

# Ocena jakości obrazu wideo z kompresją H.264/AVC po transmisji w sieci IP

Celem niniejszego artykułu jest przedstawienie wyników otrzymanych w trakcie badań wpływu parametrów transmisyjnych sieci takich jak pakietowa stopa błędów, opóźnienie oraz wielkość pasma transmisyjnego na jakość percepcyjną obrazu wideo. Aby ocenić wpływ tych parametrów na jakość odbieranego przez klienta materiału wideo, dokonano pomiarów obiektywnych w oparciu o metryki VQM, MSE oraz SSIM. W przeprowadzonych badaniach, jako standard kodowania sekwencji wizyjnych został wykorzystany H.264/AVC.

## 1. Wprowadzenie

Dynamiczny rozwój technik transmisyjnych, powoduje systematyczny wzrost przepustowości sieci, a co za tym idzie możliwość przesyłania za ich pomocą wysokiej jakości wideo i dźwięku. Dzięki zastosowaniu odpowiednich technik kompresji oraz protokołów transmisji możliwe jest strumieniowanie wideo poprzez sieć IP. W związku z tym, że parametry takiej sieci mogą ulegać chwilowemu lub długotrwałemu pogorszeniu, istotne jest aby ocenić jakie są krytyczne wartości tych parametrów i jaki wpływ mają one na jakość przesyłanego obrazu. Najbardziej istotnymi z punktu widzenia transmisji strumieniowej są następujące parametry sieci [1,2]:

- **współczynnik utraty pakietu strumieniowego PLR** (ang. *Packet Loss Ratio*) - wyrażony przez stosunek liczby pakietów utraconych w sieci do całkowitej liczby przesyłanych pakietów;
- **opóźnienie** - stanowiące różnicę między czasem nadawania i czasem odbioru pakietu;
- **jitter** - czyli wahania opóźnienia.

W artykule przedstawiono wpływ wybranych parametrów sieci takich jak liczba utraconych pakietów, opóźnienie oraz przepustowość sieci, na jakość odbieranego przez klienta materiału wideo. Ważnym aspektem jest również ocena wpływu jittera na jakość odbieranego obrazu, ale na obecnym etapie badań takiej oceny nie dokonano.

## 2. Wskaźniki jakości obrazu

Metody obiektywne oceny jakości, polegają na wyznaczeniu charakterystycznych parametrów jakościowych materiału wideo. W zależności od ilości informacji posiadanych na temat sekwencji wideo, można je podzielić na trzy grupy [3]:

- z pełną referencją (ang. *full reference*) – polegające na obliczeniu różnic pomiędzy obrazem oryginalnym i zdegradowanym;
- z częściową referencją (ang. *reduced reference*) – polegające na porównaniu pewnych parametrów sygnału oryginalnego i zdegradowanego;
- bez referencji (ang. *no-reference*) – polegające na analizie jedynie sygnału zdegradowanego.

W przeprowadzonych badaniach pod uwagę brano metody należące do grupy pierwszej, gdyż na chwilę obecną metody z tej właśnie grupy pozwalają na uzyskanie wyników najbardziej skorelowanych z wynikami uzyskanymi w procesie subiektywnej oceny materiału multimedialnego. W badaniach jako miarę postrzeganej jakości odbieranego obrazu wideo wykorzystano metryki: VQM, MSE oraz SSIM.

### **Metryka VQM**

VQM (ang. *Video Quality Metric*) [4] jest opracowaną przez ITS (ang. *Institute for Telecommunication Science*) metryką umożliwiającą dokonanie obiektywnej oceny postrzeganej jakości wideo. Pomiar dokonywany dzięki wykorzystaniu metryki VQM uwzględnia takie zaburzenia sygnału wideo jak: rozmycie, efekt blokowania obrazu, zniekształcenia konturów i ich kombinacje.

### **Metryka MSE**

Metryka MSE (ang. *Mean Squared Error*) wyznacza wartość błędu średniokwadratowego dla poszczególnych pikseli obrazu. Jej wartość oblicza się na podstawie zależności [5]:

$$MSE = \frac{1}{M \cdot N \cdot T} \sum_{t=1}^T \sum_{y=1}^N \sum_{x=1}^M [I(x, y, t) - I_s(x, y, t)]^2 \quad (1)$$

gdzie:

- I – oryginalna sekwencja wideo;
- $I_s$  – odebrana sekwencja wideo;
- T – liczba obrazów w sekwencji wideo;
- M, N – rozmiar obrazu w pikselach.

### **Metryka SSIM**

Metryka SSIM (ang. *Structural SIMilarity*) uwzględnia trzy typy zniekształceń: luminancji, kontrastu i struktury. Zakładając, że  $x = \{x_i | i = 1, 2, \dots, N\}$  jest sygnałem oryginalnym a  $y = \{y_i | i = 1, 2, \dots, N\}$  sygnałem zniekształconym, wartość SSIM można wyznaczyć na podstawie zależności [6]:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (2)$$

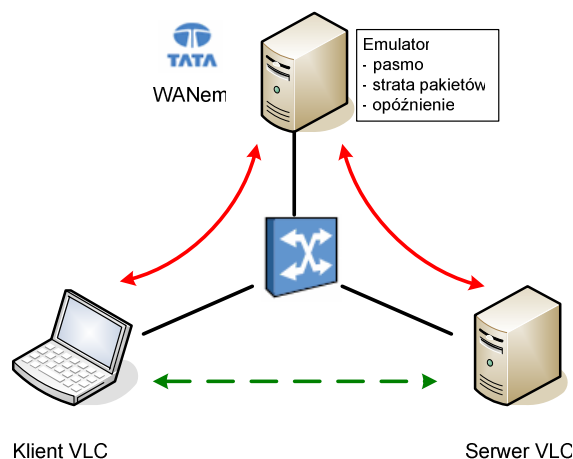
gdzie:

- |                              |   |
|------------------------------|---|
| $\mu_x$ - wartość średnia x; | $\sigma_{xy}$ - kowariancja x i y;      |
| $\mu_y$ - wartość średnia y; | $C_1 = (K_1L)^2$ , $C_2 = (K_2L)^2$ ;   |
| $\sigma_x^2$ - wariancja x;  | L - dynamiczny zakres wartości pikseli; |
| $\sigma_y^2$ - wariancja y;  | $K_1, K_2$ - wartości stałe.            |

## **3. Środowisko testowe**

Wykonanie odpowiednich pomiarów było możliwe dzięki opracowanemu stanowisku pomiarowemu. Zastosowanie emulatora sieci rozległych WANem [7] pozwoliło na zmianę parametrów sieci mających wpływ na jakość odbieranego przez klienta obrazu. Ze względu na fakt, że aplikacja ta została przygotowana wraz z systemem operacyjnym Linux jako samodzielna dystrybucja, parametry oprogramowania zostały zoptymalizowane pod względem pełnienia funkcji symulatora. Cecha ta pozwoliła na zapewnienie powtarzalności warunków panujących w sieci podczas wykonywania pomiarów.

Na Rys. 1 przedstawiono architekturę stanowiska badawczego.



**Rys. 1 Architektura stanowiska badawczego**

W skład architektury stanowiska badawczego wchodziły trzy elementy: serwer, emulator sieci oraz stacja kliencka. Wykorzystanie narzędzia VLC media player pozwoliło na strumieniowanie danych multimedialnych przez serwer oraz ich odbiór i zapisanie po stronie klienta. Wszystkie testowane sekwencje były skompresowane przy użyciu kodera H.264/AVC oraz transmitowane poprzez sieć IP przy wykorzystaniu protokołu UDP. Oceny jakościowej odbieranego materiału wideo dokonywano za pomocą narzędzia *Video Quality Estimator*. Pozwala ono na porównanie zdegradowanego sygnału z sygnałem referencyjnym i wyznaczenie zadanych wskaźników jakości obrazu.

#### 4. Badane sekwencje

W badaniach wykorzystano należące do NTIA/ITS [8] źródłowe sekwencje wizyjne. Spośród dostępnych plików źródłowych zostały wybrane 4 sekwencje, które poddane zostały kompresji H.264/AVC zgodnie z parametrami przedstawionymi w Tab.1.

**Tab. 1 Parametry kompresji badanych materiałów wizyjnych**

L.p.	Sekwencja	Charakterystyka	Rozmiar [piksel]	FPS [s <sup>-1</sup> ]	Przepływność [kbit/s]	Kodowanie	Profil i poziom
1	RedKayak	Dynamiczny obraz ze średnią ilością szczegółów	720×576	25	5008	CABAC	HP@L3
2	SnowMnt	Obraz o niewielkiej ilości ruchu			2711		
3	ControlledBurn	Obraz ze średnią ilością ruchu			4112		
4	WestWindEasy	Obraz z dużą ilością szczegółów			5373		

Wszystkie sekwencje zostały skompresowane z wykorzystaniem takich samych ustawień i zostały dobrane w taki sposób, aby różniły się pod względem ilości występującego w nich ruchu oraz stopnia szczegółowości. W zależności od tego jak dynamiczny jest obraz oraz jak wiele zawiera szczegółów, wybrane pliki charakteryzują się określoną przepływnością, która ma priorytetowe znaczenie podczas strumieniowania materiału wideo poprzez sieć IP.



RedKayak

SnowMnt

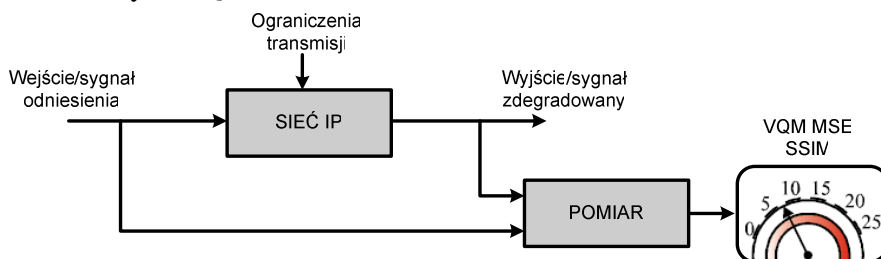
ControlledBurn

WestWindEasy

Rys. 2 Widok testowanych sekwencji wideo

## 5. Metodyka pomiarowa

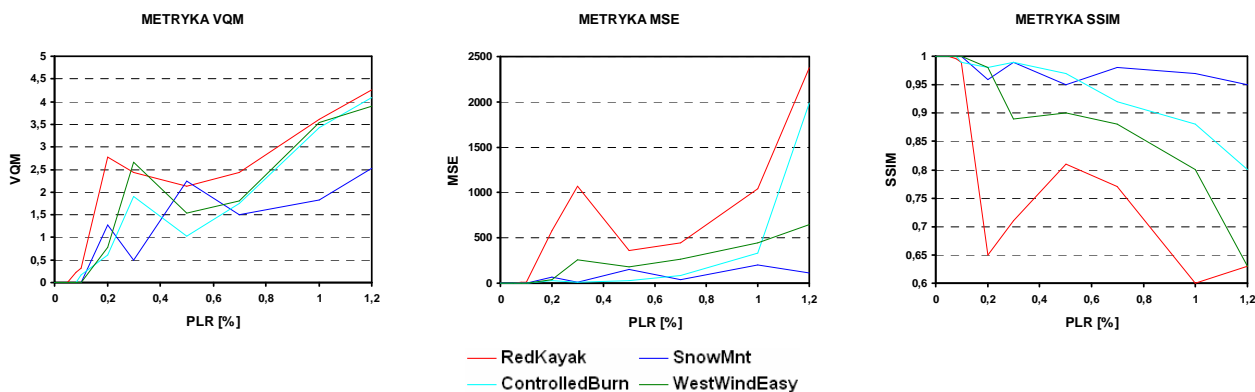
Obiektywne pomiary jakości odbieranego materiału wideo zostały wykonane dla sieci o zadanym opóźnieniu, o ograniczonej przepustowości oraz o zadeklarowanej liczbie traconych pakietów. Wartość pasma, ustawiana przy pomocy emulatora dobierana była w zależności od przepływności materiału multimedialnego, natomiast wartość PLR (ang. *Packet Loss Ratio*) zmieniano w zakresie od 0,05% do 1,2%. Oceny jakości odebranego przez klienta materiału multimedialnego dokonano przy wykorzystaniu trzech metryk: VQM, MSE i SSIM.



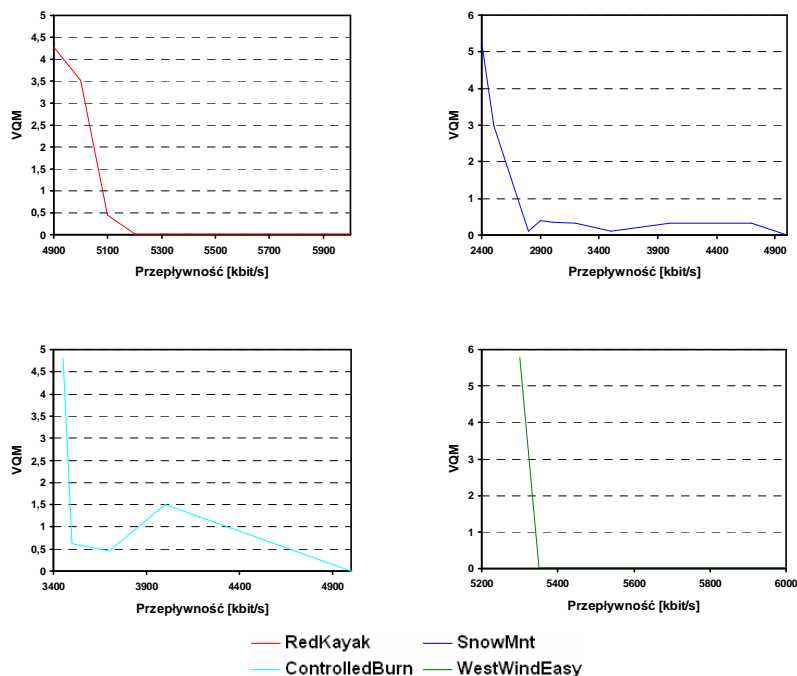
Rys. 3 Zastosowanie metody z pełną referencją do oceny jakości wideo [9]

## 6. Ocena jakości testowanych sekwencji wideo

Na rysunkach 4, 5 zostały przedstawione wyniki obiektywnych pomiarów jakości obrazu z wykorzystaniem opisanych wcześniej metryk. Okazuje się, że nawet bardzo niewielka strata pakietów rzędu 0,05% ma wpływ na jakość odbieranego przez klienta materiału wideo. Ze względu na dużą dynamikę obrazu najbardziej wrażliwą na utratę pakietów była sekwencja RedKayak. Z kolei sekwencja SnowMnt ze względu na niewielką ilość występującego w niej ruchu, wykazała się największą odpornością na występujące w kanale transmisyjnym straty pakietów. Zwiększenie wartości PLR do poziomu 1% powoduje znaczne uszkodzenie przesyłanych przez sieć sekwencji. W efekcie, w odbieranym obrazie pojawiają się efekty blokowe, staje się on rozmyty i nieczytelny. W tym momencie szczególnego znaczenia, nabiera zastosowana podczas kodowania sekwencji struktura GOP (ang. *Group of Pictures*). Uszkodzenie ramki typu I (ang. *Intra*) ma decydujący wpływ na całą strukturę GOP, gdyż powoduje uszkodzenie kolejnych ramek typu P (ang. *Predicted*) oraz B (ang. *Bidirectional*). Uszkodzenie ramek kluczowych prowadzi do propagacji błędów w kolejnych ramkach, a to z kolei do obniżenia jakości odbieranego materiału wideo. Natomiast uszkodzenie ramki typu B może pozostać w ogóle niezauważone.



**Rys. 4 Zależność pomiędzy jakością odbieranego obrazu i pakietową stopą błędów PLR**



**Rys. 5 Zależność pomiędzy jakością odbieranego obrazu i dostępnym pasmem transmisyjnym**

Przeprowadzone badania jakości odbieranego materiału wideo w funkcji przepływności kanału transmisyjnego wykazują znaczne pogorszenie jakości materiału w przypadku gdy przepływność materiału jest wyższa, niż dostępna w kanale. Dodatkowy narzut związany z nagłówkami oraz kodowaniem korekcyjnym powoduje, że wartość pasma przeznaczonego do transmisji wideo powinna mieć wartość wyższą niż chwilowa przepływność sekwencji wideo. Niespełnienie tego warunku prowadzi do natychmiastowego pogorszenia jakości odbieranego materiału, powodując jego całkowitą bezużyteczność.

Ostatnim elementem badań, było określenie wpływu opóźnienia występującego w sieci na jakość odbieranego przez klienta materiału wideo. Wykonane pomiary wykazały, że występujące w sieci opóźnienie powoduje jedynie jednorazowe odsunięcie w czasie momentu odtwarzania sekwencji wizyjnej po stronie klienta. Nie ma natomiast żadnego wpływu na samą jakość przesyłanego przez sieć materiału multimedialnego.

## 7. Podsumowanie

Transmisja danych multimedialnych w sieciach IP wiąże się z szeregiem ograniczeń. Otrzymane w trakcie badań wyniki pokazały, że zarówno utarta pakietów jak i zbyt niska przepływność kanału powodują znaczne pogorszenie jakości odbieranego przez klienta materiału wideo. Aby efektywnie transmitować treść multimedialną przez sieć, należy zagwarantować minimalną przepływność toru transmisyjnego. W przypadku jej braku, jakość odbioru znacznie się pogarsza, występują efekty poklatkowe, interpolacje, utrata synchronizacji a w konsekwencji zerwanie połączenia. Nie bez znaczenia pozostaje również sama treść transmitowanego pliku wideo. W przypadku gdy obraz jest dynamiczny i zawiera dużą ilość szczegółów, wymagania dotyczące parametrów sieci są wyższe niż dla obrazu o małej ilości występującego w nim ruchu. Podobne wnioski do przedstawionych powyżej wyciągnięto podczas badań wpływu parametrów sieci na jakość sygnału wideo z kompresją MPEG-2 przeprowadzonych w Instytucie Telekomunikacji Politechniki Warszawskiej [10].

W przeprowadzonych badaniach nie wykorzystano najnowszych metod oceny jakości obrazu zaproponowanych w zaleceniu J.247. Istotną cechą przedstawionych tam metod jest fakt, że do obliczenia różnic pomiędzy obrazem zdegradowanym a referencyjnym wykorzystują one szereg wskaźników cząstkowych, uwzględniających dokładność odwzorowania luminancji, chrominancji, krawędzi oraz detekcję efektów blokowych i utraconych elementów obrazu [11]. W związku z tym są to metody o dużej wiarygodności, ale ze względu na fakt, że brakuje obecnie oprogramowania z ich implementacją, w przeprowadzonych badaniach zrezygnowano z ich wykorzystania.

Przedstawione w artykule wyniki badań pokazują, że odbiór materiału multimedialnego o założonej jakości wymaga wprowadzenia do sieci dodatkowych elementów zapewniających wymaganą jakość usługi. Jest to możliwe do osiągnięcia dzięki zastosowaniu odpowiedniej architektury sieci i skojarzonych z nią mechanizmów QoS (ang. *Quality of Service*).

## Literatura

- [1] ITU-T Rec. G. 1010: *End-user multimedia QoS Categories*, 2001
- [2] ITU-T Rec. J. 241: *Quality of Service ranking and measurements methods for digital video services delivered over broadband IP networks*, 2005
- [3] Engelke U., Zepernick H.: *Perceptual-based Quality Metrics for Image and Video Services: A Survey*. 3rd Euro-NGI Conference on Next Generation Internet Networks, June 2007
- [4] Chikkerur S., Sundaram V., Reisslein M., Karam L. J.: *Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison*. IEEE Transactions on Broadcasting, Vol. 57, No. 2, Feb. 2011
- [5] Bhat A., Richardson I., Kannangara S.: *A New Perceptual Quality Metric for Compressed Video*. Signal Processing: Image Communication, Vol. 25, No. 8, Sep. 2010
- [6] Wang Y.: *Survey of Objective Video Quality Measurements*. Tech report, Worcester Polytechnic Institute, June 2006
- [7] [www.wanem.sourceforge.net](http://www.wanem.sourceforge.net) (*emulator WANem*), May 2011
- [8] [www.vqeg.org](http://www.vqeg.org) (*Video Quality Experts Group*), May 2011
- [9] ITU-T Rec. J. 247 *Objective Perceptual multimedia video quality measurements in the presence of a full reference*, 2008
- [10] Kula S., Dymarski P., Szmyd G.: *Wpływ parametrów sieci na jakość sygnałów wideo*. Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Sep. 2009
- [11] Dymarski P., Kula S.: *Metody i standardy badania postrzeganej jakości sygnałów audio i wideo*. Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Sep. 2009