

Andrzej POPŁAWSKI

UNIwersytet Zielonogórski, Instytut Informatyki i Elektroniki, ul. Podgórna 50, 65-246 Zielona Góra

Optymalny dobór parametrów koderów falkowych ze względu na czas kodowania

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

W artykule przedstawiono wyniki badań eksperymentalnych falkowego koderów sekwencji wizyjnych zorientowane na optymalny dobór parametrów koderów ze względu na czas kodowania. Celem badań było wskazanie takich wartości parametrów kodowania, dla których można uzyskać możliwie największą wartość współczynnika PSNR przy założonym – zredukowanym czasie kodowania. Wykazano, że możliwa jest redukcja czasu kodowania przy stosunkowo niewielkiej utracie wartości współczynnika PSNR. Badania przeprowadzono na reprezentatywnej grupie testowych sekwencji wizyjnych, stosowanych i akceptowanych w wielu środowiskach naukowych.

Słowa kluczowe: falkowy koder sekwencji wizyjnych, skalowalność

Selection of wavelet video codec parameters to optimize a time of coding

Abstract

In the paper there are presented results of experimental research on wavelet video codec. Author put emphasis on optimal selection of the coder parameters to optimize a time of coding. Author used six test sequences (Basket, City, Crew, Harbour, Ice and Soccer in 352x288, 30 fps) for experiments. Sequences used in research are commonly used to assess the video sequences compression effectiveness. Measurement of image quality is based on PSNR. There are presented results of optimal selection of wavelet video codec parameters due to the time of coding. Main goal of this paper is to show the coder parameters that provides the best values of PSNR for reduced coding time.

Keywords: wavelet video coding, scalability.

1. Wprowadzenie

Kompresja danych wizyjnych jest w dniu dzisiejszym powszechnie wykorzystywana w wielu urządzeniach i zastosowaniach, począwszy od prostych urządzeń rejestrujących obrazy, przez telefony komórkowe, kamery i aparaty cyfrowe (amatorskie i profesjonalne), aż do naziemnych i satelitarnych cyfrowych systemów przekazywania obrazów (DVB-T, DVB-S). Urządzenia umożliwiające rejestrowanie i odtwarzanie tych obrazów w większości przypadków wyposażone są w jednostki przetwarzające, zapewniające obróbkę napływających danych w czasie rzeczywistym. Tańsze urządzenia zazwyczaj posiadają mniej wydajne jednostki przetwarzające, co wiąże się z koniecznością stosowania technik kompresji mniej efektywnych, a tym samym o niższym koszcie numerycznym.

W ostatnich latach powstało wiele standardów kodowania (np.: MPEG-4, H.263, H.264), wykorzystujących techniki kodowania hybrydowego. Wraz z rozwojem standardów i wzrostem efektywności kompresji, rosną również wymagania wydajnościowe jednostek obliczeniowych, przez co pewne rozwiązania nie mogą zostać zastosowane w urządzeniach o niewystarczającej mocy obliczeniowej. W procesie kodowania hybrydowego wykorzystywana jest dyskretna transformacja kosinusowa (DCT – discrete cosine transform) [1]. Na przestrzeni ostatnich lat wielu badaczy skupiło swoją

uwagę na kodery wizyjne wykorzystujące dyskretną transformację falkową (DWT – discrete wavelet transform) [2, 3, 4]. Główną przyczyną, dla której nastąpił wzrost zainteresowania tą techniką kompresji, są zalety koderów falkowych. Do najważniejszych z nich można zaliczyć:

- osadzony strumień bitowy, umożliwiający progresywne dekodowanie,
- naturalna skalowalność w dziedzinie przestrzennej,
- naturalna skalowalność w dziedzinie czasu,
- łatwość uzyskania ściśle określonej prędkości transmisji.

Wymienione zalety koderów falkowych mają szczególne znaczenie np. w systemach bezprzewodowych o zmieniającej się w czasie przepustowości kanału transmisyjnego, a także w transmisji sekwencji wizyjnych poprzez sieci komunikacyjne o niejednorodnej strukturze (tzw. sieci heterogeniczne) [5].

2. Falkowy koder sekwencji wizyjnych

W celu uzyskania efektywnej kompresji w falkowych koderach sekwencji wizyjnych, obrazy wejściowe w pierwszej kolejności są grupowane w tzw. grupy obrazów, a następnie wspólnie przetwarzane. Rozmiar grupy obrazów (GOP) ustalany jest na podstawie założonej liczby poziomów analizy falkowej w dziedzinie czasu i wpływa bezpośrednio na złożoność i efektywność kodowania. Wielkość grupy obrazów wynosi 2^k , gdzie k jest liczbą poziomów analizy falkowej w czasie [6].

Największą część czasu kodowania koderów falkowych (około 80%) pochłania estymacja ruchu. Podstawowym elementem predykcji z kompensacją ruchu w koderach falkowych jest podział obrazu na mniejsze elementy, tzn. obraz sekwencji dzielony jest według ustalonego algorytmu na tak zwane bloki. Następnie wyznaczane są wektory ruchu dla każdego bloku obrazu, tak aby dany blok po przesunięciu o wyznaczony wektor ruchu był możliwie najbardziej zbliżony do obrazu odniesienia. Rozmiar bloku może mieć rozmiar od 64×64 do 4×4 . Im mniejszy jest rozmiar bloku, tym większą liczbę wektorów ruchu należy wyznaczyć. Wektory ruchu wyznaczane są niezależnie na każdym poziomie dekompozycji w czasie, tak więc im więcej poziomów dekompozycji w dziedzinie czasu, tym większy nakład pracy wymagany jest na przeprowadzenie procedury estymacji ruchu.

W celu wykonania estymacji ruchu z dokładnością większą niż jeden odstęp próbkowania (czyli z dokładnością większą niż jeden punkt) w pierwszej kolejności wykonywana jest estymacja ruchu z dokładnością jednego punktu. Następnie poprzez dokonanie interpolacji obrazów, dwukrotnie zwiększa się ich rozmiary przestrzenne i dokonuje poszukiwań nowego wektora ruchu w sąsiedztwie już wcześniej wyznaczonego.

Pośród wielu parametrów wpływających na przebieg kodowania, cztery z nich mają kluczowe znaczenie z punktu widzenia złożoności oraz efektywności procesu kodowania. W zależności od ich wartości rośnie lub maleje złożoność obliczeniowa, a tym samym czas potrzebny na zakodowanie danej sekwencji wizyjnej. Poniżej scharakteryzowano znaczenie tych parametrów [7].

tPL – określa liczbę poziomów dekompozycji falkowej w dziedzinie czasu. Wartość tego parametru zazwyczaj wynosi 2 do 5. Większa wartość parametru zazwyczaj oznacza większą efektywność kodowania oraz wzrost złożoności obliczeniowej wraz ze wzrostem tej wartości.

MV – oznacza dokładność kodowania wektorów ruchu. Możliwe do wykorzystania wartości to: 1, 1/2, 1/4, 1/8 odstępu próbkowania. Dokładniejsze kodowanie wektorów ruchu zwiększa dokładność estymacji ruchu.

BM – określa stopień podziału bloku na podbloki. Standardowo rozmiar bloku obrazu wynosi 64×64 i może przyjmować wartości: 64×64 , 32×32 , 16×16 , 8×8 , 4×4 . Przykładowo dla wartości parametru równego 3, rozmiar podbloków jest nie mniejszy niż 16×16 punktów. Dla każdego podbloku wyznaczany jest wektor ruchu,

stąd większa liczba mniejszych podbloków powoduje konieczność wyznaczenia większej liczby wektorów ruchu – co jest czynnością czasochłonną. Wartość tego parametru: 3, 4, 5.

SR – określa obszar poszukiwań bieżącego bloku o obrazie odniesienia (wyrażone w punktach) dla pierwszego poziomu dekompozycji w czasie. Wartość parametru jest podwajana na każdym kolejnym poziomie dekompozycji czasowej. Zwiększenie wartości parametru z jednej strony powoduje znaczny wzrost złożoności obliczeniowej, z drugiej zaś może przyczynić się do dokładniejszego wyznaczenia wektora ruchu dla danego bloku. Wartość tego parametru to 8 lub 16 punktów.

3. Badania eksperymentalne

Celem badań było określenie takich wartości parametrów kodowania, dla których można uzyskać istotne zredukowanie czasu kodowania (czasu przetwarzania danych np. przez procesor) przy możliwie małym spadku efektywności kompresji. Rozwiązania prezentowane w literaturze skupiają się w głównej mierze na ograniczeniu opóźnienia kodowania, wprowadzanego przez zespół filtrów analizy/syntezy w dziedzinie czasu [8]. Polegają na takiej modyfikacji układu filtrów, aby zapewnić mniejszą wartość opóźnienia kodowania, co zasadniczo nie wpływa na czas kodowania związany z przetworzeniem danych przez CPU.

W pracy jako T oznaczono czas kodowania, dla którego uzyskano największą wartość współczynnika PSNR. Następnie wyznaczono najkorzystniejsze parametry kodowania dla czasu kodowania wynoszącego odpowiednio 50% i 25% czasu T (1/2T oraz 1/4T). Do badań wykorzystano zestaw sześciu sekwencji wizyjnych (Basket, City, Crew, Harbour, Ice, Soccer) o rozdzielczości przestrzennej CIF (352x288) i częstotliwości występowania obrazów równej 30Hz. Sekwencje te są szeroko akceptowane i stosowane przez innych badaczy do oceny efektywności kompresji sekwencji wizyjnych, charakteryzując się zróżnicowaną dynamiką, zarówno na pierwszym jak i na drugim planie.

Eksperyment polegał na zakodowaniu a następnie zdekodowaniu testowych sekwencji dla założonych wartości parametrów kodowania:

- tPL – wartości 2, 3, 4, 5,
- MV – wartości 1/2, 1/4, 1/8,
- MB – wartości 3, 4, 5,
- SR – wartości 8, 16.

W czasie kodowania dokonywano pomiaru czasu kodowania, eksperymenty wykonano na komputerze PC wyposażonym w procesor Intel Core i7 2700K. Dla każdej sekwencji wizyjnej wykonano 72 eksperymenty (w sumie 432), dokonując sprawdzenia wszystkich możliwych wariantów kodowania.

4. Wyniki badań eksperymentalnych

W pierwszym etapie analizy uzyskanych wyników badań eksperymentalnych dokonano zestawienia wpływu poszczególnych parametrów kodowania na czas kompresji, pomijając zagadnienia związane z jakością kompresji. Największy wpływ na zwiększenie czasu kodowania ma parametr tPL, dla którego zmiana wartości z 2 na 5 przekłada się na zwiększenie czasu kodowania w zależności od badanej sekwencji średnio od 213 do 427%. Zmiana parametru SR z wartości 8 na wartość 16 powoduje zwiększenie czasu kodowania średnio od 64 do 106%. Zmiana pozostałych parametrów ma już stosunkowo niewielki wpływ na czas kodowania. Dla parametru MV, jego zmiana z 1/2 na 1/8 powoduje zwiększenie czasu kodowania średnio o 11%, zaś zmiana parametru MB z 3 do 5 powoduje zwiększenie czasu kodowania średnio o 13%.

Następnym etapem był analiza oraz wybór takich parametrów kodera, przy których spadek efektywności kodowania jest możliwie najmniejszy. W tabeli 1 przedstawiono optymalne (ze względu na współczynnik PSNR) wartości parametrów kodowania dla zredukowanego czasu kodowania równego 1/2T.

Z wyników przedstawionych w tabeli wynika, że w większości przypadków dla parametrów: tPL=4, MV=1/4, MB=4, SR=8 czasy kodowania skracają się o 50% w stosunku do najdłuższych czasów kodowania. W tabeli 2 przedstawiono wartości współczynnika

PSNR badanych sekwencji wizyjnych dla czasów równych T oraz 1/2T dla parametrów jak w tab. 1 i prędkości transmisji 512kb/s W każdym badanym przypadku nastąpił spadek wartości współczynnika PSNR, największy o 0,84dB w przypadku sekwencji City, najmniejszy o 0,09dB w przypadku sekwencji Ice, średnio o 0,27dB. Jednak poza sekwencją City, spadek ten jest stosunkowo niewielki jeżeli wziąć pod uwagę fakt, że czas kodowania uległ skróceniu o 50%.

Tab. 1. Optymalne wartości parametrów kodera dla czasu kodowania 1/2T
Tab. 1. Optimal values of coder parameters for time of coding equal 1/2 T

Nazwa sekwencji	tPL	MV	MB	SR
Basket	5	1/4	3	8
City	4	1/4	4	8
Crew	5	1/4	3	8
Harbour	4	1/4	4	8
Ice	4	1/4	4	8
Soccer	4	1/4	4	8

Tab. 2. Wartości współczynnika PSNR dla czasów kodowania T oraz 1/2T
Tab. 2. PSNR for time of coding T and 1/2T

Nazwa sekwencji	PSNR [dB], T	PSNR [dB], 1/2 T
Basket	26,64	26,54
City	38,06	37,22
Crew	34,33	34,29
Harbour	31,09	30,97
Ice	39,20	39,11
Soccer	34,25	33,81

Podobne, choć nie takie same zależności uzyskano dla czasu kodowania 1/4T. Wyniki te zostaną przedstawione w pełnej wersji artykułu.

5. Podsumowanie

W artykule przedstawiono wyniki badań eksperymentalnych falukowego kodera sekwencji wizyjnych, podczas których dokonywano zmian parametrów kodowania, mających na celu skrócenie czasu kodowania. Wskazano takie wartości parametrów kodowania, dla których można uzyskać możliwie największą wartość współczynnika PSNR przy założonym – zredukowanym czasie kodowania. Wykazano, że możliwa jest redukcja czasu kodowania przy stosunkowo niewielkiej utracie wartości współczynnika PSNR.

6. Literatura

- [1] Ahmed N., Natarajan T., Rao K.R.: Discrete cosine transform. – IEEE Trans. on Computer, str. 90–93, 1974
- [2] Vaidyanathan P.P.: Multirate Systems and Filter Banks, Englewood Cliffs, NJ: Prentice Hall, 1993.
- [3] Vetterli M., Kovacević J.: Wavelets and Subband Coding, Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [4] Woods J. W.: Subband Image Coding. – Kluwer Academic Publisher, 1991.
- [5] Wu D., Hou Y., Zhang Y., "Scalable video coding and transport over broad-band wireless networks", Proc. of the IEEE, vol. 89, 2001.
- [6] Pesquet-Popescu B., Bottreau V.: Three-dimensional lifting schemes for motion compensated video compression, Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal, 2001.
- [7] Hsiang S.-T., Woods J. W.: Embedded video coding using invertible motion compensated 3-D subband/wavelet filter bank. – Signal Processing: Image Communication, Vol. 16, str. 705–724, 2001.
- [8] Pau G., Viéron J., Pesquet-Popescu B., "Video Coding with Flexible MCTF Structures for Low End-to-End Delay", ICIP 2005, Genova, vol. 3, 241-244.