

Możliwości prognozowania wpływu strumienia binarnego na wskaźnik PSNR w procesie wytwarzania płyt Blu-ray i DVD-Video

Rafał Kłoda*, Sabina Żebrowska-Łucyk**

*Przemysłowy Instytut Automatyki i Pomiarów PIAP

** Instytut Metrologii i Inżynierii Biomedycznej, Politechnika Warszawska

Streszczenie: W procesie produkcyjnym płyt DVD-Video i Blu-ray zakodowany automatycznie materiał filmowy podlega oglądowi przez operatora systemu kodującego w celu wykrycia miejsc wymagających poprawienia. Decyzja o tym, czy należy wykonać powtórne kodowanie oraz dobranie nowych parametrów kodowania pozostają w gestii operatora koderza. Współczesne systemy kodujące mają wbudowane narzędzie umożliwiające automatyczne wskazywanie fragmentów, które prawdopodobnie wymagają powtórnego zakodowania, jednak brakuje ciągle narzędzi wspierających operatora w kwestii doboru parametrów kodowania. Przedstawiona w artykule metoda prognozowania wpływu strumienia binarnego S na wartość wskaźnika PSNR pozwala operatorowi systemu, kodującego materiał filmowy w trybie wieloprzebiegowym na przewidywanie wartości, jaką osiągnie wskaźnik PSNR po powtórnym kodowaniu przy innej wartości strumienia binarnego. Przeprowadzony eksperyment potwierdził praktyczną użyteczność metody.

Słowa kluczowe: prognozowanie, PSNR, Blu-ray, DVD-Video, proces wytwarzania płyt

DOI: 10.14313/PAR_207/140

1. Wprowadzenie

Wytwarzanie płyt DVD-Video oraz Blu-ray jest wieloetapowym procesem produkcyjnym [3, 11, 12], a jednym z jego etapów jest kodowanie materiału filmowego. W fazie kodowania obraz źródłowy jest poddawany procesowi kompresji stratnej z parametrami ustalonymi wstępnie podczas sporządzenia budżetu bitów [3, 10–12]. Jest to ważny element procesu przygotowania płyt, gdyż informacja wizualna zajmuje największą część przestrzeni dyskowej nośnika a jakość obrazu jest istotnym elementem oceny produktu finalnego. Zaawansowane kodery – wymagane w takim procesie – wykorzystują złożone automatyczne techniki doboru parametrów kompresji [3, 10], a ponadto system kodujący pozwala operatorowi modyfikować parametry kodowania wybranych fragmentów materiału i kodować je powtórnie (ang. *segment-based re-encoding*) [2, 3, 10]. Etap kodowania materiału filmowego jest sprzężony z weryfikowaniem jakości materiału uzyskanego po kompresji. Z tego powodu istotną rolę odgrywa operator

systemu kodującego, którym jest wykwalifikowany specjalista w zakresie kodowania materiałów multimedialnych [14]. Operator systemu, w razie stwierdzenia niezadowalającej jakości fragmentów materiału, poddaje je zabiegom wstępnym (np. filtrowaniu) oraz tak modyfikuje automatycznie dobrane wartości parametrów kodowania, aby uzyskać lepszą jakość materiałów po ich ponownym zakodowaniu [2]. Wiedza i doświadczenie eksperckie operatora systemu kodującego mają znaczący wpływ na końcową jakość płyty.

Producenci współczesnych systemów kodujących, dostrzegając potrzebę ułatwienia pracy operatora, wprowadzają dodatkowe narzędzia do kontroli jakości procesu kodowania (ang. *Quality Control*). Umożliwiają one automatyczne wskazywanie fragmentów, które prawdopodobnie wymagają powtórnego zakodowania. Takie rozwiązanie zostało zastosowane m.in. w programowym koderze TotalCode Professional firmy Rovi [2], znanym też pod wcześniejszą nazwą Sonic CineVision. Koder umożliwia przygotowanie strumieni zgodnych z DVD-Video i Blu-ray w trzech standardach: MPEG-2 oraz H.264/MPEG-4 (AVC) i SMPTE 421M (VC-1). Proces kodowania rozpoczyna się od wczytania nieskompresowanego materiału filmowego w postaci pliku w formacie AVI, QuickTime lub Planar YUV (możliwe jest również wczytanie sekwencji obrazów w formacie DPX) oraz ustawienia parametrów kodowania obliczonych na podstawie budżetu bitów. Opcjonalnie materiał źródłowy można poddać procesom filtracji. Uruchomienie procesu kodowania rozpoczyna się przez wywołanie polecenia systemowego [2].

Głównymi czynnikami wpływającymi na końcową jakość obrazu w omawianym koderze są parametry *Target Bit Rate* oraz *Max Quantization*. Ustawiona wartość średnia strumienia binarnego (parametr *Target Bit Rate*) ma zawsze pierwszeństwo przed parametrem *Max Quantization*. Ten dodatkowy parametr (*Max Quantization*) pozwala na ustalenie maksymalnego poziomu kwantyzacji, który będzie używany przy kodowaniu. Jego niższe wartości zapewniają uzyskanie lepszej jakości obrazu lecz w większości przypadków ustawia się wartość tego parametru na 100 %, aby umożliwić koderowi używanie pełnej skali kwantyzacji [2].

Skutkiem kompresji stratnej obrazu są różnice między materiałem źródłowym a materiałem zdekodowanym [8].

Niektóre kodery, wśród nich TotalCode Professional, automatycznie obliczają wskaźnik *PSNR* (ang. *Peak Signal to Noise Ratio*), który jest miarą tych różnic. Wprawdzie jego wartość nie przekłada się jednoznacznie na stopień zniekształceń postrzeganych przez człowieka, lecz wiadomo, że jest on dodatnio skorelowany z percepcyjnym poziomem jakości [1, 6, 13, 15] i dlatego stanowi dla operatora wskazówkę przy wyszukiwaniu segmentów o niskiej jakości. Fragmenty te są poddawane powtórnemu kodowaniu po ustawieniu przez operatora innych parametrów kodowania, na ogół większej wartości strumienia binarnego. Źródła bibliograficzne informują o kierunku zmian wskaźnika *PSNR* w funkcji strumienia binarnego [4, 5, 7] (wzrost strumienia skutkuje wzrostem *PSNR*), jednak brakuje doniesień na temat możliwości przewidywania jak bardzo zmieni się wskaźnik *PSNR* po ustawieniu nowej wartości strumienia w zastosowaniu do tego typu procesu.

W Instytucie Metrologii i Inżynierii Biomedycznej przeprowadzono eksperymenty, które miały na celu zbadanie czy taka możliwość istnieje. Doprowadziły one do znalezienia prostej formuły matematycznej wiążącej zmianę strumienia ze spodziewaną zmianą wartości wskaźnika *PSNR*.

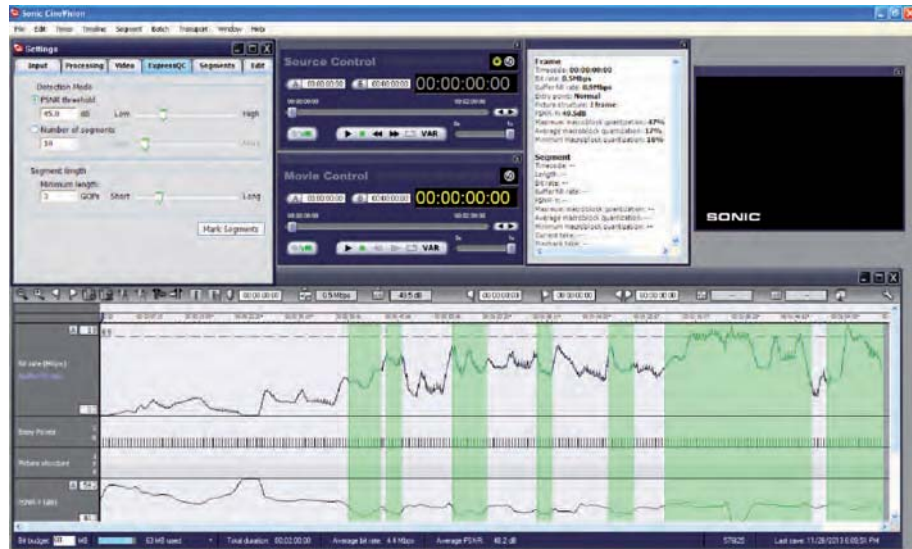
2. Eksperyment

Do przeprowadzenia eksperymentu wykorzystano fragment pełnometrażowego materiału filmowego w formacie PAL, który poddano procesowi kodowania na koderze TotalCode Professional w trybie zmiennego strumienia bitowego VBR (ang. *Variable Bit Rate*) ustalonego na średnim poziomie $S = 3,5 \text{ Mb/s}$. Zastosowano grupę obrazową GOP (ang. *Group of Pictures*) o długości 12 obrazów z dwiema ramkami typu B.

Zaimplementowany w koderze algorytm identyfikacji segmentów o najgorszej jakości wykorzystuje wskaźnik *PSNR*, którego dolna wartość progowa jest ustawiana przez operatora systemu. Zgodnie z definicją [9], wskaźnik ten jest stosunkiem maksymalnej możliwej mocy sygnału do mocy szumu zakłócającego wierność odwzorowania tego sygnału i wyznacza się go według wzoru:

$$PSNR = 10 \cdot \log_{10} \left(\frac{(2^b - 1)^2}{MSE} \right) \quad [dB] \quad (1)$$

w którym b oznacza liczbę bitów wykorzystywaną do reprezentacji składowej sygnału ($b = 8$), a MSE (ang. *Mean Squared Error*) to miara błędu średniokwadratowego wyznaczona na podstawie różnic pikseli w ramach obrazu materiału po procesie kodowania i ramach źródłowych.



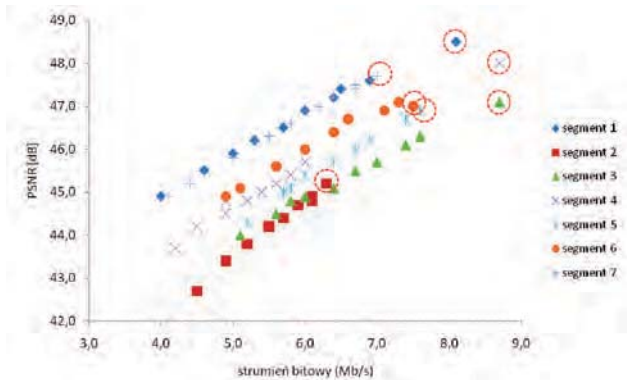
Rys. 1. Interfejs graficzny koder TotalCode Professional. Widoczna z lewej strony zakładka *ExpressQC* umożliwia identyfikację segmentów o najgorszej jakości. Są one wyróżnione kolorem zielonym na wykresie w dolnej części ekranu

Fig. 1. TotalCode Professional encoder GUI. On the left side *ExpressQC* tab for finding segments that require improvement. They are marked in green on the bottom part of the image above

Wartości wskaźnika *PSNR* dla badanego materiału filmowego uzyskane po procesie kodowania zawierały się od 39,8 dB do 58,2 dB. Wykorzystano wbudowane narzędzie *ExpressQC* do kontroli efektów kodowania (rys. 1) i dokonano za jego pomocą detekcji fragmentów o najgorszej jakości. Założono progową wartość wskaźnika $PSNR_{threshold} = 45 \text{ dB}$ oraz minimalną długość segmentu wynoszącą trzy długości GOP. Wartości te wprowadzono do odpowiednich komórek interfejsu programu kodującego (rys. 1).

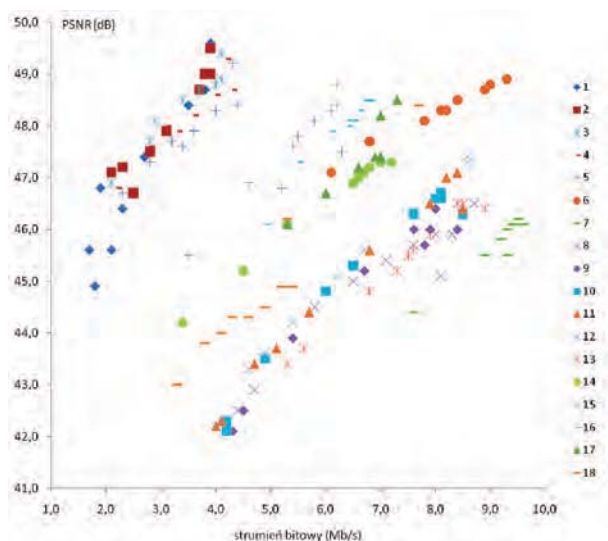
W efekcie działania algorytmu otrzymano zbiór siedmiu segmentów o różnym czasie trwania, dla których spełniona była nierówność $PSNR_{base\ code} \leq PSNR_{threshold}$. Na rys. 1 fragmenty te są oznaczone kolorem zielonym. Segment o najkrótszym czasie trwania wynosi nieco powyżej 2 s, a najdłuższy nieco ponad 22 s. Sumarycznie czas segmentów stanowił ponad 41 % długości całego badanego materiału. Wprowadzenie w dużej części materiału nowych wartości strumienia (często ustalanych przez operatora na maksimum) stwarza niebezpieczeństwo przekroczenia dopuszczalnych wartości wynikających z budżetu bitów i pojemności nośnika, dlatego jest celowe podjęcie badań nad opracowaniem formuły prognozującej skutki modyfikowania strumienia.

W celu pozyskania danych do analizy statystycznej każdy z segmentów kodowano 9-krotnie, każdorazowo ustawiając inną wartość średniego strumienia binarnego S z zakresu od 4,0 Mb/s do 8,7 Mb/s. Parametr kwantyzacji *Max Quantization* był, oprócz kilku przypadków, ustawiony na wartość domyślną 100 %.



Rys. 2. Wykresy średnich wartości wskaźnika $PSNR$ dla segmentów kodowanych na dziewięciu poziomach strumienia binarnego S . Punkty obwiedzione kółkiem odpowiadają wariantom kodowania, w których dodatkowo zmieniono parametr $Max\ Quantization$

Fig. 2. Plots of average values of $PSNR$ for segments coded on nine levels of bitrate S . The encircled points correspond to the coding with various value of the $Max\ Quantization$ parameter



Rys. 3. Wykresy średnich wartości $PSNR$ w funkcji strumienia binarnego S dla kolejnych grup obrazowych w segmencie 7. Numerami oznaczono kolejne GOP

Fig. 3. Plots of average $PSNR$ as a function of a bitrate S for individual GOPs in the segment 7. Numbers denote consecutive GOPs in the segment

3. Analiza wyników

W pierwszym etapie analizy danych sporządzono wykresy ilustrujące powiązanie średnich wartości $PSNR$ ze średnimi wartościami strumienia binarnego S dla każdego z siedmiu segmentów (rys. 2).

Jak widać, zwiększenie wartości strumienia bitowego S powoduje zwiększenie wartości wskaźnika $PSNR$ a układ punktów dla poszczególnych segmentów ma charakter prawie liniowy. Wykresy wyznaczone dla poszczególnych segmen-

tów są względem siebie wyraźnie przesunięte, ale współczynnik kierunkowy jest zbliżony.

Zbadano czy podobne prawidłowości można zaobserwować wewnątrz segmentów. W tym celu sporządzono analogiczny wykres, ale dla wycinków o długości tylko jednej grupy obrazowej. Wybrano segment nr 7 o długości ok. 8 s, który zawiera 19 GOP. Rezultaty zamieszczono na wykresie (rys. 3). Są one podobne do wyników otrzymanych przy analizie całych segmentów (rys. 2).

Przedstawione wykresy pozwalają zaproponować zależność, która umożliwi na podstawie wyników pierwszego kodowania wyznaczyć teoretyczną (prognozowaną) wartość wskaźnika $PSNR_{teoretyczny}$ odpowiadającą zmienionej wartości strumienia S , a mianowicie:

$$PSNR_{teoretyczny} = PSNR_{base\ encode} + k \cdot (S - S_{base\ encode}) \quad [dB] \quad (2)$$

gdzie: $S_{base\ encode}$ – strumień binarny zastosowany do pierwszego kodowania (bazowy), $PSNR_{base\ encode}$ – wskaźnik $PSNR$ otrzymany podczas pierwszego kodowania, k – współczynnik kierunkowy wyznaczony eksperymentalnie.

Na podstawie wyników uzyskanych z pierwszego doświadczenia wyznaczono wartość współczynnika k jako średnią arytmetyczną ze współczynników kierunkowych równań regresji dla każdego z siedmiu zbiorów punktów. Otrzymano wartość $k = 1,02$. Przy budowie modelu pominięto punkty odpowiadające wariantom kodowania, w których dodatkowo zmieniono parametr $Max\ Quantization$ (na rys. 2 są to punkty obwiedzione kółkiem).

Dokładność predykcji wskaźnika $PSNR$ wyznaczono poprzez wyliczenie według wzoru (2) wartości $PSNR_{teoretyczny}$ dla rzeczywistych średnich wartości strumienia binarnego S i wartości strumienia binarnego bazowego $S_{base\ encode}$, które następnie porównano ze znanymi wartościami wskaźnika $PSNR$ dla powtórnie zakodowanych segmentów, wyznaczając różnice:

$$\Delta PSNR = PSNR - PSNR_{teoretyczny} \quad (3)$$

Przeciętna wartość $|\Delta PSNR|$ w grupie wszystkich segmentów $\Delta PSNR$ nie przekroczyła 0,1 dB, a maksymalna wartość $|\Delta PSNR|$ wyniosła 0,48 dB.

4. Podsumowanie

W procesie produkcyjnym płyt DVD-Video i Blu-ray zakodowany automatycznie film podlega oglądowi przez operatora systemu kodującego, w celu wykrycia miejsc wymagających poprawienia. Decyzja o tym czy należy wykonać powtórne kodowanie oraz dobranie nowych parametrów kodowania pozostają w gestii operatora koder. Niektóre kodery, w tym TotalCode Professional stosowany do opisanych w artykule badań, mają wbudowane narzędzie umożliwiające automatyczne wskazywanie fragmentów, które prawdopodobnie wymagają powtórnej zakodowania, jednak brakuje ciągle narzędzi wspierających operatora w kwestii doboru parametrów kodowania.

Badania nad ustaleniem zależności między wartością wskaźnika $PSNR$ a strumieniem binarnym prowadzą liczne

ośrodki. Niektóre doniesienia wskazują na liniowy charakter takiego związku, jednak nie było dotychczas publikacji przedstawiających model, który mógłby znaleźć zastosowanie w procesie wytwarzania płyt cyfrowych.

Istotną nowością w przedstawionym podejściu jest prognozowanie wpływu strumienia binarnego S na wartość wskaźnika $PSNR$ bez konieczności uwzględnienia wielu parametrów związanych z obrazem, a bazowanie jedynie na efektach pierwszego etapu kodowania oraz wartości współczynnika k charakteryzującego koder i ustawione w nim parametry. Przedstawiona metoda pozwala operatorowi systemu kodującego materiał filmowy w trybie wieloprzebiegowym na przewidywanie wartości, jaką osiągnie wskaźnik $PSNR$ po powtórnym kodowaniu przy innej wartości strumienia binarnego. Zastosowanie podanej w artykule zależności (2) daje więc operatorowi narzędzie do szybkiego ustalenia czy jest możliwe takie zwiększenie strumienia binarnego aby wartość wskaźnika $PSNR$ przekroczyła ustaloną wartość progową.

Przedstawiony sposób postępowania umożliwia bardzo łatwe obliczenie spodziewanych wartości $PSNR$. Podany model matematyczny cechują prostota i uniwersalność. Przeprowadzony eksperyment potwierdził praktyczną użyteczność modelu. Wartość współczynnika k można wyznaczyć empirycznie dla dowolnego systemu kodującego i zakresów strumienia binarnego stosowanych w danym procesie kodowania.

Bibliografia

1. Avcibas I., Sankur B., Sayood K., *Statistical evaluation of image quality measures*, Journal of Electronic Imaging, 11(2), 2002, 206–223.
2. *CineVision User Guide*, Sonic Solutions, 2011.
3. Ely M., Block D., *Publishing in the Age of DVD*, Sonic Solutions, 1998.
4. *HDTV Contribution Codecs technology evaluation*, EBU Technical Report 008, 2010.
5. Huynh-Thu Q., Ghanbari M., *The accuracy of PSNR in predicting video quality for different video scenes and frame rates*, Telecommunication Systems, vol. 49, 2012, 35–48.
6. Le Callet P., Viard-Gaudin C., Pechard S., Caillaud E., *No reference and reduced reference video quality metrics for end to end QoS monitoring*, IEICE Trans. Commun., vol. E89-B, no. 2, 2006.
7. Maćkowiak S., Domański M., *Badania eksperymentalne kodeka MPEG cyfrowego sygnału telewizyjnego*, Krajowe Sympozjum Telekomunikacji, 1998, 190–198.
8. Ostaszewska A., Żebrowska-Łucyk S., *Filtracja sygnału w metodach ciągłej oceny błędów kompresji materiału filmowego na rzecz automatyzacji procesu monitorowania jakości*, „Pomiary Automatyka Robotyka”, nr 1, 2009, 10–13.
9. Przelaskowski A., *Kompresja danych*, [www.ire.pw.edu.pl/~arturp/Dydaktyka/koda/skrypt.html], 2002.
10. *Sonic DVD Creator User Guide*, Sonic Solutions, 1999.
11. Taylor J., Johnson M.R., Crawford C.G., *DVD Demystified*, Third Edition, McGraw-Hill, 2006.
12. Taylor J., Zink M., Crawford C.G., Armbrust C.M., *Blu-ray Disc Demystified*, McGraw-Hill, 2008.
13. Wang Z., Sheikh H.R., Bovik A.C., *No-reference perceptual quality assessment of JPEG compressed images*, Proc. IEEE Int. Conf. Image Proc, 2002.
14. Westwater R., Talisman M., *Subjective quality monitoring*, Conf. HPA Technology Retreat, 2004.
15. Winkler S., Mohandas P., *The evolution of video quality measurement: from PSNR to hybrid metrics*, IEEE Transactions on Broadcasting 54 (3), 2008, 660–668. ■

The possibility of forecasting the impact of the bitrate on the PSNR value in the production of Blu-ray and DVD-Video discs

Abstract: During the DVD-Video and Blu-ray discs production process, the film encoded automatically is always inspected by the human operator, in order to detect fragments that require improvement. The decision of whether to perform a re-encoding as well as selection of new encoding parameters are up to the operator. Although modern encoding systems have built-in tool for automatically identifying fragments that are likely to require re-encoding, but still lacks tools to support the operator with setting new encoding parameters. The method presented in the paper allows the operator of the video encoding system to predict the value that the PSNR quality metric will reach after re-encoding video material with a chosen bit rate. Experiments confirmed the utility of the method.

Keywords: video quality forecasting, PSNR, Blu-ray, DVD-Video, mastering, bit

Artykuł recenzowany; nadesłany 14.01.2014 r., przyjęty do druku 17.03.2014 r.

mgr inż. Rafał Kłoda

Specjalista w zakresie masteringu płyt Blu-ray oraz DVD-Video. Zajmuje się badaniami nad oceną jakości obrazu oraz uczestniczy w wielu projektach związanych z nowymi technologiami w dziedzinie multimedialnych.

e-mail: kloda@mchtr.pw.edu.pl



dr hab. Sabina Żebrowska-Łucyk, prof. PW

Prowadzi badania naukowe i kieruje projektami z zakresu metrologii oraz inżynierii jakości z uwzględnieniem technik multimedialnych. Zajmuje się też zastosowaniami metod statystycznych w badaniach naukowych i praktyce laboratoryjnej.

e-mail: szl@mchtr.pw.edu.pl

