

Andrzej POPŁAWSKI

UNIwersytet Zielonogórski, Instytut Informatyki i Elektroniki

Filtracja czasowa dla zastosowań w trójwymiarowej falkowej kompresji sekwencji wizyjnych z małym opóźnieniem kodowania

Mgr inż. Andrzej POPŁAWSKI

Tytuł magistra otrzymał na Politechnice Zielonogórskiej w roku 1998, autor lub współautor kilku prac z zakresu kompresji sekwencji wizyjnych. Pracuje na stanowisku asystenta w Instytucie Informatyki i Elektroniki Uniwersytetu Zielonogórskiego. Główne zainteresowania naukowe dotyczą zagadnień trójwymiarowej falkowej kompresji sekwencji wizyjnych dla małych opóźnień kodowania.



e-mail: A.Poplawski@iie.uz.zgora.pl

Streszczenie

Analiza falkowa z kompensacją ruchu nabrała w ostatnich latach dużego znaczenia za względu na liczne zastosowania, między innymi w koderach skalowalnych. Skalowalność jest ważną cechą wykorzystywaną podczas transmisji obrazów ruchomych poprzez sieci telekomunikacyjne o niejednorodnej strukturze. Analiza falkowa w dziedzinie czasu wprowadza opóźnienie kodowania, które wyklucza kody falkowe z wielu zastosowań. W pracy przedstawiono dotychczasowe rozwiązania zmniejszające opóźnienia kodowania, jak również przedstawiono autorską propozycję modyfikacji filtracji w dziedzinie czasu prowadzącą do redukcji opóźnienia wprowadzanego przez analizę falkową.

Słowa kluczowe: analiza falkowa, kompresja sekwencji wizyjnych z małym opóźnieniem kodowania, filtracja w dziedzinie czasu.

Three-dimensional wavelet temporal filtering for low delay video coding

Abstract

The significance of the wavelet analysis with motion compensation, due to numerous applications in the area of scalable video coders, has grown significantly in importance in the recent years. The scalability of the coders is a fundamental feature while deploying within the heterogeneous networks. Unfortunately the time domain wavelet processing is coupled with the delay, which efficiently eliminates its use in many applications. The paper presents various existing ways of decreasing the delay of video coding in wavelet temporal filtering and provides their comparison with the solution proposed by the author.

Keywords: wavelet analysis, low delay video coding, temporal filtering.

1. Wstęp

Skalowalna kompresja sekwencji wizyjnych oparta na przestrzenno-czasowej analizie falkowej jest obecnie przedmiotem szerokiego zainteresowania wielu badaczy. Kodery wykorzystujące analizę falkową stanowią interesującą alternatywę w stosunku do klasycznych koderów hybrydowych, gdyż umożliwiają w naturalny sposób uzyskanie skalowalności czasowej i przestrzennej [1, 2, 3] oraz charakteryzują się zbliżoną efektywnością kodowania.

Skalowalność jest istotna w wielu zastosowaniach np. w systemach bezprzewodowych o zmieniającej się w czasie przepustowości kanału transmisyjnego, w transmisji sekwencji wizyjnych poprzez sieci komunikacyjne o niejednorodnej strukturze [12, 16].

Warto podkreślić, że o ile analiza falkowa w dziedzinie przestrzennej posiada długą historię, o tyle analiza falkowa w dziedzinie czasu dopiero za sprawą Ohma [6], który zaproponował efektywną analizę falkową z kompensacją ruchu stała się przedmiotem szerokiego zainteresowania. Późniejsze prace [10, 11, 17] dotyczyły wykorzystania liftingu w filtracji czasowej, zwiększenia dokładności kodowania wektorów ruchu oraz zastąpienie filtrów Haara filtrami 5/3.

Filtracja w dziedzinie czasu jest zawsze związana z opóźnieniem tym większym, im większy jest rząd wykorzystanych filtrów

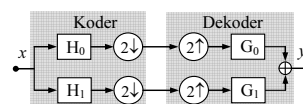
oraz im więcej stosuje się poziomów analizy czasowej. Zwiększenie rzędu filtrów polepsza co prawda separację subpasma, lecz powoduje wzrost opóźnienia kodowania. Opisane w najnowszej literaturze techniki filtracji z kompensacją ruchu poprawiając efektywność kompresji przyczyniają się do zwiększenia opóźnienia kodowania związanego z wykorzystaniem dłuższych filtrów.

W wielu zastosowaniach jak np.: wideotelefonii, systemy wideokonferencyjne, systemy zdalnej obserwacji i nadzoru wizyjnego, wymagane jest niewielkie opóźnienie kodowania. Przykładowo zalecenie IUT-T H.261 [4] określa wartość opóźnienia kodowania dla usługi wideokonferencji na poziomie nie większym niż 150 ms. To ograniczenie wymusza stosowanie bądź to innych technik kompresji sekwencji wizyjnych, bądź też modyfikacje istniejących algorytmów falkowej filtracji z kompensacją ruchu.

Celem pracy jest porównanie charakterystyk częstotliwościowych opisanych w literaturze [7, 9, 13] zespołów filtrów analizy i syntezy o zredukowanych opóźnieniach. Niestety w tych pozycjach literaturowych brakuje analizy skutków zmiany struktur filtru na charakterystyki częstotliwościowe. Niniejsza praca ma uzupełnić ten brak. Ponadto autor proponuje nowy zespół filtrów analizy subpasmovej, który również charakteryzuje się zredukowanym opóźnieniem, ale także korzystnymi właściwościami charakterystyk częstotliwościowych.

2. Obecne schematy kodowania i analiza opóźnienia

Dla uproszczenia przeanalizowano układ dzielący pasmo na dwa subpasma (rys. 1). Taki zespół filtrów stanowi podstawowy składnik układów analizy falkowej. Wejściowy sygnał x jest poddawany po stronie kodera filtracji za pomocą pary filtrów: dolno-przepustowego H_0 oraz górno-przepustowego H_1 . Ponieważ praca dotyczy charakterystyk częstotliwościowych, to pominięto zagadnienia związane z realizacją polifazową omawianych zespołów filtrów [14, 15].



Rys. 1. Ogólny schemat analizy/syntezy falkowej

Fig. 1. A general block diagram of the wavelet analysis/synthesis system

Sygnały zostają decymowane, po czym są poddawane kompresji. Sygnały w tej postaci są przesyłane do dekodera, gdzie następuje (po uprzedniej dekompresji) przywrócenie częstotliwości próbkowania oraz filtracja dolno-przepustowa za pomocą filtru G_0 i górno-przepustowa za pomocą filtru G_1 . W dziedzinie transformacji Z sygnał wyjściowy opisuje równanie:

$$Y(z) = \frac{1}{2} \left([H_0(z)G_0(z) + H_1(z)G_1(z)]X(z) + [H_0(-z)G_0(z) + H_1(-z)G_1(z)]X(-z) \right). \quad (1)$$

Idealna rekonstrukcja występuje w przypadku, gdy transmitanca układu $T(z) = H_0(z)G_0(z) + H_1(z)G_1(z) = 2z^{-q}$, oraz składowa powstała w wyniku nakładania się widm (aliasingu) w badanym sygnale $S(z) = H_0(-z)G_0(z) + H_1(-z)G_1(z) = 0$.

Filtracja falkowa z kompensacją ruchu ogrywa kluczową rolę w koderach falkowych. Obecnie wyróżnić można dwa główne schematy filtracji w dziedzinie czasu: filtracja z wykorzystaniem filtrów Haara oraz filtracja z wykorzystaniem filtrów 5/3. Jeżeli przyjmujemy następujące oznaczenia: x_t – obraz wejściowej sekwencji wizyjnej w czasie t , h_t^l – sygnał subpasma dużych częstotliwości w czasie t na poziomie l , h_t^l – sygnał subpasma niskich częstotliwości

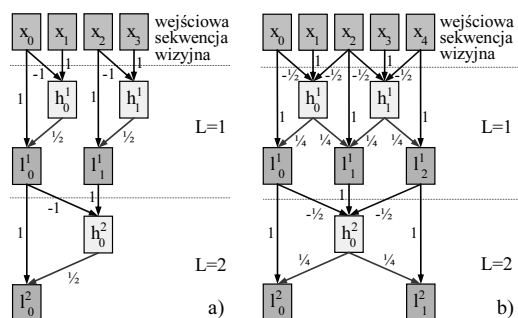
ści w czasie t na poziomie l , wówczas dla filtrów Haara uzyskujemy (rys. 2a):

$$h_l^l = \frac{1}{\sqrt{2}}(x_{2l+1} - x_{2l}), \quad l_l^l = \sqrt{2}x_{2l} + h_l^l. \quad (2)$$

Niech L oznacza numer poziomu czasowej dekompozycji falkowej. Wówczas opóźnienie kodowania N wyrażone w liczbie obrazów sekwencji dane jest zależnością: $N=2^L-1$. W przypadku zastosowania filtrów 5/3 (rys. 2b) uzyskujemy:

$$h_l^l = \frac{1}{\sqrt{2}}(x_{2l+1} - \frac{1}{2}(x_{2l} + x_{2l+2})) \quad (3)$$

$$l_l^l = \sqrt{2}x_{2l} + \frac{1}{2}(h_{l-1} + h_l) \quad (4)$$



Rys. 2. Schemat czasowej analizy falkowej z użyciem filtrów: a) Haara, b) 5/3
Fig. 2. Scheme of the temporal wavelet analysis with usage of:
a) Haar filters, b) 5/3 filters

Opóźnienie kodowania przy wykorzystaniu filtrów 5/3 dane jest zależnością: $N=2^{L+1}-1$. Opóźnienie to wynika w głównej mierze z konieczności przetworzenia przyszłych obrazów sekwencji wizyjnej, które w danym momencie nie są jeszcze znane. Im większa jest liczba poziomów dekompozycji czasowej – a tym samym liczba przetwarzanych obrazów – tym większe jest opóźnienie.

W dziedzinie transformaty Z filtry Haara (analizy) opisane są równaniami (5), zaś filtry 5/3 odpowiednio równaniami (6, 7):

$$H_0(z) = \frac{1}{\sqrt{2}}(1 + z^{-1}), \quad H_1(z) = \frac{1}{\sqrt{2}}(1 - z^{-1}), \quad (5)$$

$$H_0(z) = \frac{1}{\sqrt{2}}(-\frac{1}{4} + \frac{1}{2}z^{-1} + \frac{6}{4}z^{-2} + \frac{1}{2}z^{-3} - \frac{1}{4}z^{-4}), \quad (6)$$

$$H_1(z) = \frac{1}{\sqrt{2}}(-\frac{1}{2} + z^{-1} - \frac{1}{2}z^{-2}). \quad (7)$$

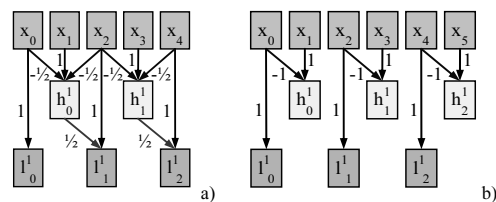
Filtry syntezy G_0 i G_1 wyznacza się na podstawie filtrów analizy [14, 15].

3. Istniejące rozwiązania redukcji opóźnienia kodowania

W literaturze można spotkać zaledwie kilka pozycji odnoszących się do problemu redukcji opóźnienia kodowania w koderach falkowych. Do najważniejszych można zaliczyć pracę [9], która jest kontynuacją propozycji zawartych w [7, 8]. Przedstawiono w niej rozwiązania umożliwiające redukcję opóźnienia kodowania w oparciu o filtry 5/3. Autorzy proponują zastosowanie trzech układów (oznaczonych jako $T1$, $T2$, $T3$) analizy/syntezy, charakteryzujących się mniejszym opóźnieniem kodowania w stosunku do filtrów 5/3, przy czym układ $T1$ (rys. 2b) wykorzystuje do analizy i syntezy filtry 5/3 bez modyfikacji.

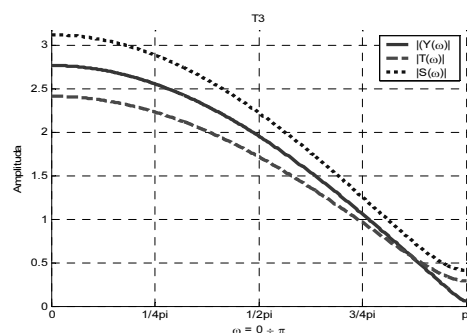
Opóźnienie kodowania w układzie $T2$ (rys. 3a) wynosi jeden obraz dla jednego poziomu dekompozycji czasowej. Układ $T3$ [9] (rys. 3b) opisany zależnościami (8) umożliwia uzyskanie opóźnienia kodowania równego zero [5], co oznacza, że każdy obraz sekwencji jest bezpośrednio przetwarzany bez konieczności oczekiwania na obrazy przyszłe:

$$H_0(z) = \sqrt{2}, \quad H_1(z) = \frac{1}{2}(1 - z^{-1}). \quad (8)$$



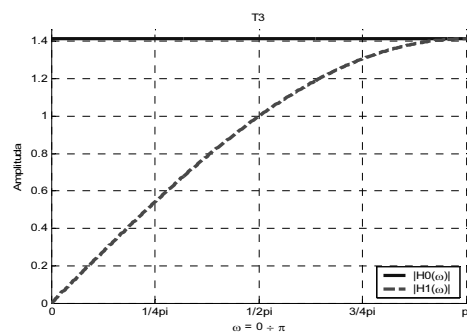
Rys. 3. Schemat czasowej analizy falkowej według układu: a) $T2$, b) $T3$
Fig. 3. Scheme of the temporal wavelet analysis with the usage of:
a) $T2$ scheme, b) $T3$ scheme

Na rys. 4 przedstawiono charakterystyki częstotliwościowe związane z transmitancją $Y(\omega)$, $T(\omega)$ oraz $S(\omega)$, zaś rys. 5 przedstawia charakterystyki częstotliwościowe filtrów górnoprzepustowego H_1 oraz dolnoprzepustowego H_0 dla układu $T3$. Jak wynika z rys. 4, zespół filtrów $T3$ nie spełnia warunków idealnej rekonstrukcji, transmitancja $T(\omega)$ jak i $S(\omega)$ jednostajnie maleje w miarę wzrostu częstotliwości ω .



Rys. 4. Charakterystyki częstotliwościowe dla układu $T3$
Fig. 4. Frequency responses of the $T3$ scheme

Jak pokazano na rys. 5, w układzie $T3$ filtracja dolnoprzepustowa praktycznie nie występuje. Należy podkreślić, że podstawą analizy falkowej jest podział pasma na dwa subpasma.



Rys. 5. Charakterystyki częstotliwościowe filtrów H_0 i H_1 dla układu $T3$
Fig. 5. Frequency responses of the H_0 and H_1 filters of the $T3$ scheme

Układy z zespołami filtrów $T2$ i $T3$ badano w pracy [9], gdzie stwierdzono mniejszą efektywność kompresji w przypadku zastosowania układu $T3$ (pozwalającego na osiągnięcie mniejszych opóźnień kodowania) w porównaniu z układem $T2$. Aby osiągnąć założone opóźnienia kodowania/dekodowania równe 150 ms, wykorzystano układ $T2$ na pierwszym i drugim poziomie dekompozycji czasowej, zaś układ $T3$ na trzecim i czwartym poziomie dekompozycji czasowej. W przypadku założonego opóźnienia o wartości 0 ms, na wszystkich poziomach dekompozycji czasowej wykorzystano układ $T3$. Dobór układu kodowania jest więc pewnego rodzaju kompromisem między efektywnością kompresji a opóźnieniem kodowania.

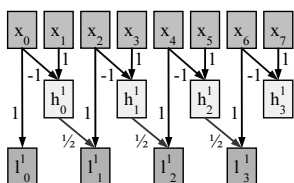
4. Propozycja zmiany filtracji w dziedzinie czasu

Ze względu na brak filtracji dolnoprzepustowej w układzie $T3$, postanowiono zmodyfikować ten układ filtracji celem dokonywa-

nia filtracji również dla niskich częstotliwości. Głównym założeniem było osiągnięcie takiego samego opóźnienia kodowania jak w układzie *T3*, stąd wynika brak możliwości filtracji z wykorzystaniem obrazów przyszłych. Na rys. 6 przedstawiono proponowany układ filtracji w dziedzinie czasu, w oparciu o filtry Haara. Zdecydowano się na ten rodzaj filtrów ze względu na mniejsze opóźnienia czasowe występujące przy wykorzystaniu tego rodzaju filtracji. Opóźnienie kodowania dla rozważanego przypadku wynosi zero obrazów. Proponowany układ oznaczono jako *B21*, opisany jest równaniami (9, 10):

$$H_0(z) = \frac{1}{\sqrt{2}}(2 + z^{-1} - z^{-2}), \quad (9)$$

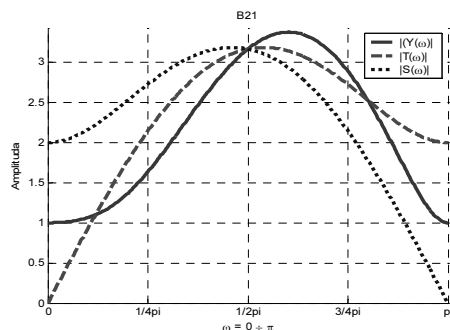
$$H_1(z) = \frac{1}{\sqrt{2}}(1 - z^{-1}). \quad (10)$$



Rys. 6. Schemat czasowej analizy falkowej dla układu *B21*

Fig. 6. Scheme of the temporal wavelet analysis with usage of the *B21* scheme

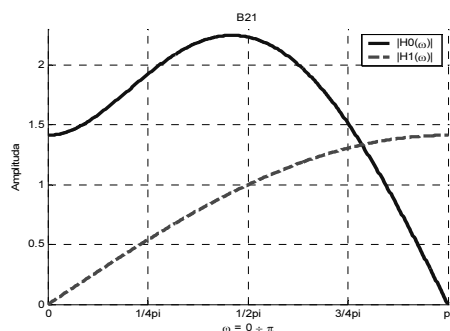
Na rys. 7 przedstawiono charakterystyki częstotliwościowe transmitancji $Y(\omega)$, $T(\omega)$ oraz $S(\omega)$, zaś rys. 8 przedstawia charakterystyki częstotliwościowe filtrów górnoprzepustowego H_1 oraz dolnoprzepustowego H_0 dla układu *B21*.



Rys. 7. Charakterystyki częstotliwościowe dla układu *B21*

Fig. 7. Frequency responses of the *B21* scheme

Jak wynika z rys. 7, również układ *B21* nie spełnia warunków idealnej rekonstrukcji, transmitancja $T(\omega)$ dla $\omega=\pi$ jest co prawda równa dwa, jednak po wzroście w zakresie średnich częstotliwości spada do zera dla $\omega=0$. Również składowa aliasingu osiąga zero jedynie dla wartości $\omega=\pi$.



Rys. 8. Charakterystyki częstotliwościowe filtrów H_0 i H_1 dla układu *B21*

Fig. 8. Frequency responses of the H_0 and H_1 filters of the *B21* scheme

W układzie *B21* w przeciwieństwie jednak do układu *T3*, występuje filtracja dolnoprzepustowa (rys. 8). Prawdą jest, że choć dla średnich częstotliwości następuje dość silny wzrost amplitudy

filtru H_0 , to należy pamiętać, że w układzie *T3* filtracja jest wykonywana tylko dla składowej wysokoczęstotliwościowej. Planuje się przeprowadzenie badań eksperymentalnych, mających na celu sprawdzenie słuszności zaproponowanego rozwiązania.

5. Podsumowanie

W pracy dokonano porównania charakterystyk częstotliwościowych opisanych w literaturze zespołów filtrów analizy i syntezy o zredukowanych opóźnieniach kodowania na tle zaproponowanego, autorskiego zespołu filtrów. Zaproponowano schemat filtracji w dziedzinie czasu, mający takie same właściwości pod względem opóźnienia kodowania w stosunku do innych rozwiązań, lecz zapewniający podział pasma na dwa subpasma, co jest podstawą analizy subpasmowej.

Zaprezentowane rozwiązanie zostało przeanalizowane teoretycznie, nie sprawdzono jego poprawności w testach eksperymentalnych. Dlatego też dalszym etapem badań będzie weryfikacja zaprezentowanego rozwiązania, polegająca na przeprowadzeniu serii eksperymentów z wykorzystaniem testowych sekwencji wizyjnych.

Praca naukowa finansowana ze środków na naukę w latach 2005-2006 jako projekt badawczy.

6. Literatura

- [1] Domański M.: Zaawansowane techniki kompresji obrazów i sekwencji wizyjnych, Poznań, WPP, 1998.
- [2] Hsiang S.-T., Woods J.W.: Embedded image coding using zeroblocks of subband/wavelet coefficients and context modeling, *ISCAS'00*, vol. 3, 662-665, 2000.
- [3] ISO/IEC International Standard 13818: Information Technology – Generic Coding of Moving Pictures and Associated Audio Information, 1994.
- [4] ITU-T Rec. H.261, Video codec for audiovisual services at p 64 kbit/s.
- [5] Ohm J.-R.: Complexity and delay analysis of MCTF interframe wavelet structures, doc. m8520, Klagenfurt MPEG meeting, July 2002.
- [6] Ohm J.-R.: Three-dimensional subband coding with motion compensation, *IEEE Transactions on Image Processing*, vol. 3, 559-589, 1994.
- [7] Pau G., Pesquet-Popescu B., van der Schaar M., Viéron J.: Delay-performance trade offs in motion-compensated scalable subband video compression, doc. m11084, Redmond MPEG meeting, July 2004.
- [8] Pau G., Pesquet-Popescu B., van der Schaar M., Viéron J.: Delay-Performance Trade-Offs in Motion-Compensated Scalable Subband Video Compression, *Proc. of Advanced Concepts for Intelligent Vision Systems (ACIVS)*, 2004.
- [9] 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.
- [10] 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.
- [11] Secker A., Taubman D.: Motion-compensated highly scalable video compression using an adaptive 3D wavelet transform based on lifting, *Proc. of the IEEE Int. Conf. on Image Processing*, 1029-1032, 2001.
- [12] Sheldon T.: Wielka encyklopedia sieci komputerowych, Robomatic, Wrocław, 1999.
- [13] Turaga D., van der Schaar M.: Unconstrained temporal scalability with multiple reference and bi-directional motion compensated temporal filtering, doc. m8388, Fairfax MPEG meeting, 2002.
- [14] Vaidyanathan P.P.: Multirate Systems and Filter Banks, Englewood Cliffs, NJ: Prentice Hall, 1993.
- [15] Vetterli M., Kovachević J.: Wavelets and Subband Coding, Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [16] Wu D., Hou Y., Zhang Y.: Scalable video coding and transport over broad-band wireless networks, *Proc. of the IEEE*, vol. 89, 6-20, 2001.
- [17] Zhan Y., Picard M., Pesquet-Popescu B., Heijmans H.: Long temporal filters in lifting schemes for scalable video coding, doc. m8680, Klagenfurt MPEG meeting, July 2002.