



## **DOBÓR PARAMETRÓW W METODZIE QE-ARM**

### **Selection of parameters in QE-ARM method**

**Stefan Brachmański**

Politechnika Wrocławska, Instytut Telekomunikacji i Akustyki  
[stefan.brachmanski@pwr.wroc.pl](mailto:stefan.brachmanski@pwr.wroc.pl)

#### **STRESZCZENIE**

Przedstawiono wyniki badań jakości transmisji mowy w sieciach wykorzystujących protokoły IP. Pomiary obiektywne wykonano metodą QE-ARM. Zbadano wpływ parametrów metody QE-ARM, takich jak rodzaj, długość i przesunięcie okna parametryzacji, klasyfikacja wstępna, liczba współczynników LPC lub Cepstrum na skuteczność rozpoznawania mowy. W celu zbadania zgodności ocen jakości mowy, wyniki otrzymane metodą QE-ARM (dla różnych parametrów) porównano z wyrazistością logatomową otrzymana metodami tradycyjną oraz z wyborem.

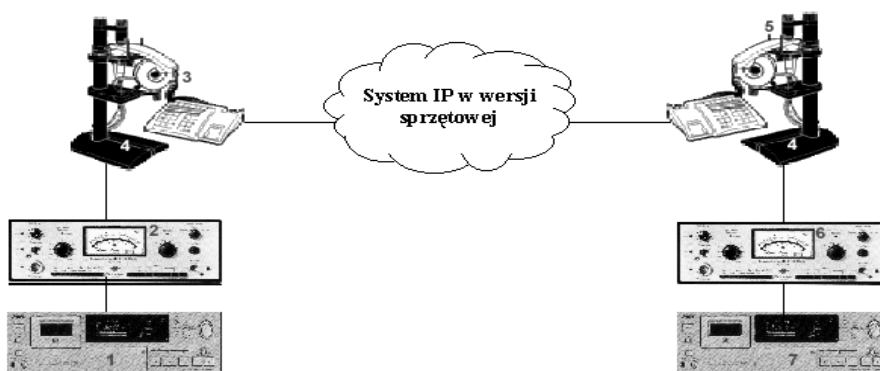
#### **1. WPROWADZENIE**

Ocena jakości mowy dokonana z wykorzystaniem metod obiektywnych powinna być zgodna z oceną uzyskaną w pomiarach subiektywnych. Spełnienie tego wymagania zależy między innymi od rodzaju obiektywnej metody wykorzystanej do oceny jakości mowy. W obiektywnej metodzie oceny jakości transmisji mowy wykorzystującej techniki automatycznego rozpoznawania mowy (QE-ARM) wpływ na zbieżność wyników pomiarów obiektywnych i subiektywnych ma dobór parametrów wykorzystywanych w rozpoznawaniu mowy. W aplikacji metody QE-ARM możliwa jest zmiana takich parametrów, jak [2]:

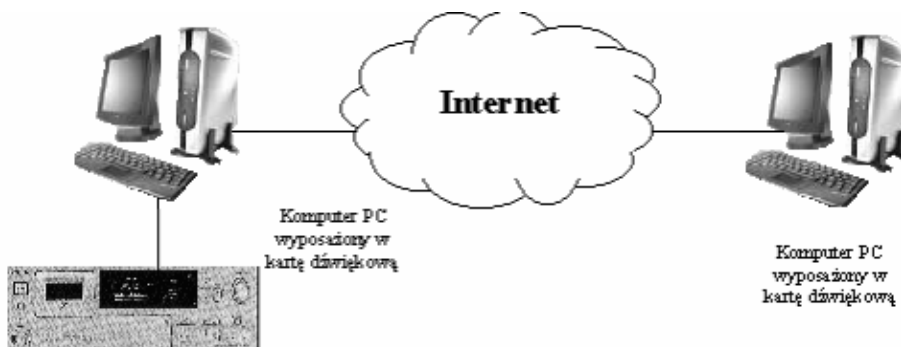
- \* typ okna czasowego (Hamming, Hanning, Blackmann, Blackmann-Harris, Bartlett, prostokąt, Gauss),
- \* długość okna czasowego ( $64 \div 4096$  próbek),
- \* przesunięcie okna czasowego ( $1 \div 50$  % długości okna),
- \* metodę parametryzacji (FFT, BF-FFT, LPC lub Cepstrum),
- \* liczbę parametrów FFT (analiza w pasmach oktawowych lub tercjowych),
- \* liczbę parametrów LPC (rzęd predykcji  $8 \div 32$ ),
- \* metrykę (miara Hamminga, Euklidesa, Camberra lub Itakury),
- \* liczbę klas czasowych (podział listy logatomowej na 2 - 5 grup czasowych).

## 2. MATERIAŁ TESTOWY

Materiałem testowym w pomiarach subiektywnych oraz obiektywnych były zrównoważone fonematycznie i strukturalnie listy logatomowe. Pomiarów wykonano w czasie nie rzeczywistym, tzn. przed wykonaniem pomiarów listy testowe zostały przesłane przez badane kanały VoIP i zarejestrowane na magnetofonie cyfrowym lub bezpośrednio na twardym dysku komputera wyposażonego w kartę przetwornika A/C. W ten sposób zapewnione zostały identyczne warunki transmisji dla wszystkich porównywanych metod oceny jakości mowy. Materiał testowy w sprzętowej realizacji sieci IP pozyskano w układach pokazanych na rys. 1 dla sprzętowej realizacji sieci IP i rys. 2 dla programowej realizacji sieci IP.



Rys. 1. Schemat blokowy układu do rejestracji materiału testowego - sprzętowa realizacja sieci IP. 1, 7 – magnetofon cyfrowy, 2 – wzmacniacz pomiarowy B&K 2606 , 3 – sztuczne usta B&K 4227, 4 - sztuczna głowa B&K 4602, 5 – sztuczne ucho B&K 4195, 6 - wzmacniacz pomiarowy B&K 2609.



Rys. 2. Schemat blokowy układu do rejestracji materiału testowego – programowa realizacja sieci IP

## 3. POMIARY

Pomiary jakości transmisji mowy wykonano dla rozwiązań software'owych realizując połączenie w sieci lokalnej oraz w sieci miejskiej opartej na telewizji kablowej (serwer AIPSA, serwer TP S.A., telewizja kablowa, modem telewizji kablowej). W drugim rozwiązaniu maksymalna szybkość do abonenta wynosi 768kb/s, a od abonenta 128 kb/s.

W ramach eksperymentu przebadano tory telefonii IP zrealizowane za pomocą oprogramowania: Microsoft NetMeeting (kodowanie CELP i G.723.1), ICQ, LabNet Phone, Paltalk, Buddyphone (przepływność 13 kb/s i 64kb/s) oraz Tlen. Natomiast wersją sprzętową była sieć korporacyjna grupy kapitałowej Howell S.A., oparta na rozwiązaniu i sprzęcie firmy CISCO. Dla tej sieci wykonano dwa rodzaje badań:

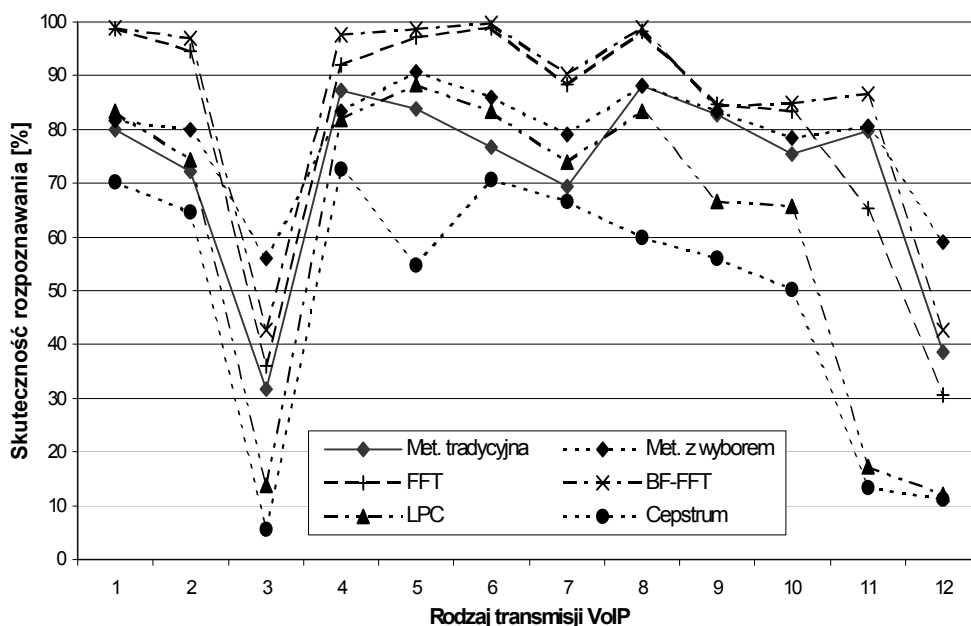
- transmisja w obrębie jednej centrali,
- transmisja z czterokrotną zmianą kompresji (Świebodzice-Poznań, Poznań-Świebodzice).  
Transmisja w tym połączeniu przebiegała następująco: od telefonu Cisco 64 kb/s, przez router Cisco 3600 - transkodowanie z 64 kb/s na 8.3 kb/s, transmisja po sieci WAN TP S.A. - 8.3 kb/s, router w oddziale - transkodowanie z 8.3 kb/s na 64 kb/s, router w oddziale - transkodowanie z 64 kb/s na 8.3 kb/s, transmisja po WAN TP S.A. - 8.3 kb/s, router Cisco 3600 - transkodowanie z 8.3 kb/s na 64 kb/s, do telefonu Cisco 64 kb/s.

Poszczególnym rozwiązaniom VoIP przypisano losowo numery od 1 do 12. Dla wszystkich rozwiązań VoIP wykonano subiektywne pomiary wyrazistości logatomowej metodami tradycyjną (zgodnie z Polską Normą PN-V-90002) [4] i z wyborem [1] oraz obiektywną metodą QE-ARM [2]. Zbadano wpływ następujących parametrów [3]:

- rodzaj parametryzacji – FFT, BF-FFT, LPC, Cepstrum,
- typ okna – Hamming, Hanning, Blackmann, Blackmann-Harris, Bartlett, prostokąt, Gauss,
- rozmiar okna – 256, 512, 1024, 2048,
- przesunięcie okna – 10%, 25%, 50%,

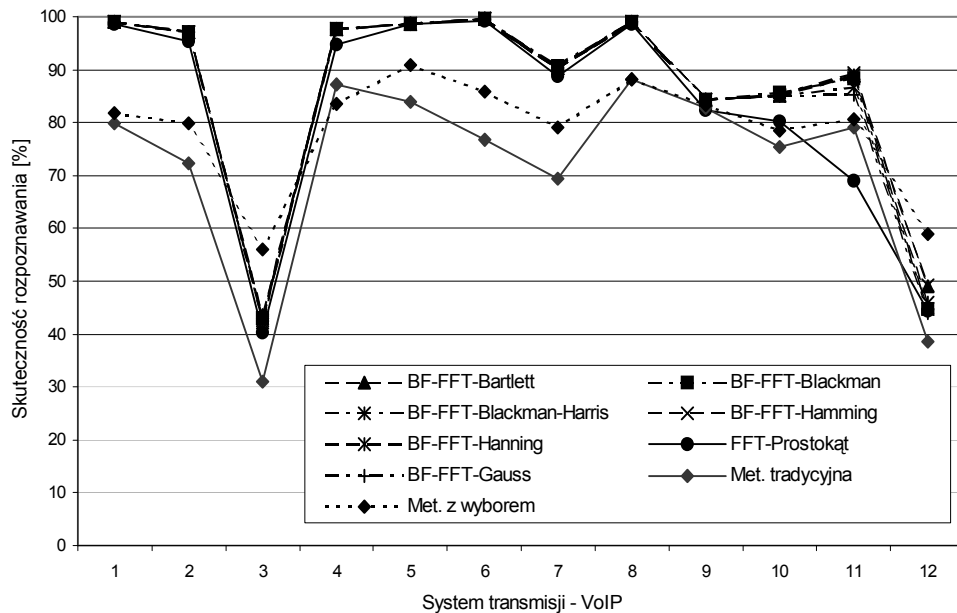
#### 4. WYNIKI

W pomiarach testowych w pierwszej kolejności zbadano wpływ miary odległości na skuteczność metody QE-ARM. Otrzymane wyniki przedstawiono na rys.3.



Rys. 3. Wpływ miary odległości na skuteczność metody QE-ARM.

Dla wszystkich rodzajów parametryzacji, przy ustalonej optymalnej mierze odległości, zbadano wpływ rodzaju okna. Przykładowe wyniki dla parametryzacji BF-FFT pokazano na rys.4. Ograniczono się do przedstawienia wyników tylko dla jednego rodzaju parametryzacji ze względu na czytelność wykresu.



Rys. 4. Wpływ rodzaju okna na skuteczność metody QE-ARM (parametryzacja BF-FFT).

## 5. PODSUMOWANIE

Analizując otrzymane wyniki można stwierdzić, że największy wpływ na skuteczność rozpoznawania ma rodzaj parametryzacji. Okazało się, że największą skuteczność uzyskano dla parametryzacji BF-FFT (transformata Fouriera w pasmach barkowych), przy czym najlepsza zbieżność wystąpiła dla parametryzacji LPC. Analizując wpływ rodzaju okna można zauważyć, że dobór okna ma zdecydowanie mniejsze znaczenie niż dobór metody parametryzacji.

## LITERATURA

1. BAŚCIUK K., BRACHMAŃSKI S., The automation of the subjective measurements of logatom intelligibility., The 102<sup>nd</sup> Convention AES, Munich 1997, Preprint 4407 (A5).
2. BRACHMAŃSKI S., Związek między wyrazistością logatomową, a obiektywną miarą QE-ARM dla mowy przesyłanej przez Internet, Mat. OSA'2004, Gdańsk, 2004.
3. WALCZAK T., Ocena jakości transmisji mowy w sieciach IP obiektywną metodą QE-ARM, Praca dyplomowa, Politechnika Wrocławska, 2003.
4. PN-V-90002 Cyfrowe łańcuchy telefoniczne. Wymagania i metoda pomiaru wyrazistości logatomowej.