

OPTIMALIZACJA METOD OBIEKTYWNEJ OCENY JAKOŚCI SYGNAŁÓW FONICZNYCH KODOWANYCH STRATNIE

Optimization for Objective Methods of Lossy Coded Audio Signals Evaluation

Piotr Kozłowski

Politechnika Wrocławska, Instytut Telekomunikacji i Akustyki, Zakład Akustyki
piotr.kozlowski@pwr.wroc.pl

STRESZCZENIE

Referat dotyczy metod obiektywnej oceny jakości sygnałów fonicznych kodowanych stratnie. W tekście zaprezentowano oprogramowanie, w którym zaimplementowano protokoły oceny obiektywnej. Rozwiązania przyjęte przy implementacji programowej umożliwiły regulację tych parametrów wdrożonych protokołów, które mają wpływ na ostateczne wyniki. W tekście zaprezentowano również wyniki badań subiektywnych. Rezultaty tych badań traktowane są jako referencyjne oceny wykorzystywane do strojenia metod obiektywnych. Przez strojenie rozumie się uzyskanie jak największej zgodności pomiędzy wynikami uzyskiwanymi w badaniach subiektywnych a rezultatami wykorzystania prezentowanego oprogramowania. Dodatkowo zaproponowano zmiany w protokołach oceny obiektywnej, mające na celu zbliżenie wskaźników obiektywnych do wyników subiektywnych. Przedstawiono rezultaty badań, ilustrujące wpływ wprowadzonych zmian na wyniki oceny obiektywnej.

1. KLASYCZNE METODY POMIAROWE

Klasyczne metody klasyfikowania jakości nie nadają się do oceny sygnałów zakodowanych stratnie przez systemy cyfrowe. Wynika to w pierwszej kolejności stąd, że złożone modele psychoakustyczne wykorzystywane przez kodeki są układami wysoce nieliniowymi. Klasyczne pomiary np. charakterystyki częstotliwościowej czy stosunku sygnał/szum SNR (Signal to Noise Ratio) mają sens w odniesieniu do systemów liniowych. Również wskaźniki nieliniowe, np. THD+N (Total Harmonic Distortion + Noise), nie uwzględniają efektu maskowania, jaki występuje w procesie percepcji sygnałów akustycznych. Nie bez znaczenia jest również fakt, że system nieliniowy należy pobudzać w sposób umożliwiający obserwację jego nieliniowego charakteru. Postulatu tego nie spełnia pobudzenie sinusoidalne, będące najprostszym sygnałem, jaki można poddać kodowaniu.

Inną, bardzo istotną, cechą kodeków jest zależność ich funkcji przenoszenia od czasu, choćby ze względu na wykorzystanie uśredniania na słowach zgromadzonych w buforze, które może prowadzić do pojawienia się słyszalnych artefaktów. Wynika stąd, że poprawna ocena kodeka może zostać przeprowadzona jedynie w naturalnych warunkach jego pracy, czyli przy pobudzeniu sygnałem mowy lub sygnałem muzycznym.

Metody subiektywne, jedyne w pełni akceptowane, wymagają spełnienia następujących wymogów:

- duża liczba odpowiednio dobranych słuchaczy,
- duża liczba powtórzeń,
- sesje odpowiednio rozłożone w czasie,
- odpowiednie pomieszczenia odsłuchowe.

Wymagania te sprawiają, że pomiary subiektywne są czasochłonne i drogie, niezależnie od zastosowanych paradygmatów oraz ilości porównywanych sygnałów.

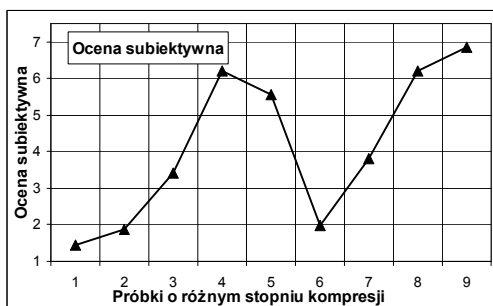
2. METODY OCENY OBIEKTYWNEJ

Struktura programu PQMS (Perceptual Quality Measurement System) prezentowanego w pracy została przedstawiona szczegółowo we wcześniejszych pracach [3], [8], [10], [11]. Dotychczas zaimplementowano następujące algorytmy obiektywnej oceny jakości sygnałów fonicznych:

- PAQM Perceptual Audio Quality Measure [1],
- PSQM Perceptual Speech Quality Measure [2],
- NMR Noise to Mask Ratio [6],
- PEAQ PErceived Audio Quality [6],
- PESQ PErcieved Evaluation of Speech Quality [7].

Przy opracowywaniu omawianego oprogramowania wykorzystano fakt, iż struktura wszystkich wymienionych powyżej algorytmów jest podobna. Wszystkie one symulują zachowanie ludzkiego kanału słuchowego. System pomiarowy został opracowany jako otwarta platforma umożliwiająca dołączanie kolejnych protokołów oceny w postaci wtyczek programowych. Wynikami poszczególnych algorytmów są wskaźniki oceny jakości będące różnicą pomiędzy wrażeniami słuchowymi powstałymi w trakcie „słuchania” badanej próbki skompresowanej oraz jej oryginału.

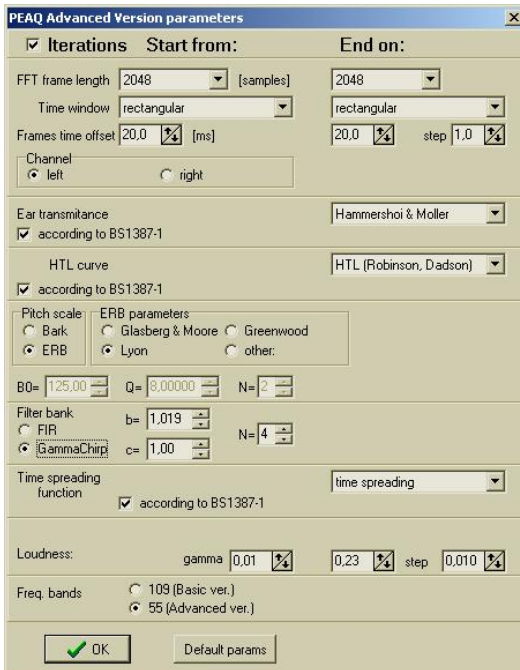
2. SPOSOBY OPTIMALIZOWANIA METOD OCENY OBIEKTYWNEJ



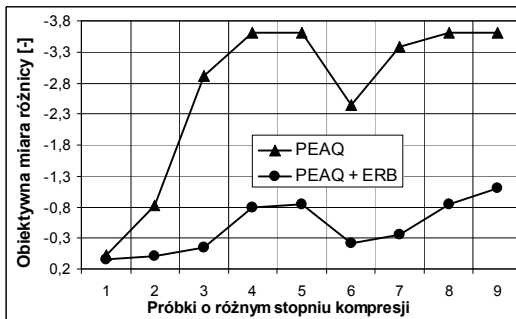
Rys. 1. Wyniki badań subiektywnych

Optymalizacja protokołów oceny obiektywnej została poprzedzona badaniami subiektywnymi. Szczegóły dotyczące metodologii odsłuchów oraz weryfikacji ich wyników przedstawiono we wcześniejszej pracy [10]. W zastosowanej skali ocena 1 oznacza, że słuchacz nie zauważał różnicy pomiędzy oryginałem a próbką ocenianą, natomiast ocena 7 – słuchacz uznał daną próbkę za najbardziej odbiegającą od oryginału. Oceny pośrednie słuchacze przyznawali zgodnie z własną oceną różnicy jakości pomiędzy

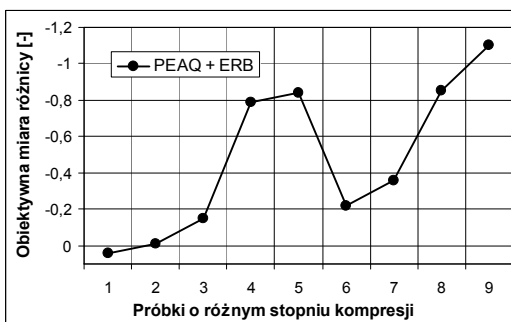
próbki. Wyniki badań subiektywnych traktowane są jako wskazania referencyjne podczas oceny rezultatów pracy protokołów obiektywnych. We wszystkich badaniach subiektywnych i obiektywnych wykorzystywane są te same próbki dźwiękowe: różne gatunki muzyki (klasyka, pop, jazz, rock, elektronika) oraz mowa żeńska i męska. Z uwagi na bardzo ograniczoną objętość tekstu przedstawiono tutaj tylko wyniki uzyskane dla muzyki pop i protokołu PEAQ. W pozostałych przypadkach również zaobserwowano prezentowane tendencje.



Rys. 2. Okno ustawień parametrów oceny obiektywnej protokołu PEAQ Advanced



Rys. 3. Wyniki dla protokołów PEAQ i, PEAQ+ERB

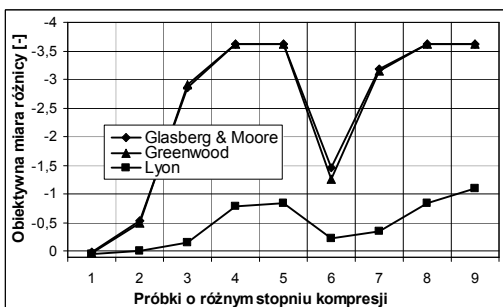


Rys. 4. Wyniki dla protokołu PEAQ+ERB

Dla każdego protokołu oceny obiektywnej stworzono okno umożliwiające niezależną regulację wewnętrznych parametrów przetwarzania sygnałowego. Możliwe jest również iteracyjne zmienianie poszczególnych parametrów w określonych granicach z zadaniem krokiem. Takie rozwiązanie w znacznym stopniu ułatwia poszukiwanie optymalnego zestrojenia poszczególnych protokołów. Poprzez optymalne zestrojenie, czy też poprawę wyników rozumiana jest minimalizacja różnic pomiędzy wynikami otrzymanymi w badaniach subiektywnych a rezultatami pracy prezentowanego oprogramowania. Słuszność tezy postawionej na początku badań mówiącej, że możliwe jest zoptymalizowanie i poprawa istniejących już protokołów, nawet tych zalecanych w oficjalnych dokumentach ITU, została potwierdzona w prowadzonych eksperymentach.

Jedną z zaproponowanych modyfikacji była zamiana wysokościowej skali barkowej wykorzystywanej w oryginalnych postaciach zaimplementowanych protokołów na skalę ERB. Szczegóły dotyczące tej zamiany przedstawiono we wcześniejszej pracy [9]. Zauważono znaczącą poprawę uzyskiwanych ocen wynikającą z tej modyfikacji. Można to zaobserwować analizując wyniki przedstawione na rys. 1, 3, 4, szczególnie relacje pomiędzy ocenami uzyskanymi dla próbek nr 2-6 oraz 8-9.

Dodatkowo zbadano wpływ ustawień parametrów filtrów ERB na wyniki oceny obiektywnej. Na rys. 5 przedstawiono oceny uzyskane przy wykorzystaniu parametrów zaproponowanych przez Glasberga i Moore'a [4], Greenwooda [5] i Lyona [12]. Wyniki zdobyte przy zastosowaniu parametrów zaproponowanych przez Lyona okazały się lepsze niż w pozostałych przypadkach. Różnica polega m. in. na gorszej ocenie próbki nr 9 od próbki nr 8, co jest zgodne z wynikami subiektywnymi.



Rys. 5. Wyniki PEAQ dla różnych parametrów ERB

3. PODSUMOWANIE

Zastosowanie skali ERB wprowadziło wyraźną poprawę ocen. Odpowiedni dobór parametrów ERB umożliwił dodatkową poprawę wyników. Po zoptymalizowaniu ERB możliwe będzie zestrojenie pozostałych wewnętrznych parametrów przetwarzania sygnałowego tak, aby uzyskiwane oceny obiektywne były maksymalnie zbliżone do wyników odsłuchów.

Praca finansowana ze Środków KBN w roku 2004 - projekt badawczy nr 3 T11D 020 26

LITERATURA

1. Beerends J. G., Stemerding J. A., "A Perceptual Audio Quality Measure Based on a Psychoacoustic Sound Representation (PAQM)", J. Audio Eng. Soc., vol. 40, No. 12, pp. 963-978, 1992.
2. Beerends J. G., Stemerding J. A., "A Perceptual Speech Quality Measure Based on a Psychoacoustic Sound Representation", J. Audio Engineering Society Vol. 42, No. 3, pp. 115-123, 1994.
3. Dobrucki A. B., Kozłowski P., "Objective, Perceptual Based Evaluation of Compressed Speech and Audio Signals", IEEE Scientific Workshop Signal Processing '2003, Poznań, 2003. str. 17-21.
4. Glasberg B. R., Moore B. C. J., "Derivation of auditory filter shapes from notched-noise data", Hearing Research, vol. 47, 1990.
5. Greenwood Donald D., "A cochlear frequency-position function for several species-29 years later", Journal of the Acoustical Society of America, Vol. 87 (6), 1990 June.
6. ITU-R Recommendation BS.1387, "Method for Objective Measurements of Perceived Audio Quality (PEAQ)", 1998.
7. ITU-T Recommendation P.862, "Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrow band telephone networks and speech codecs", 2001.
8. Kozłowski P., "Obiektywna ocena jakości sygnałów fonicznych kodowanych stratnie", Elektronizacja nr 4/2003, Warszawa 2003, str.: 14-16.
9. Kozłowski P., Dobrucki A. B., "Proposed Changes to the Methods of Objective, Perceptual Based Evaluation of Compressed Speech and Audio Signals". 116th Convention of the Audio Engineering Society, Berlin 2004.
10. Kozłowski P., Dobrucki A. B., "The Comparison between Subjective and Objective, Perceptual Based Evaluation of Compressed Speech and Audio Signals", International Congress on Acoustics ICA 2004, Kyoto 2004 str II-1811-1814.
11. Kozłowski P., Dobrucki A. B., „Oprogramowanie do obiektywnej oceny jakości sygnałów fonicznych kodowanych stratnie”. Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji KKRRiT 2004, Warszawa 2004, str. 99-102.
12. Slaney Malcolm, "Lyon's Cochlear Model", Apple Technical Report #13, Apple Computer Corporate Library, Cupertino, CA 95014, 1988.