

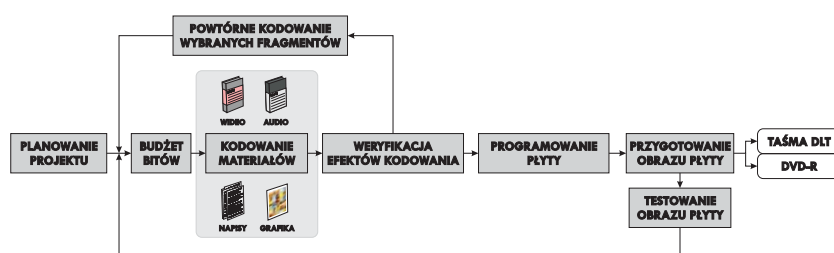
Metoda i system automatycznej detekcji błędów spowodowanych kompresją cyfrową w procesie wytwarzania płyt DVD-Video

Rafał Kłoda
Sabina Żebrowska-Łucyk

Widoczne błędy kompresji powstające przy kodowaniu materiału filmowego, dyskwalifikują ten produkt z rynku, dlatego zapobieganie takim usterkom ma istotne znaczenie – szczególnie dla producentów płyt DVD-Video. W artykule przedstawiono propozycję metody i systemu do automatycznego wykrywania błędów kompresji w procesie premasteringu DVD. Zaproponowano również algorytm prognozujący jakość wizualną, który zostanie wykorzystany w planowanym systemie do wskazania miejsc o najgorszej jakości w materiale filmowym. Wskazane przez algorytm obszary zostaną zweryfikowane i poprawione przez operatora systemu kodującego.

Do kodowania materiałów filmowych umieszczanych na płytach DVD-Video stosuje się algorytm kompresji MPEG-2. Jest to algorytm kompresji stratnej, w wyniku którego następuje znaczne zmniejszenie objętości plików, ale pojawiają się błędy kompresji, takie jak: blokowa struktura obrazu, pulsowanie bloków, artefakty na krawędziach (przypadkowe kropki i kreski nazywane też „fruwającymi komarami”). Dostrzegalne przez widza błędy kompresji stanowią usterkę produktu. Ponieważ płyty DVD-Video należą do najpopularniejszych wyrobów rynkowych, zapobieganie takim usterkom ma istotne znaczenie zarówno dla producentów, jak i nabywców płyt. Dlatego ważnym elementem procesu wytwarzania płyt DVD-Video jest kontrola jakości efektów kodowania.

Przygotowanie materiałów audiowizualnych zgodnie z przyjętymi normami i standardami oraz umieszczenie ich na obrazie dysku [1, 2] nosi nazwę procesu premasteringu DVD. Wytworzony w tym procesie obraz dysku powielany jest w procesie tłoczenia. Proces technologiczny premasteringu DVD można podzielić na etapy (rys. 1): planowanie projektu, budżet bitów, kodowanie materiałów audiowizualnych, programowanie płyty DVD-Video, przygotowanie obrazu oraz testowanie, którego ważnym elementem jest weryfikacja jakości skompresowanego materiału. Poszczególne etapy mogą być realizowane w różnych wyspecjalizowanych wytwórniach, ale integracja i kontrola tych etapów w jednym miejscu czyni cały proces bardziej



Rys. 1. Schemat procesu premasteringu DVD

skutecznym i wydajnym. Przykładem takiego systemu premasteringowego jest Sonic DVD Creator™.

Profesjonalne systemy kodujące pozwalają na różne tryby pracy kodera, w zależności od czasu przeznaczanego na kodowanie. Parametry kodowania – zwłaszcza strumień bitowy (*bitrate*) wynikają z budżetu bitów. W celu uzyskania optymalnej jakości długich materiałów filmowych, koder wykorzystuje technikę kodowania ze zmiennym strumieniem bitowym (VBR). Polega ona na utrzymywaniu średniej wartości strumienia bitowego i podwyższanie lub obniżanie jego poziomu, w celu dopasowania wartości do złożoności materiału źródłowego. Standard dopuszcza chwilowe zwiększanie strumienia bitowego w krótkich przedziałach czasowych do maksymalnej wartości $V_{bmax} = 9,8$ Mbps. Takie dostosowywanie strumienia zapobiega powstawaniu widzialnych błędów kompresji w obszarach trudnych do kodowania (o dużej dynamice i wysokich częstotliwościach przestrzennych), a przekroczenie ustalonej wartości średniej strumienia bitowego jest kompensowane przez jego obniżenie przy kodowaniu mniej wymagających fragmentów.

Maksymalne wykorzystanie możliwości kodera, to zastosowanie kodowania VBR z wieloprzebiegową analizą materiału źródłowego przed kompresją. Przy takiej konfiguracji, kodowanie odbywa się etapami. Pierwszy etap to analiza materiału źródłowego. Koder dokonuje przeglądu materiału w celu detekcji zmian scen i znalezienia miejsc wymagających obecności

mgr inż. Rafał Kłoda, Przemysłowy Instytut Automatyki i Pomiarów, Instytut Metrologii i Inżynierii Biomedycznej PW
dr hab. inż. Sabina Żebrowska-Łucyk, prof. nzw. PW, Instytut Metrologii i Inżynierii Biomedycznej, Politechnika Warszawska

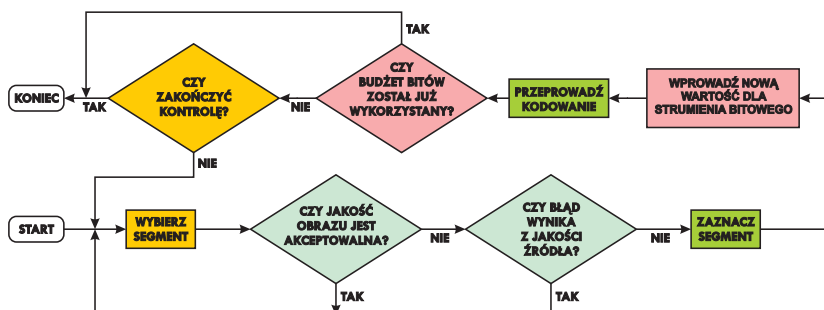
dotatkowych ramek oraz rejestruje wahania wartości strumienia bitów w zależności od stopnia złożoności obrazu. W następnym etapie koder wykonuje wielopięciową analizę wybranych wcześniej z materiału źródłowego krótkich fragmentów, o największym i najmniejszym stopniu złożoności obrazu. Rezultatem tych działań jest profil strumienia bitowego dla analizowanego materiału z przydzieloną większą liczbą bitów dla trudnych do kodowania obszarów i mniejszą dla mniej wymagających fragmentów. Na ostatnim etapie analizy, koder próbuje dopasować średnią i maksymalną wartość strumienia bitowego, która wynika z uwarunkowań budżetu bitów. Proces ten odbywa się stopniowo poprzez regulację skali kwantyzacji i stopniowe jej zawężanie, aby jednocześnie zachować profil i możliwie najdokładniej zbliżyć się do wartości strumienia bitowego wynikającej z parametrów kodowania. Po zakończeniu tych operacji następuje właściwy proces kodowania materiału filmowego.

Procedura wykrywania błędów w typowym procesie wytwarzania

Nawet taki złożony proces kodowania nie gwarantuje zadowalającej jakości całego materiału. Wykorzystanie miar obiektywnych do wskazania miejsc o najgorszej jakości nie jest skuteczne, gdyż znane wskaźniki są słabo skorelowane z jakością wizualną, dlatego konieczne jest sprawdzenie skompresowanego filmu przez operatora.

Zastosowanie ciągłych w czasie, percepcyjnych metod oceny jakości kompresji do uzyskania kompletnej informacji o jakości materiału nie jest stosowane, gdyż byłoby to trudne do realizacji i kłopotliwe w użyciu, ze względu na długość typowych materiałów filmowych umieszczanych na płytach DVD-Video.

W związku z tym standardem postępowania przy kontroli jakości tego etapu procesu technologicznego jest obecnie wyrywkowy ogląd zakodowanego materiału, w celu wychwycenia miejsc o najgorszej jakości i – w przypadku stwierdzenia miejsc o nieakceptowalnej jakości – ręcznego wprowadzenia korekt parametru kodowania (rys. 2).



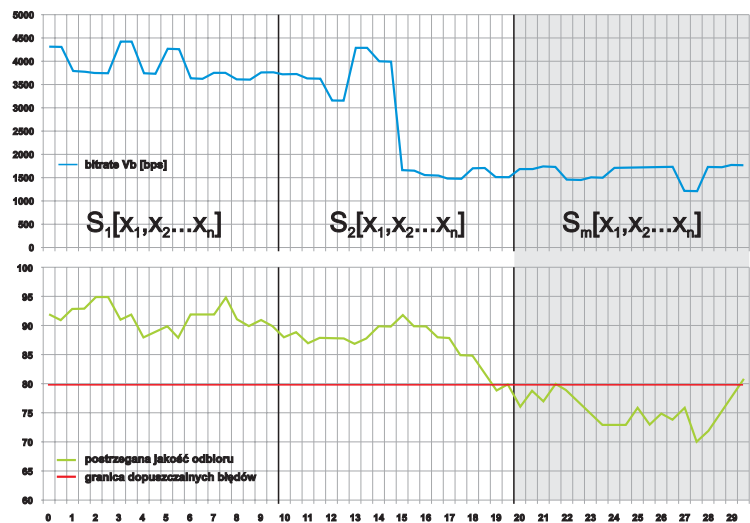
Rys. 2. Procedura kontroli jakości kodowania

Procedura przeglądu materiału wymaga szczególnej uwagi, wiedzy i skupienia osoby odpowiedzialnej za ten etap procesu. Co więcej, absorbuje ona czas pracy koderów potrzebny na przeprowadzenie kontroli, wstrzymując tym samym cały proces technologiczny. W efekcie, po powtórnym kodowaniu, następuje wprowadzenie lepsze wykorzystanie pasma w obrębie dostępnego jeszcze budżetu bitów, jednak część usterek pozostaje niewykryta.

W Instytucie Metrologii i Inżynierii Biomedycznej podjęto badania, których celem jest opracowanie metody pozwalającej na automatyczne wykrywanie usterek powstałych w procesie wytwarzania płyt DVD-Video i związanych z kompresją materiału filmowego.

Metoda i system wykrywania usterek wywołanych kompresją

Opracowana metoda detekcji usterek opiera się na prognozowaniu ocen jakości wizualnej na podstawie zbioru parametrów obliczeniowych charakteryzujących: a) materiał oryginalny (np. TI, SI itp.), b) parametry kompresji (Vb, długość GOP, liczba ramek B) oraz c) różnice



Rys. 3. Idea proponowanej metody

między materiałem oryginalnym i skompresowanym (np. PSNR, MSE, SSIM itp.). Eksperyment prowadzący do wyznaczenia algorytmu prognozującego ocenę opisano w dalszej części.

Drugą istotną cechą proponowanej metody detekcji usterek jest podział materiału filmowego na równomierne segmenty (S). Każdy z segmentów jest opisany wyżej wymienionymi wskaźnikami numerycznymi ($x_1, x_2, \dots, x_n$), które stanowią parametry wejściowe dla algorytmu prognozującego (rys. 3). Efektem działania algorytmu jest przypisanie segmentom rang odpowiadających przewidywanej jakości i wydzielenie zbioru segmentów o prognozowanym najniższym poziomie jakości.

Segmenty te są następnie wyświetlone ekspertowi, który je ocenia. Sposób wyświetlania segmentów jest zbliżony do metod stosowanych w punktowych badaniach percepcyjnych [8]. Sposób ułożenia i zawartość powstałej sesji zależą od przypisanych rang, dostępnego budżetu bitów i czasu przewidzianego na kontrolę jakości.

Metoda wyznaczenia parametrów algorytmu prognozującego jakość wizualną

W celu wyznaczenia parametrów algorytmu prognozującego (modelu prognozowanej jakości wizualnej), przeprowadzono serię eksperymentów, w których wykorzystano zbiór czterech sekwencji testowych z kodowanych z różnymi parametrami kompresji. Na specjalnie zbudowanym stanowisku pomiarowym [4], przeprowadzono badania z udziałem 52 obserwatorów, którzy dokonywali oceny wyświetlanego materiału zgodnie z metodą *Degradation Category Rating* (DCR) według zalecenia ITU-T [8]. Następnie wyznaczono średnie ocen obserwatorów (*Mean Opinion Score* MOS) dla poszczególnych sekwencji i parametrów kodowania.

Wykorzystując analizę regresji wielorakiej, która pozwoliła badać związki MOS z całym zbiorem wybranych wskaźników numerycznych, opracowano model matematyczny opisujący związek między wyselekcjonowanymi parametrami obliczeniowymi a średnią oceną obserwatorów. Model ten zawiera parametry należące do dwóch kategorii metryk: a) z pełnym odniesieniem oraz b) bez odniesienia. W grupie metryk z pełnym odniesieniem znalazły się parametry PSNR i SSIM. Parametry te bazują na porównaniu obrazu źródłowego i poddanego kompresji. Powszechnie znany parametr PSNR [5] mierzy różnicę pomiędzy dwoma obrazami (każdy o n próbkach) zgodnie z zależnością:

$$PSNR = 10 \log_{10} \frac{255^2 n^2}{\sum_{i=1}^n (O_i - Z_i)^2} \quad [dB]$$

gdzie O_i , Z_i – luminancja pikseli w obrazie źródłowym i po kompresji. Im większa wartość PSNR, tym większy jest stosunek sygnału do szumu (tym lepiej). SSIM wykorzystuje znacznie bardziej skomplikowaną procedurę obliczeń [6].

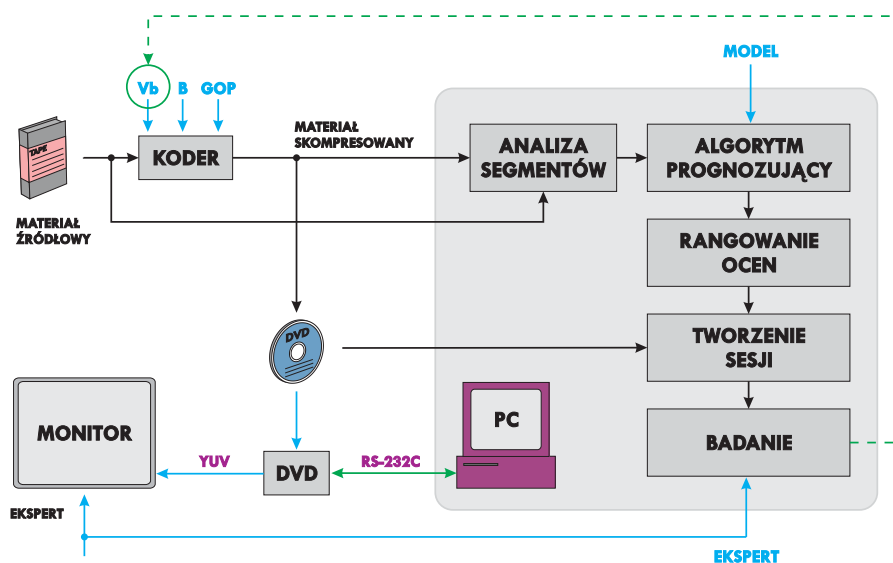
Z rodziny metryk bez odniesienia do modelu weszły parametry SI_q i TI_q [7]. Parametry z tej grupy bazują tylko na analizie obrazu po kompresji. Są one wynikiem niewielkiej modyfikacji parametrów SI i TI , znanych z literatury [8], służących do opisu charakteru materiału. Parametr SI_q opisuje liczbę szczegółów w materiale. Do wyliczenia tego parametru zastosowano krawędziowy filtr Sobela, złożony z dwóch masek 3×3 . Jedna maska jest przeznaczona do wykrycia poziomych krawędzi obrazu, druga do pionowych. Analiza obrazu dotyczy wartości luminancji, zgodnie z zależnością:

$$SI_q = \text{upper quartile} \left[\text{std} \left(\text{Sobel} (F_n) \right) \right]$$

Parametr TI_q opisuje chwilowe zmiany w materiale, czyli określa jego dynamikę i jest obliczany jako różnica wartości luminancji poszczególnych pikseli sąsiednich ramek, zgodnie z zależnością:

$$TI_q = \text{upper quartile} \left[\text{std} (F_n - F_{n-1}) \right]$$

Wprowadzenie modyfikacji, polegających na zastąpieniu wartości maksymalnych kwartylem górnym, zminimalizowało wpływ zmiany ujęć na wartości parametrów, co jest jedną z wad metryk TI i SI . Wartości tych parametrów wyliczone bezpośrednio z definicji posłużyły do analizy i selekcji typowych materiałów wizualnych i w końcowym efekcie doprowadziły do wyboru czterech sekwencji, pokrywających szeroki przekrój zmienności parametrów TI i SI .



Rys. 4. Stanowisko do kontroli jakości z wykorzystaniem proponowanej metody

W wyniku zastosowania analizy regresji wielorakiej otrzymano następujący model:

$$MOS = 0,003 \cdot TI_q \cdot SI_q \cdot SSIM + 0,358 \cdot PSNR - 11,4$$

Wszystkie parametry modelu są statystycznie wysoce istotne ($p < 0,00018$), a sama funkcja dobrze dopasowana ($R^2 = 0,8537$). Model ten zweryfikowano empirycznie prognozując oceny dla kilku innych sekwencji i stwierdzono, że pozwala trafnie przewidywać oceny obserwatorów.

Podsumowanie

Uzyskane dotąd wyniki prac podjętych w Instytucie Metrologii i Inżynierii Biomedycznej, tworzą podstawy do wprowadzenia znacznych usprawnień w procesie kontroli jakości płyt DVD-Video. Proponowany system nie tylko wspomaga proces detekcji, ale również pozwala odseparować proces kontroli od procesu kodowania a to umożliwia bardziej efektywne wykorzystanie systemu premasteringowego. Dodatkową korzyścią wynikającą z zastosowania omawianego systemu będzie generowanie raportów zawierających zapis zmian wprowadzonych przez operatora.

Dalszym etapem prac będzie opracowanie metody automatycznego wprowadzania korekt w materiale filmowym, na podstawie analizy wyników z przeprowadzonej kontroli.

Praca naukowa finansowana ze środków na naukę w latach 2007-2009 jako projekt badawczy.

Bibliografia

1. ISO/IEC 13818-1:2007 *Generic coding of moving pictures and associated audio information: Systems*.
2. Ely M., Block D.: *Publishing in the Age of DVD*, Sonic Solutions, 1998.
3. Kłoda R., Żebrowska-Łucyk S., Ostaszewska A.: *A dedicated station for MPEG-2 quality evaluation with ACR method*, International PhD Conference, Republika Czeska, Pilzno 2007.
4. Kłoda R., Ostaszewska A.: *Subjective video quality evaluation: an influence of a number of subjects on the measurement stability*, Recent Advances in Mechatronics, Springer Verlag, Berlin Heidelberg 2007, pp. 611-615.
5. Fisher Y.: *Fractal Image Compression*, Springer Verlag, 1995.
6. Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli: *Image quality assessment: From error visibility to structural similarity*, IEEE Transactions on Image Processing, 2004.
7. Ostaszewska A., Kłoda R.: *Quantifying the amount of spatial and temporal information in video test sequences*, Recent Advances in Mechatronics, Springer Verlag, Berlin Heidelberg 2007, pp. 11-15.
8. ITU-T Recommendation P.910 (1996), *Subjective video quality assessment methods for multimedia applications*.

