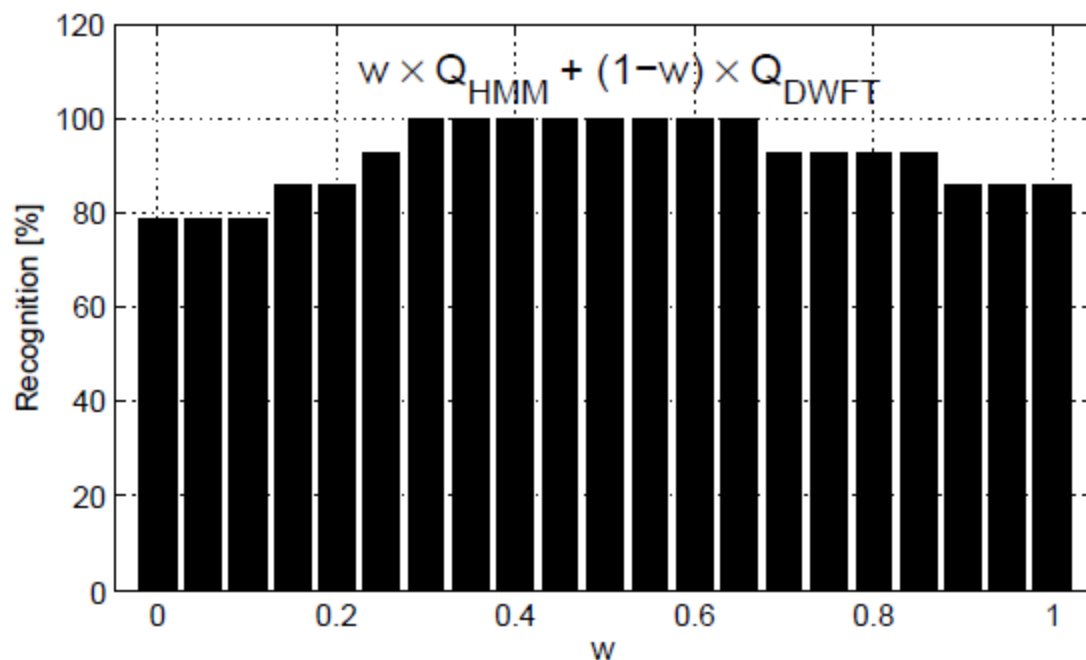
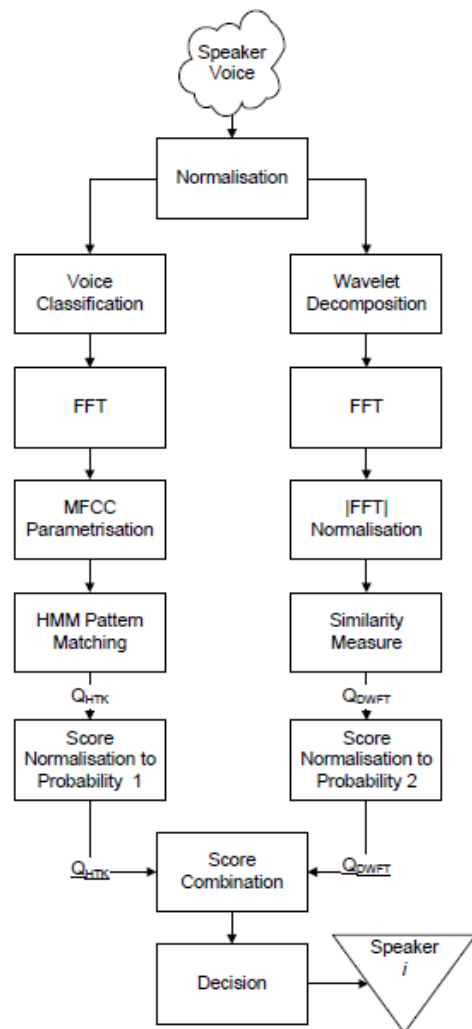
System rozpoznawania mówcy



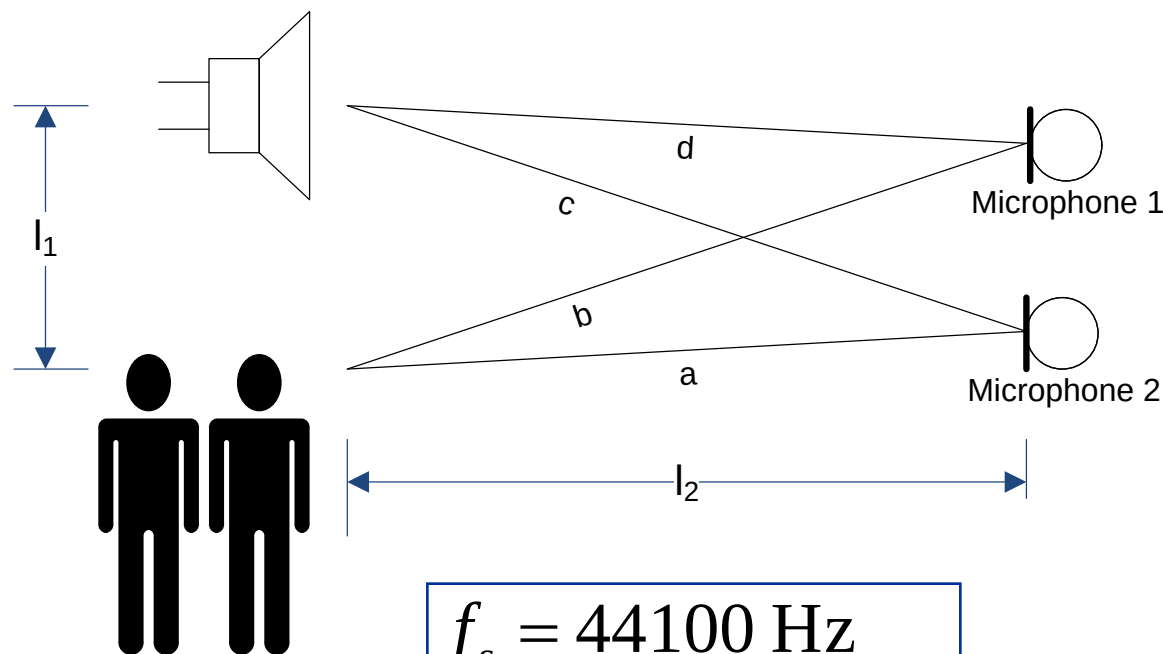
Opis jakości systemu – stopień rozpoznania



Dwie równoległe metody rozpoznawania



Systemy wielomikrofonowe do lokalizacji źródeł dźwięku



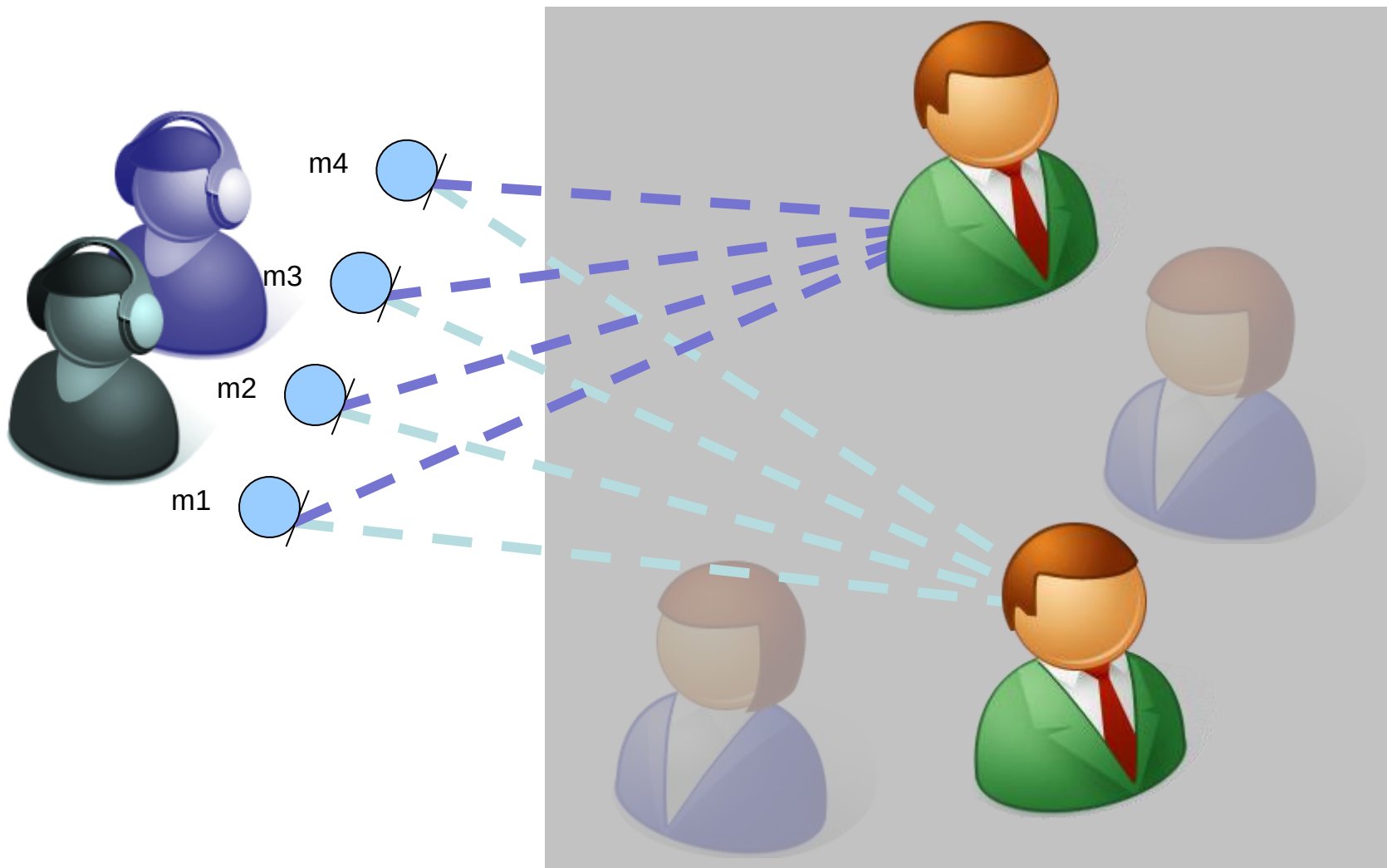
$$f_s = 44100 \text{ Hz}$$

$$\Delta t \approx 23 \mu\text{s}$$

$$v = 330 \text{ m/s}$$

$$\rho = v \Delta t \approx 7.5 \text{ mm}$$

Możliwość pracy kilku niezależnych operatorów



Q. D. Atkinson: "Phonetic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa". Science, 32, 346-348, 2011.

Cele projektu :

- opracowanie nowej metody ekstrakcji fonemów i określania ich charakterystyk,
- zbadanie różnic pomiędzy fonemami dla około 300 języków świata,
- badanie pochodzenia i ekspansji języków,
- stworzenie drzewa genealogicznego języków świata w oparciu o multijęzyczną analizę fonemów i metody stosowane w ewolucyjnej genetyce,
- zwiększenie efektywności rozpoznawania fonemów mowy polskiej.

Zarówno wyniki prac jak i zebrany materiał lingwistyczny będzie dostępny dla naukowców całego świata na stronie internetowej
<http://speechsamples.agh.edu.pl>

Raytracer audio i video (NCBiR, Innotech, 2012-2014)



m.in.

- Symulowanie wpływu akustyki pomieszczeń na dialogi.
- Modyfikacja brzmienia nagrania w celu rozróżnienia różnych trzecioplanowych postaci posługujących się nagraniami pochodzącymi od tego samego aktora.

System dialogowy człowiek-komputer oparty na łączach telefonicznych (NCBiR, Lider, 2012-2014)

- Korpus mowy z rozmów telefonicznych;
- Udoskonalenie parametryzacji i segmentacji (specyfika mowy zniekształconej przez połączenie telefoniczne);
- Klasyfikacja (dynamiczne sieci Bayesa, warunkowe pole losowe, głębokie sieci neuronowe);
- Detekcja przerw pomiędzy słowami (przy braku przerwy akustycznej – statystyki fonemów i logika rozmyta);
- Udoskonalenie modelu n-gramowego;
- Model semantyczny oparty o: macierze słowo-temat, Słownik Semantyczny i SłowoSieć;
- Podejmowanie decyzji i rozumienie wypowiedzi w oparciu o rozpoznaną mowę;
- Przyjazna i wciągająca forma interakcji z użytkownikiem.

Lingwistyczny warsztat do analizy i rozpoznawania mowy (NCN, Sonata, 2012-2014)



- Wykrywanie zdań wtrąconych, końców zdań, przerw itp.;
- Dostosowanie gramatyki Spejda do języka mówionego;
- Wykonanie probabilistycznej wersji Spejda;
- Analiza małych grup słownych w celu rozcinania siatki słów;
- Analiza metod użycia słownika fleksyjnego w modelowaniu mowy;
- Zbadanie powiązań między słownikiem kolokacji a modelem n-gramowym;
- Rozwinięcie modelu n-gramowego z wykorzystaniem informacji typu syntaktycznego.

Biometryczna weryfikacja i identyfikacja głosu (NCBiR, 2012-2014)

Stworzenie konkurencyjnego zarówno ekonomicznie jak i technologicznie produktu z dziedziny biometrii głosowej. Badania mają doprowadzić do opracowania zaawansowanego technologicznie systemu rozpoznawania i identyfikacji mówcy. Zakłada się, że opracowany system korzystać będzie z informacji akustycznej, bez uwzględnienia innych modalności.

Podstawowymi zastosowaniami systemu będą infolinie oraz telefoniczne centra obsługi klienta. Drugim zastosowaniem systemu będzie wykorzystanie weryfikacji głosu jako elektronicznego klucza w budynkach inteligentnych, zamkach elektronicznych, itp.

Algorytmy audiowizualnego przetwarzania sygnału mowy polskiej na potrzeby interfejsu człowiek- komputer (NCN, PRELUDIUM, 2012-2014)



Książkę opiniuję z przyjemnością, ponieważ mam miłą świadomość, że uczestniczę w przygotowaniu publikacji, która ma gwarantowane duże powodzenie u czytelników i która będzie miała realną szansę stać się godną następczynią mojej książki zatytułowanej „Sygnał mowy”, wydanej blisko ćwierć wieku temu. Publikacja stanowi bardzo wartościowe osiągnięcie twórcze i jest dawno oczekiwanym podręcznikiem akademickim dotyczącym dziedziny mającej bardzo duże znaczenie dla rozwoju współczesnych systemów naturalnej komunikacji między ludźmi i szeroko rozumianym światem techniki. Jest to niewątpliwie najlepsza książka o przetwarzaniu mowy polskiej, jaką kiedykolwiek napisano. Takiego podręcznika i takiej monografii zdecydowanie brakowało, dlatego jestem pewien, że będzie to kolejny sukces zarówno Autorów, jak i Wydawnictwa.

Ryszard Tadeusiewicz

Główną zaletą opracowania jest przedstawienie bardzo szerokiej gamy zagadnień w czytelnej i przejrzystej formie. Autorom udało się powiązać teorię dotyczącą metod analizy, przetwarzania i syntezy mowy z ich praktycznymi zastosowaniami. Przystępna forma prezentacji nawet trudniejszych zagadnień, powinna zainteresować szeroki krąg osób. Szczególną uwagę zwrócono na problemy charakterystyczne dla języka polskiego. Monografia jest bardzo wartościową pozycją, zawierającą oprócz szerokiego przeglądu aktualnego stanu wiedzy i uwzględniającego osiągnięcia Autorów, również wytyczne dla kierunków dalszych badań w dziedzinie przetwarzania mowy. Uważam, że zostanie ona przyjęta z uznaniem na polskim rynku wydawniczym.

Zygmunt Ciota

4E 61

70 6F 63 7A C485 74 6B 75

62 79 C582 6F

73 C582 6F 77 6F

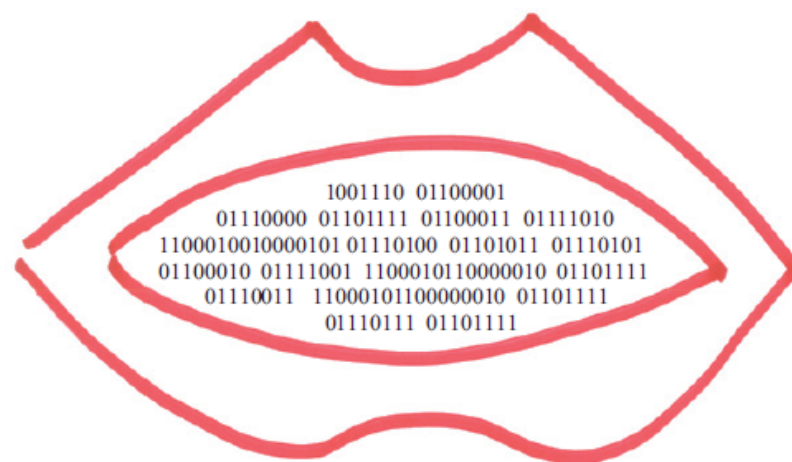
ISBN 978-83-7464-441-9

PRZETWARZANIE MOWY

Bartosz Ziółko, Mariusz Ziółko

PRZETWARZANIE MOWY

Bartosz Ziółko, Mariusz Ziółko



WYDAWNICTWA AGH

KRAKÓW 2011

Kamienie milowe w historii świata i ludzkości

"**Na początku było Słowo**" (J 1, 1). Tymi słowami Jan rozpoczyna swą Ewangelię, sięgając poza granice wyznaczające początek naszego czasu, aż do Boskiej wieczności.

Język jest jednym z najnowszych wytworów ewolucji - zdecydowana większość lingwistów z obozu ewolucjonistów umieszcza **pojawienie się pierwszego dosyć dobrze rozwiniętego języka** nie wcześniej niż 40 czy 50 tysięcy lat temu.

Gutenberg Johannes, właściwie **Johannes Gensfleisch zum Gutenberg** (ok. 1399-1468), niemiecki złotnik i mincerz. Uważany za **wynalazcę drukowania** za pomocą ruchomych czcionek, aparatu do ich odlewania i prasy drukarskiej.

Początki **Internetu** wiążą się z powstaniem sieci rozległej ARPANET i sięgają końca lat 60. XX wieku.

Proletariusze wszystkich krajów, łączcie się! –
fragment z Manifestu Komunistycznego Karola
Marksa i Fryderyka Engelsa

Lingwiści w całej Polsce, łączcie się! –
AGH



Dziękujemy za wzmożoną uwagę!