

Andrzej POPŁAWSKI

UNIwersytet Zielonogórski, Instytut Informatyki i Elektroniki

System wspomagający subiektywną ocenę jakości kompresji sekwencji wizyjnych

dr inż. Andrzej POPŁAWSKI

Tytuł doktora otrzymał na Politechnice Poznańskiej w roku 2006, autor lub współautor prac z zakresu kompresji sekwencji wizyjnych. Pracuje na stanowisku adiunkta w Instytucie Informatyki i Elektroniki Uniwersytetu Zielonogórskiego. Głównie zainteresowania naukowe dotyczą zagadnień trójwymiarowej falkowej kompresji sekwencji wizyjnych dla małych opóźnień kodowania.

e-mail: A.Poplawski@iie.uz.zgora.pl



Streszczenie

Kompresja cyfrowego sygnału wizyjnego, pomimo wnoszenia różnego rodzaju zakłóceń, jest obecnie szeroko wykorzystywana w wielu zastosowaniach; pozwala na znaczne ograniczenie prędkości transmisji w stosunku do sygnału nie poddanego kompresji. W pracy zostały omówione zalecane przez International Telecommunication Union (ITU) najczęściej stosowane metody do oceny jakości kompresji obrazów. Zaprezentowano system, którego zadaniem jest ułatwienie przeprowadzenia subiektywnej oceny jakości kompresji sekwencji wizyjnych. Umożliwia on przeprowadzenie oceny jakości zdalnie, poprzez sieć Internet, co poszerza grono osób mogących brać udział w teście a także ułatwia uzyskanie dużej liczby ocen.

Słowa kluczowe: subiektywna ocena jakości sekwencji wizyjnych, kompresja sekwencji wizyjnych

Computer-aided system for subjective video quality assessment

Abstract

A video signal compression is today widely used in many applications. Such compression introduced artefacts to the signal, such as: blocking effect, loss of details, colour distortions, mosquito effect, dirty window effect, motion irregularity, geometric distortions [1, 2]. There is a need to compare various codecs in order to choose the best algorithms of compression. Nowadays many types of objective video quality assessments are done by measuring differences between the original video and the reconstructed video. The most commonly used objective measure of image quality is the PSNR (1) [1, 5]. Objective methods are simple and fast, but sometimes not reliable. That is the reason why it is recommended to do subjective quality assessment [3]. There exists standardized test methods for subjective video quality assessment; the most accepted methods are produced by ITU-R [6]. These tests require a large group of test persons that are non-experts on video. The most commonly used methods are: Double Stimulus Impairment Scale, Double Stimulus Continuous Quality Scale, Single Stimulus. In the paper software that is helpful in doing subjective video quality assessment there is proposed and presented. The software has possibility to do subjective tests through the Internet and have client-server architecture. Such construction makes quality assessment much easier and more accessible for assessors. Views of some windows of the software are presented in fig. 3 to fig. 6.

Keywords: subjective video quality assessment, video sequence compression

1. Wstęp

Współczesne społeczeństwo informacyjne, którego cechami są m.in. przetwarzanie, przekazywanie oraz pobieranie informacji, wykorzystuje różnego rodzaju usługi telekomunikacyjne, zarówno do przekazywania informacji dźwiękowej jak i wizyjnej. Przesyłanie obrazów ruchomych, zwłaszcza w wysokich rozdzielczościach, wymaga stosowania bardzo szybkich łącz. Ze względu na ograniczenia prędkości transmisji powszechnie stosuje się zazwyczaj stratną kompresję sygnału wizyjnego, co z jednej strony

pozwala na wykorzystanie wolniejszych łącz, z drugiej zaś, wprowadza do transmitowanego sygnału wizyjnego różnego rodzaju zniekształcenia. Wielkość oraz rodzaj tych zniekształceń zależy między innymi od stopnia kompresji, rodzaju zastosowanego kodeka oraz treści obrazu ruchomego. Do najczęściej występujących zniekształceń możemy zaliczyć: efekt blokowy, efekt komarów, efekt brudnej szyby, nieregularność ruchu, zniekształcenia geometryczne, zniekształcenia koloru [1, 2].

Ze względu na dostępność wielu różnych kodeków sekwencji wizyjnych, wykorzystujących różne techniki kompresji, jak również przy wprowadzaniu zmian do już istniejących kodeków, zachodzi potrzeba ich porównania i znalezienia rozwiązań dających najlepsze rezultaty. Potrzebne są więc odpowiednia narzędzia do porównania różnych rozwiązań. W chwili obecnej do oceny jakości algorytmów kompresji pod kątem wnoszonych przez nie deformacji obrazu stosuje się zazwyczaj ilościowe miary obiektywne, oparte na porównaniu różnic w obrazie oryginalnym i zrekonstruowanym (po kompresji a następnie dekompresji) [3, 4]. Powszechnie jako miarę jakości wykorzystuje się impulsowy współczynnik sygnału do szumu PSNR (*Peak Signal-to-Noise Ratio*) zdefiniowany równaniem [1, 5]:

$$PSNR = 10 \log \frac{255^2 N^2}{\sum_i e_i^2} [dB], \quad (1)$$

gdzie:

- 255 – zakres dynamiczny sygnału,
- N – liczba punktów w obrazie,
- e_i – różnica wartości między i -tym punktem obrazu oryginalnego i zdekodowanego.

Zaletą metod obiektywnych są proste obliczeniowo algorytmy, jednakże wyniki oceny stopnia zniekształceń (określającymi stopień podobieństwa pomiędzy obrazem oryginalnym, a obrazem zrekonstruowanym), uzyskane przy zastosowaniu tych miar jakości, mogą w pewnych przypadkach znacząco odbiegać od subiektywnych ocen obserwatorów ze względu na to, że dotychczas nie udało się znaleźć dobrej metody odzwierciedlającej sposoby percepcji różnych zakłóceń przez ludzi [1]. Dlatego też zaleca się dodatkowe stosowanie miar subiektywnych, będących uzupełnieniem miar obiektywnych [3].

2. Subiektywna ocena jakości kompresji sekwencji wizyjnych

Metoda oceny subiektywnej jakości obrazów zalecana przez International Telecommunication Union ITU-R [6], polega na ocenie prezentowanego materiału przez osoby oceniające (widzów), a następnie na uśrednieniu uzyskanych w ten sposób ocen. Widzowie mogą oceniać jakość sekwencji wizyjnej lub oceniać różnice jakości w stosunku do sekwencji odniesienia. Uzyskiwana miara jest oznaczona jako MOS (*mean opinion score*).

Wyróżnia się cztery rodzaje obrazów prezentowanych podczas przeprowadzania oceny:

- T1 – obraz wzorcowy,
- T2 – obraz szary,
- T3 – obraz oceniany,
- T4 – obraz szary.

Czas prezentowania obrazów T1 oraz T3 wynosi zazwyczaj 10 sekund, obrazu T2 – 3 sekundy oraz obrazu T4 od 5 do 11 sekund.

Ocenę jakości sekwencji wizyjnych można przeprowadzać w różnych konfiguracjach, zależnie od potrzeb. W tym celu wykorzystuje się odpowiednio zdefiniowane metody pomiarowe. Najczęściej wykorzystywane metody pomiarowe to:

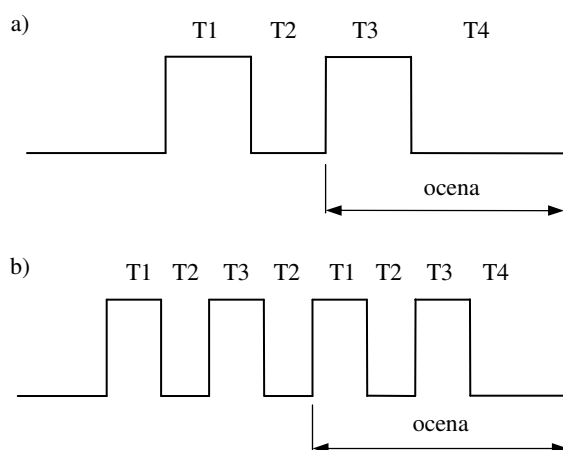
- DSIS (*Double Stimulus Impairment Scale*),
- DSCQS (*Double Stimulus Continuous Quality Scale*),
- SS (*Single Stimulus*).

W metodzie DSIS obserwatorom przedstawiana jest następująca kolejność scen: sekwencja oryginalna, trzysekundowy obraz szary, następnie sekwencja zdekodowana i na zakończenie obraz szary trwający od 5 do 8 sekund. Wyniki reprezentowane są w skali pięciostopniowej przedstawionej w tab. 1. Obserwatorzy oceniają różnicę pomiędzy sekwencją wzorcową, a sekwencją zdekodowaną. Test metodą DSIS przeprowadzać można w dwóch wariantach: wariant I – sekwencja wzorcową oraz sekwencja zdekodowana prezentowane są tylko raz, co przedstawia rys. 1a), wariant II – sekwencja wzorcową oraz sekwencja zdekodowana prezentowane są dwukrotnie, co przedstawia rys. 1b),

Tab. 1. Skala ocen w metodzie DSIS

Tab. 1. Grade scale in DSIS method

Wartość oceny	Ocena
5	różnice niezauważalne
4	różnice zauważalne lecz niedokuczliwe
3	różnice lekko dokuczliwe
2	różnice dokuczliwe
1	różnice bardzo dokuczliwe

Rys. 1. Kolejność prezentacji obrazów w metodzie DSIS, a) wariant I, b) wariant II
Fig. 1. Pictures presentation sequence in method DSIS, a) variant I, b) variant II

Metoda DSCQS polega na ocenie obu prezentowanych obrazów (zarówno sekwencji wzorcowej jak i sekwencji ocenianej). Obserwator w związku z tym nie wie nawet, który obraz poddawany jest testom. Ocena jakości obrazu może być dokonywana albo jako ocena słowna jakości obrazu, co przedstawiono w tab. 2, albo jako oceny w skali od 1 do 100 określające bezwzględną jakość obrazu. Kolejność prezentowanych obrazów w tej metodzie jest analogiczna jak w wariantie II metody DSIS.

Tab. 2. Skala ocen w metodzie DSCQS

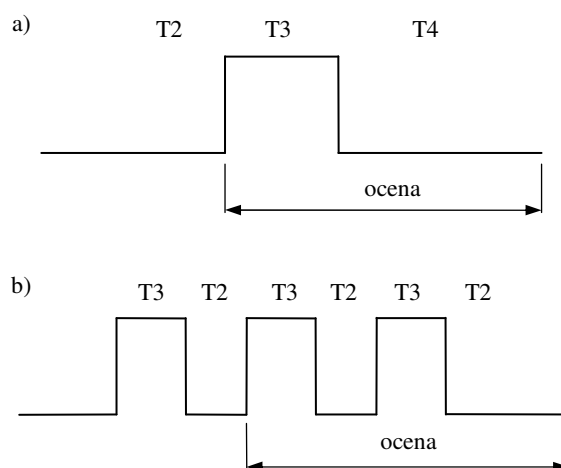
Tab. 2. Grade scale in DSCQS method

Wartość oceny	Ocena
5	jakość bardzo dobra
4	jakość dobra
3	jakość dostateczna
2	jakość mierna
1	jakość niedostateczna

W metodzie SS pokazywane są zarówno obrazy wzorcowe jak i obrazy oceniane. Odtwarzanie może być wykonywane w dwóch wariantach: wariant I – bez powtarzania obrazów zdekodowanych, kolejność prezentowanych obrazów przedstawiono na rys. 2a), wariant II – z wielokrotnym powtarzaniem obrazów zdekodowa-

nych, kolejność prezentowania obrazów przedstawiono na rys. 2b).

Materiały do oceny należy przygotować w taki sposób, aby dane sesja testowa nie trwała dłużej niż 30 minut. Oceny materiału dokonuje grupa składająca się z minimum 15 osób (przy czym nie powinni to być eksperci z danej dziedziny), najlepiej w różnym przedziale wiekowym. Należy poinformować osoby oceniające o metodzie oceny oraz sposobie oddawania ocen. Przed rozpoczęciem testu właściwego przeprowadzany jest test próbny, mający na celu adaptację wzroku oceniających do aktualnie panujących warunków. Podczas testu próbnego oceniający mają możliwość oddawania ocen, przy czym oceny te nie są uwzględniane w teście właściwym.

Rys. 2. Kolejność prezentacji obrazów w metodzie SS, a) wariant I, b) wariant II
Fig. 2. Pictures presentation sequence in method SS, a) variant I, b) variant II

Prawidłowo przeprowadzone pomiary subiektywnej oceny jakości sekwencji wizyjnych, są uważane za najbardziej wiarygodny fundament oceny jakości różnych technik kompresji [1]. Niestety, takie pomiary są żmudne, czasochłonne i kosztowne. Dlatego też wykorzystanie oprogramowania wspomagającego wykonanie takich testów może przyczynić się do znacznego usprawnienia ich przeprowadzenia.

3. Oprogramowanie wspomagające subiektywną ocenę jakości kompresji sekwencji wizyjnych

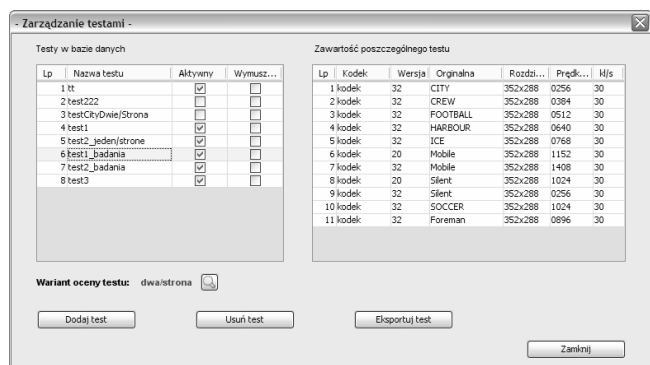
Podstawową funkcją stworzonego systemu jest możliwość przeprowadzenia oceny jakości sekwencji wizyjnych z wykorzystaniem komputerów klasy PC. Aplikacja została napisana w architekturze klient-serwer, gdzie klientem jest komputer osoby dokonującej oceny, zaś serwerem komputer osoby nadzorującej system subiektywnej oceny jakości sekwencji wizyjnych. Transmisja danych niezbędnych do przeprowadzenia testów między klientem a serwerem może odbywać się przy wykorzystaniu zarówno sieci LAN jak i sieci Internet. Takie rozwiązanie pozwala na dużą elastyczność w przeprowadzaniu testów. Możliwe jest bowiem udostępnienie zestawów testowych osobom oceniającym nie tylko w obrębie jednego pomieszczenia czy budynku, ale również osobom będącym w różnych, często odległych miejscach z dostępem do sieci Internet.

System po stronie serwera umożliwia:

- zarządzanie zbiorem sekwencji wizyjnych poddawanych ocenie,
- definiowanie oraz edycję testów do przeprowadzenia,
- udostępnianie osobom oceniającym wybranych testów do przeprowadzenia,
- definiowanie dowolnych wariantów i metod oceny,
- zarządzanie ocenami uzyskanymi podczas przeprowadzanych testów,

- sporządzanie raportów oraz wykresów na podstawie przeprowadzonych testów.

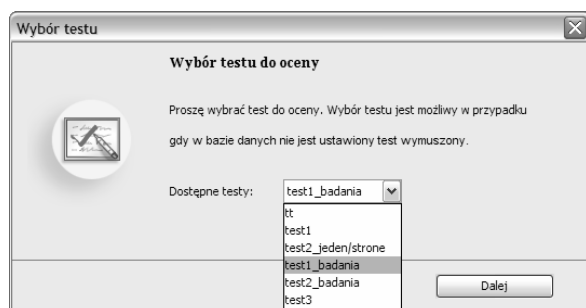
Administrator systemu po wprowadzeniu sekwencji wizyjnych do systemu (sekwencje oryginalne oraz zdekodowane) ma możliwość definiowania testów z wybranymi sekwencjami (rys. 3), określenia metody/wariantu oceny danych sekwencji. Test taki po udostępnieniu dla oceniających może zostać przesłany na stanowisko klienta, zaś po jego zakończeniu wyniki oceny są automatycznie przesyłane do komputera serwera i zapisywane w bazie wyników.



Rys. 3. Okno programu zarządzania testami

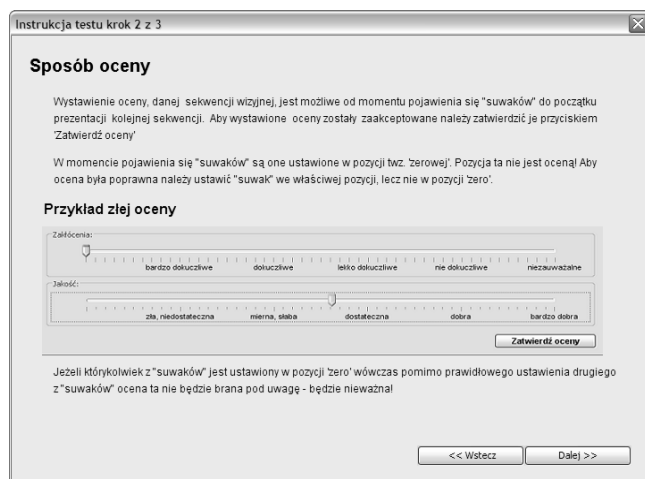
Fig. 3. Test management window

Aplikacja po stronie klienta, czyli osoby oceniającej, po uruchomieniu wymaga nawiązania połączenia się z aplikacją serwera. Po nawiązaniu połączenia osoba oceniająca ma możliwość wyboru oraz pobrania testu do oceny (rys. 4).



Rys. 4. Okno programu wyboru testu

Fig. 4. Test select window



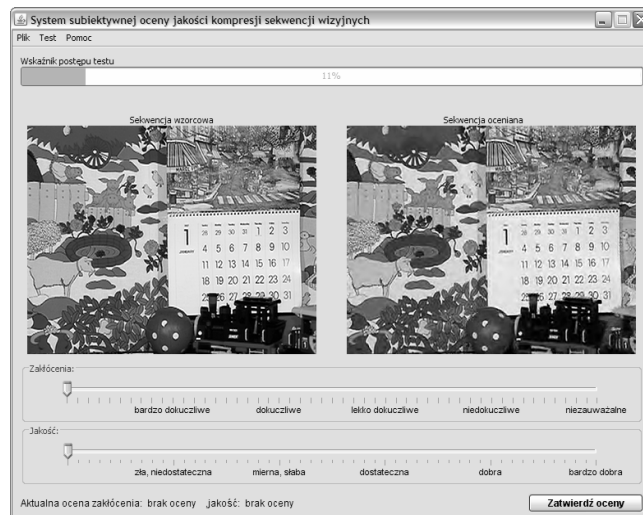
Rys. 5. Okno programu z instrukcją do testu

Fig. 5. Program instruction window

W dalszej kolejności prezentowana jest oceniającemu instrukcja, opisująca w jaki sposób należy poprawnie oddawać oceny w danym teście (rys. 5). Prezentowany jest również przykładowy

test, którego celem jest zaznajomienie oceniającego ze sposobem przeprowadzenia testu właściwego, jak również adaptacji wzroku do panujących warunków.

Po prezentacji instrukcji rozpoczyna się test właściwy. Oceniającemu prezentowane są wybrane do testu sekwencje wizyjne zgodnie z przyjętym wariantem oceny. Osoba oceniająca dokonuje subiektywnej oceny poprzez przesunięcie suwaka (rys. 6) znajdującego się poniżej prezentowanych sekwencji w czasie przeznaczonym na oddawanie ocen.



Rys. 6. Okno programu testu właściwego

Fig. 6. Main test window

Po zakończeniu testu na ekranie prezentowane jest okno podsumowujące, wyświetlające liczbę prawidłowo oddanych ocen. Osoba oceniająca ma możliwość rozpoczęcia kolejnego testu lub zamknięcia aplikacji.

4. Wnioski

Zaprezentowany system wspomagający ocenę jakości kompresji sekwencji wizyjnych jest narzędziem uniwersalnym, mogącym mieć zastosowanie zarówno do oceny nowych lub istniejących koderów sekwencji wizyjnych, jak również w celach dydaktycznych. Możliwość przeprowadzenia oceny zdalnie, poprzez sieć Internet, poszerza grono osób mogących brać udział w teście a także ułatwia uzyskanie dużej liczby ocen. Oceny subiektywne oddawane przez poszczególnych oceniających cechują się niepowtarzalnością oddawanej oceny, co jest cechą naturalną. Dlatego też wiarygodność przeprowadzonego testu wzrasta wraz ze wzrostem liczby oddanych ocen. Pomimo tego, że oddawane oceny zależą m.in. od kontekstu innych sekwencji przedstawianych w teście, uważa się je za najbardziej wiarygodne wyniki oceny jakości kompresji sekwencji wizyjnych.

5. Literatura

- [1] M. Domański: Zaawansowane techniki kompresji obrazów i sekwencji wizyjnych. Politechnika Poznańska, Poznań 2001.
- [2] K. Sayood: Kompresja danych. Wydawnictwo READ ME, Warszawa 2002.
- [3] Z. Barańczuk: Obiektywna ocena jakości kompresowanych stratnie obrazów statycznych. Konferencja MISSI 2000.
- [4] W. Skarbek: Multimedia. Algorytmy i standardy kompresji. Akademicka Oficyna Wydawnicza PLJ, Warszawa 1998.
- [5] J.-R. Ohm: Multimedia Communication Technology. Berlin, New York, Springer 2004.
- [6] ITU-R, Recommendation BT.500-11: Methodology for subjective assessment of the quality of television pictures, 2002.