



Gdańsk, 14-16 maja 2012

## OCENA JAKOŚCI VIDEO 3D-HD ZNIEKSZTAŁCONEGO PRZEZ KOMPRESJĘ I STRATY PAKIETÓW

**Streszczenie:** Obecnie, coraz większą popularność zdobywają urządzenia umożliwiające wyświetlanie treści 3D. Dla rozwoju usług związanych z dystrybucją wideo tego typu, fundamentalne znaczenie ma jakość postrzegana przez użytkowników (ang. *Quality of Experience, QoE*). Zależy ona od wielu czynników, które są ze sobą powiązane. W artykule zaprezentowano badania postrzeganej jakości dla wideo 3D-HD w reprezentacji „up side down side”, które zostało zniekształcone przez kompresję i straty pakietów w sieciach telekomunikacyjnych.

### 1. WSTĘP

Sukces i konkurencyjność aplikacji multimedialnej zależy nie tylko od innowacyjności zastosowanych rozwiązań, ale przede wszystkim od ich jakości postrzeganej przez użytkownika. Jeśli jego oczekiwania z tym związane nie zostaną spełnione, to nie zainwestuje on w dostarczoną mu usługę (bądź urządzenie) albo po prostu nie będzie z niej korzystał w przyszłości.

W artykule skoncentrowano się na problemie wpływu kompresji i strat pakietów (PLR) na jakość wideo trójwymiarowego (3D). Faktem powszechnie znanym jest duże podobieństwo łańcuchów dostarczania treści 2D i 3D. Niestety, ze względu na specyficzną naturę treści 3D ich zakłócenia i zniekształcenia występujące w kolejnych etapach łańcucha dostarczania są diametralnie inne od tych znanych dla łańcucha 2D [1].

Zaprezentowano wyniki badań i wnioski uzyskane w wyniku przeprowadzenia testu jakości postrzeganej przez użytkowników. Na podstawie tych badań sformułowano wnioski dotyczące zależności pomiędzy utratą pakietów i przepływnością kodowania a jakością subiektywną sekwencji 3D.

Struktura artykułu jest następująca: w sekcji 2 przedstawiono aktualny stan wiedzy tj. podobne badania przeprowadzone w innych ośrodkach naukowych i płynące z nich wnioski. W sekcji 3 omówiono szczegóły związane z przygotowaniem eksperymentu – sposób wygenerowania materiału testowego. Metodę przeprowadzenia testów subiektywnych zaprezentowano w sekcji 4, po której tj. w sekcji 5 przeprowadzono analizę wyników badań i zebrano uzyskane wnioski.

### 2. STAN WIEDZY

W kilku laboratoriach na świecie zostały przeprowadzone eksperymenty o zbliżonym zakresie obejmują-

cym zarówno wpływ kompresji jak i strat pakietów podczas transmisji.

Wang i in. [2] przeprowadzili badania postrzeganej jakości dla sekwencji wideo 3D-HD zakodowanych różnymi metodami kodowania (Simulcast, MVC oraz Simulcast z mniejszą rozdzielczością w czasie i przestrzeni). W jednym z wniosków wskazano, że taka sama jakość może być zapewniona przy znacznej redukcji przepływności, która może być osiągnięta przez obniżenie rozdzielczości. Wpływ próbkowania w dół (ang. *downsampling*) na jakość wideo 3D był badany również przez Stelmacha i in. [3]. W swoich eksperymentach uwzględnili oni różne kombinacje poziomego i pionowego próbkowania w dół i poddali ocenie ich wpływ na jakość oraz wrażenie głębi i ostrości. Tego typu przestrzenne próbkowanie w dół jest obecnie stosowane dla transmisji 3D HDTV z wykorzystaniem reprezentacji „side by side” (obraz dla lewego oka sąsiaduje w jednej ramce z obrazem przeznaczonym dla prawego oka). Tego typu reprezentację w swoich badaniach zastosowali Lebreton i in. [4]. Na ich podstawie potwierdzili, że jest ona najbardziej efektywnym rozwiązaniem dla transmisji sekwencji stereoskopowych 3D-HD – wymaga węższego pasma niż metody Simulcast czy MVC wykorzystujące pełną rozdzielczość ramki. Dodatkowo badania wskazały, że testerzy mają szczególne trudności przy ocenie dyskomfortu, pojęcia, które nie jest dla nich jasne. W drugiej części badań porównywano jakość wideo 2D i 3D zniekształconych przez straty pakietów. Testy nie wykazały znacznych różnic w postrzeganej jakości. Na tej podstawie naukowcy wywnioskowali, że teoria „*binocular suppression*” nie sprawdza się dla sekwencji zniekształconych błędami transmisji. Zakłada ona, że percepcja pary obrazów przeznaczonych odpowiednio dla lewego i prawego oka jest zdominowana przez ten o wyższej jakości, zatem teoretycznie wystarczy zapewnić wysoką jakość obrazu tylko dla jednego oka.

### 3. PRZYGOTOWANIE EKSPERYMENTU

W tej sekcji pokrótce omówiono sposób przygotowania sekwencji testowych, które podlegały ocenie przez testerów. Celem badań była analiza wpływu kompresji i strat pakietów na jakość stereoskopowego wideo FullHD prezentowanego na migawkowym wyświetlaczu 3D.

### 3.1. Sekwencje źródłowe (SRC)

Sekwencje źródłowe (SRC, *Source Video Sequences*) zostały wyprodukowane przez japońską korporację NTT i udostępnione w ramach współpracy w Video Quality Experts Group (VQEG). Dostarczono zbiory nieskompresowanych plików. Każdy plik w formacie TIFF i rozdzielczości 4K zawierał pojedynczą klatkę dla prawego bądź lewego oka. Sekwencje wideo ostatecznie nagrano z prędkością 30 FPS. Następnie podzielono je na sceny, z których wybrano 9 najbardziej charakterystycznych. Za pomocą skryptu Perl dokonano konwersji z rozdzielczości 4K do FullHD w trybie „up side down side” (obraz dla lewego oka i obraz dla prawego oka znajdują się jednej ramce jeden nad drugim) do formatu PNG. W ten sposób uzyskano pliki 1920 x 2160 pikseli zawierające ramki widoku lewego i prawego – jedna nad drugą. Ostatnim krokiem było przekształcenie plików PNG do postaci sekwencji wideo w formacie YUV 420p przy użyciu narzędzia ffmpeg.

### 3.2. Hipotetyczne cykle referencyjne (HRC)

Aby wygenerować materiał testowy dla celów eksperymentu sekwencje źródłowe zostały zmodyfikowane według 18 różnych cykli referencyjnych (HRC, *Hypothetical Reference Circuits*). W ten sposób powstały 162 sekwencje użyte następnie w trakcie testów (PVS). Sekwencje źródłowe zostały podzielone na dwie grupy i każda grupa została zakodowana dla 6 różnych przepływności. Dodatkowo utworzono dwie grupy współczynników strat pakietów i zastosowano je do zniekształcenia wideo o różnych przepływnościach.

### 3.3. Generowanie sekwencji testowych (PVS)

Pierwszym krokiem była dwuetapowa kompresja sekwencji źródłowych. Najpierw zostały one zakodowane przy użyciu kodeka x264 do strumienia H.264 Annex B. Uzyskane pliki zostały zniekształcone przez dodanie strat pakietów przy użyciu oprogramowania Sirannon [5]. W celu zapewnienia pełnej kontroli transmisji eksperymenty zostały przeprowadzone przy użyciu dwóch komputerów PC połączonych bezpośrednio kablem Ethernet. Serwer posiadał zainstalowany system operacyjny Ubuntu 10.10, natomiast klient Microsoft Windows 7. Zostało zastosowane strumieniowanie RTP dla plików wideo opisane w dokumencie RFC3984 [6]. Plikami wejściowymi dla serwera były pliki XML i sekwencje PVS. Struktura serwera składała się z następujących bloków zdefiniowanych w plikach XML: AVC-reader, AVC-packetizer, basic-scheduler i nadajnika RTP. Po stronie klienta uruchomiono oprogramowanie Wireshark w celu przechwytywania ruchu wideo.

## 4. PRZEBIEG EKSPERYMENTU

Środowisko testowe przygotowano zgodnie z Rekomendacjami ITU-R BT.500-12 [9], ITU-T P.910 [10] i ITU-R BT.1438 [11]. W teście użyto wyświetlacza 3D BENQ 24 wraz z okularami migawkowymi.

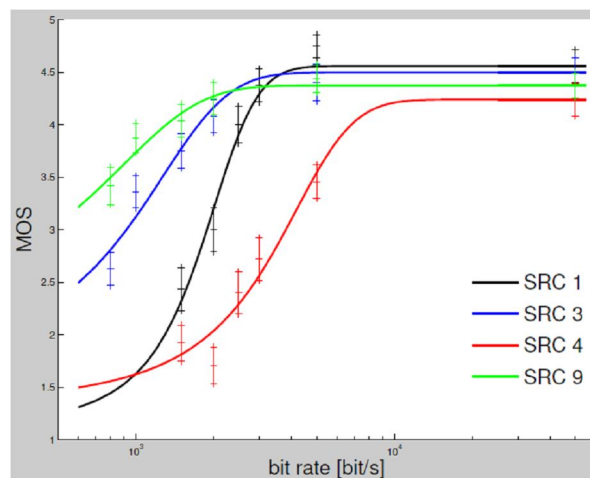
Badanie zostało wykonane zgodnie z metodą ACR (*Absolute Category Rating*) opisaną w Rekomendacji ITU-T P.910 [10]. Testerzy (łącznie 30 osób) wykonywali badanie indywidualnie. Na etapie poszukiwań, kryterium selekcji do uczestnictwa w teście był pozytywny wynik badania percepcji głębi. Każdy tester po obejrzeniu trwającej 10 s sekwencji wideo miał 14 s na wskazanie subiektywnej oceny w skali MOS jej jakości, po tym czasie na wyświetlaczu prezentowano kolejną sekwencję do oceny. Czas trwania pojedynczego testu był równy około godzinę i 45 minut. Uczestnicy byli proszeni o przerwanie eksperymentu, gdyby poczuli się źle, wystąpiły u nich zawroty głowy lub inne symptomy specyficzne podczas korzystania z technologii 3D. Po wykonaniu testu przeprowadzono wywiad, w którym pytano o wrażenia z testu, dotychczasowe doświadczenia z technologią 3D i kryteria wyboru ocen ze skali. Uzyskano także informacje charakterystyczne dla osoby testującej takie jak wiek, płeć, poziom wykształcenia. Wykonano test ostrości wzroku oraz test na występowanie daltonizmu.

## 5. DYSKUSJA WYNIKÓW BADAŃ

W tej sekcji zaprezentowano analizę wpływu przepływności i strat pakietów na jakość wideo 3D. W celu zwiększenia przejrzystości wykresy przedstawiają zależności tylko dla czterech różnych sekwencji źródłowych. Są to wykresy funkcji logarytmicznej:

$$f(x) = 1 + \frac{MOS_{\max} - 1}{1 + \exp(ax + b)} \quad (1)$$

gdzie  $x$  jest równe odpowiednio wartości przepływności bądź wartości strat pakietów,  $MOS_{\max}$  jest wartością oceny MOS odpowiednio dla przepływności równej 50 Mbit/s w przypadku funkcji przepływności lub dla braku strat w przypadku funkcji strat pakietów. Zastosowanie tej funkcji umożliwia obserwację trendów pozbawioną wpływu ocen pojedynczych testerów.



Rys. 1. Oceny MOS dla różnych sekwencji źródłowych i przepływności.

Rys. 1. przedstawia wykres ocen MOS w funkcji przepływności kodowania. Na jego podstawie można wnioskować o znacznym wpływie treści sekwencji źródłowych na oceny MOS testerów. Najbardziej znacząca różnica może być zaobserwowana dla przepływności równej 1,5 Mbit/s, gdzie SRC4 została oceniona w skali MOS na mniej niż 2, podczas gdy SRC9 na około 4,25, czyli wyżej niż najwyższa ocena osiągana przez SRC4 dla kolejnych przepływności. Taka sytuacja jest spowodowana przede wszystkim różnicami w stopniu skomplikowania scen. Sekwencja SRC4 jest bardziej szczegółowa i dynamiczna zarówno jeśli chodzi o obiekty jak i pozycję kamery. Z kolei sekwencja SRC9 została nagrana bez zmian pozycji kamery, tylko jeden aktor poruszał się powoli, a tło było pozbawione wielu detali. Drugi wniosek jest bardziej zaskakujący. Teoretycznie zastosowana metoda kodowania dla transmisji treści 3D powinna wymagać zastosowania dwukrotnie większego pasma niż w przypadku transmisji treści 2D, jednakże otrzymane wyniki wskazują, że już 5 Mbit/s pozwala uzyskać jakość najwyższą ocenianą przez testerów. Statystycznie podobna sytuacja zachodzi dla 50 Mbit/s z wyjątkiem SRC4. Nasze własne obserwacje wykazały, że pojedynczy widok dla przepływności 5 Mbit/s jest już znacznie zniekształcony przez artefakty kompresji, niemniej jednak układ wzorkowy człowieka jest zdolny do znacznie lepszej percepcji, gdy jest stymulowany dwoma różnymi widokami (zgodnie z teorią „*binocular suppression*”).

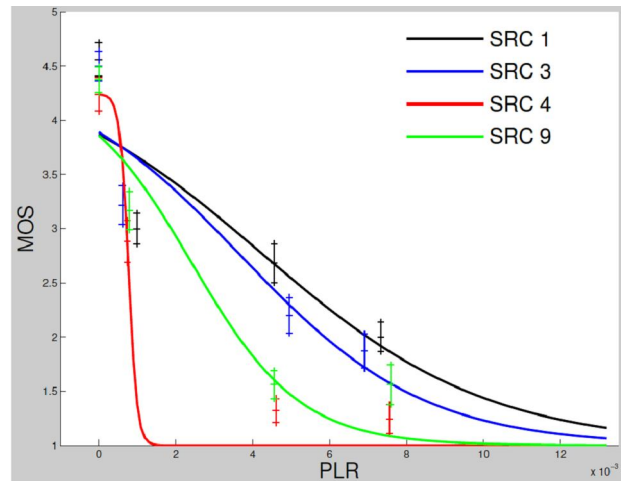
Ograniczone pasmo skutkuje zastosowaniem metod kompresji, które wprowadzają artefakty w transmitowanych obrazach. Innym zjawiskiem zachodzącym w sieciach telekomunikacyjnych, które powoduje zniekształcenia wideo jest utrata pakietów. W badaniach uwzględniono różne wartości przepływności bitowej ze względu na silną zależność pomiędzy stratami pakietów a przepływnością kodowania sekwencji. Na rys. 2 przedstawiono wyniki uzyskane dla różnych przepływności i współczynników strat pakietów.

Wpływ strat pakietów jest bardzo interesujący i silnie zależy od przepływności. Dla przepływności 50 Mbit/s (rys. 2a) niewielkie straty pakietów powodują znaczne pogorszenie jakości. Najmniejsze wprowadzone w badaniu straty były na poziomie 0,1%. Spadek ocen MOS jest statystycznie istotny dla wszystkich sekwencji i ostatecznie zatrzymuje się przy wartości około 1,25 tj. wynosi więcej niż 30% możliwej zmiany w skali MOS. Dla wyższego prawdopodobieństwa strat ta zmiana nie jest tak znaczna, a przynajmniej w przypadku sekwencji SRC1 i SRC3. Co więcej, zmiana ta nie zależy tak bardzo od sekwencji jak od przepływności. To zjawisko pokazuje, jak bardzo ważna jest kontrola strat pakietów i jak zasadnicze znaczenie powinna mieć dla operatorów.

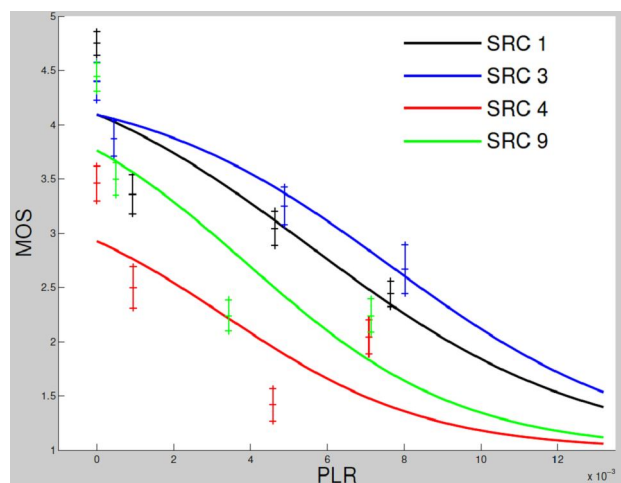
W przypadku silniejszych kompresji można wyciągnąć podobne wnioski. Jedyną różnicą polega na tym, że występują konkretne zakresy, w których zakłócenia spowodowane stratami nie są zauważalne dla testerów. Na przykład na rys. 2d ocena MOS dla sekwencji ze

stratami jest wyższa (ze względu na wartość, a nie statystycznie) niż ocena dla sekwencji bez strat.

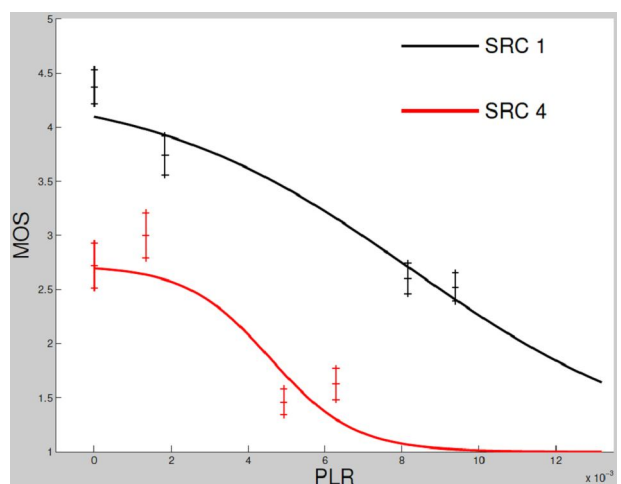
Kolejnym wnioskiem z przeprowadzonych badań jest brak możliwości „ukrycia” za pomocą drugiego widoku zniekształceń spowodowanych przez straty pakietów. W każdym razie nie jest to możliwe za pomocą zastosowanych w eksperymencie metod kodowania i dekodowania.



a) dla przepływności 50 Mbit/s

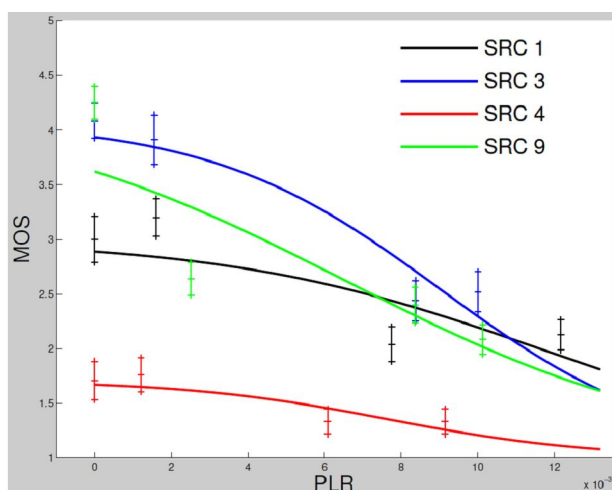


b) dla przepływności 5 Mbit/s

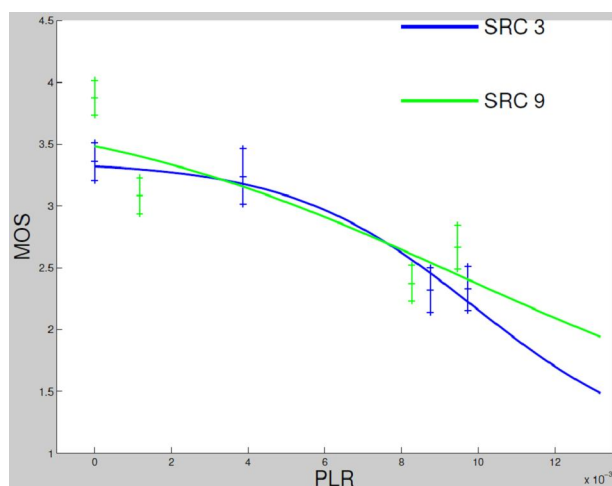




c) dla przepływności 3 Mbit/s



d) dla przepływności 2 Mbit/s



e) dla przepływności 1 Mbit/s

Rys. 2. Oceny MOS w funkcji współczynnika strat pakietów dla różnych SRC i przepływności.

Ostatnią konkluzją jest stwierdzenie, że wpływ sekwencji źródłowych w przypadku strat pakietów jest inny niż w przypadku różnych przepływności. Sekwencja SRC1, która wypadła najgorzej w przypadku analizy wpływu przepływności, była z kolei najbardziej odporna na zniekształcenia spowodowane utratą pakietów. To pokazuje jak trudno jest przewidzieć jakość mając do dyspozycji tylko informację o przepływności i współczynniku strat pakietów.

## 6. PODSUMOWANIE

W artykule przedstawiono badanie jakości QoE w funkcji straty pakietów (PLR), przepływności i rodzaju treści 3D. Pokazano, że wpływ strat pakietów jest silnie zależny od przepływności. Zaobserwowano, że przy zastosowanych w badaniu metodach kodowania i dekodowania straty pakietów nie mogą zostać „ukryte” przez

dostarczenie drugiego widoku. Wskazano, że bardzo trudno przewidzieć jakość postrzegana przez użytkownika jedynie na podstawie informacji o poziomie strat pakietów i przepływności – potrzebna jest ponadto informacja o typie i rodzaju sekwencji.

## 7. PODZIĘKOWANIA

Artykuł powstał w ramach Programu Operacyjnego Innowacyjna Gospodarka - projekt nr POIG.01.01.01-00-045/09-00 „Inżynieria Internetu Przyszłości” sfinansowanego przez Ministerstwo Nauki i Szkolnictwa Wyższego z Europejskiego Funduszu Rozwoju Regionalnego.

## SPIS LITERATURY

- [1] D. Juszka, „Badania subiektywne jakości QoE dla usług telewizji trójwymiarowej”, *Przegląd Telekomunikacyjny*, *Wiadomości Telekomunikacyjne*; ISSN 1230-3496, 2011, R. 84 nr 6 s. 633–636.
- [2] K. Wang, M. Barkowsky, R. Cousseau, K. Brunnström, R. Olsson, P. L. Callet, M. Sjöström, “Subjective evaluation of HDTV stereoscopic videos in IPTV scenarios using absolute category rating”, A. J. Woods, N. S. Holliman, N. A. Dodgson, Eds., *Proc. SPIE*, vol. 7863, no. 1. SPIE, 2011
- [3] L. B. Stelmach, W. J. Tam, D. V. Meegan, A. Vincent, “Stereo image quality: effects of mixed spatio-temporal resolution”, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, Issue 2, str. 188 - 193, 2000
- [4] P. Lebreton, A. Raake, M. Barkowsky, P. L. Callet, “A subjective evaluation of 3D IPTV broadcasting implementations considering coding and transmission degradation”, *International Symposium on Multimedia*, vol. 0, str. 506-511, 2011
- [5] A. Rombaut, SIRANNON 0.6.10: Modular Multimedia Streaming, University of Ghent, IBCN., May 2011, <http://sirannon.atlantis.ugent.be/>
- [6] T. S. M. W. S. Wenger, M. M. Hannuksela, D. Singer, RTP Payload Format for H.264 Video, The Internet Society, February 2005, <http://www.ietf.org/rfc/rfc3984.txt/>
- [7] B. Szczerba, D. Ziobro, Generating of packet loss in the video sequences encoded with H.264 video codec, AGH University of Science and Technology, Atlanta, USA, November 2010, <ftp://vqeg.its.bldrdoc.gov/Documents/>
- [8] FFmpeg, <http://ffmpeg.org/>
- [9] ITU-R BT.500-12, Methodology for the subjective assessment of the quality of television pictures, International Telecommunication Union Recommendation, Rev. 12, 2009
- [10] ITU-T P.910, Subjective video quality assessment methods for multimedia applications, International Telecommunication Union Recommendation, 1999
- [11] ITU-R BT.1438, Subjective assessment of stereoscopic television pictures, International Telecommunication Union Recommendation, 2000