

# ULEPSZONY KODER SKALOWALNY DLA SYSTEMU MPEG-2

MAREK DOMAŃSKI  
ADAM ŁUCZAK  
SŁAWOMIR MAĆKOWIAK

Instytut Elektroniki i Telekomunikacji Politechniki Poznańskiej  
ul. Piotrowo 3A 60-965 Poznań  
*e-mail: [domanski, aluczak, smack]@et.put.poznan.pl*

**Streszczenie:** Artykuł prezentuje problem skalowalnego kodowania sekwencji wizyjnych. Istniejące i poddane standaryzacji rozwiązania nie są zadowalające ze względu na produkowaną zbyt dużą liczbę bitów wyjściowego strumienia danych w stosunku do koderów jednowarstwowych nieskalowalnych. Dlatego też szeroko poszukiwane są nowe rozwiązania poprawiające efektywność kodowania. Celem tego artykułu jest przedstawienie udoskonalonego systemu skalowania przestrzennego i czasowego kodera progresywnych sekwencji wizyjnych w pełni zgodnego w warstwie podstawowej ze standardem MPEG-2.

**Słowa kluczowe:** skalowalność przestrzenno-czasowa, MPEG-2, analiza subpasmowa.

## 1. WPROWADZENIE

Jesteśmy obecnie świadkami powstawania nowych usług multimedialnych wykorzystujących transmisję sygnałów wizyjnych w sieciach telekomunikacyjnych o zróżnicowanych parametrach. Konieczność dopasowania procesu kompresji cyfrowych sygnałów wizyjnych do warunków narzuconych przez takie heterogeniczne sieci telekomunikacyjne, a także wymaganie zwiększenia odporności na błędy transmisji stymulują zainteresowanie skalowalnym kodowaniem sekwencji wizyjnych. W trakcie takiej kompresji wytwarzane są dwa strumienie binarne: strumień warstwy podstawowej, który reprezentuje sekwencje obrazów o małej rozdzielczości lub niskiej jakości oraz strumień warstwy rozszerzającej reprezentujący dodatkową informację potrzebną do odtworzenia obrazów o pełnej rozdzielczości i pełnej jakości. Wydzielenie w transmitowanym sygnale pewnej informacji podstawowej umożliwi zwiększenie odporności na błędy transmisji. Warstwa podstawowa może zostać zakodowana z większą nadmiarowością podczas, gdy informacje umożliwiające odtworzenie sekwencji o pełnej jakości będą odbierane prawidłowo tylko w przypadku braku zakłóceń w kanale. Warstwy

rozszerzające mogą zostać także pominięte w czasie transmisji w warunkach dużego nasilenia ruchu w sieciach komputerowych. Istotną cechą kodowania skalowalnego jest możliwość niezależnego zdekodowania strumienia warstwy podstawowej bez potrzeby dekodowania strumienia warstwy rozszerzającej.

Skalowalność jest ujęta w standardzie MPEG-2 [1,2], który określa cztery typy skalowalności [10]:

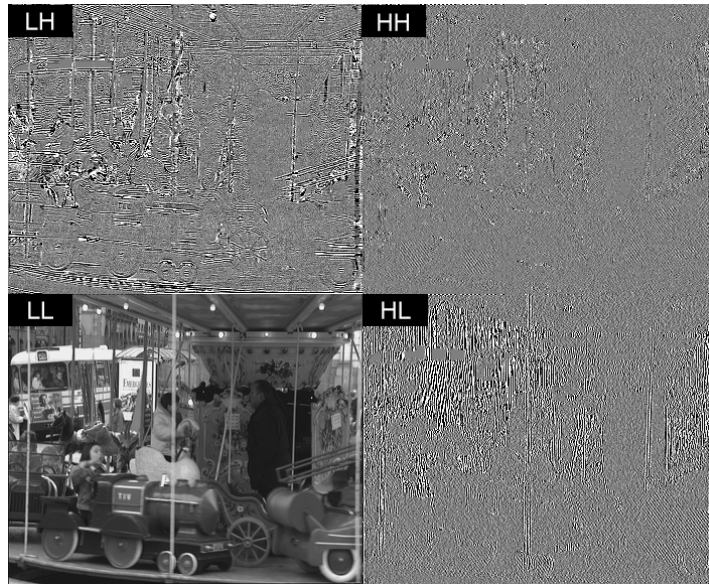
- *Skalowalność przestrzenną*, w której warstwa podstawowa reprezentuje sekwencje obrazów o mniejszej rozdzielczości przestrzennej, a warstwa rozszerzająca błąd predykcji obrazu potrzebny do zrekonstruowania obrazu o pełnej rozdzielczości przestrzennej. W rozwiązaniu tym stosuje się ideę kodowania piramidalnego.
- *Skalowalność czasową*, w której warstwa podstawowa reprezentuje sekwencje obrazów o mniejszej liczbie obrazów na sekundę. Zawarte w warstwie rozszerzającej dodatkowe obrazy kodowane są z wykorzystaniem predykcji względem obrazów warstwy podstawowej i innych obrazów wewnątrz warstwy rozszerzającej.
- *Skalowalność SNR*, w której warstwa rozszerzająca zawiera poprawki do zgrubnie skwantowanych współczynników transformaty błędu predykcji przesyłanych w warstwie podstawowej.
- *Podział danych*, w której dane zakodowane koderem nieskalowalnym podzielone są między warstwę podstawową i rozszerzającą.

Standard MPEG-2 zezwala także na łączenie różnych typów skalowań. Powstaje w takim przypadku sygnał złożony przynajmniej z kilku warstw. Skalowanie takie nazywamy hybrydowym.

Wśród opisanych powyżej czterech rodzajów skalowania, skalowalność przestrzenna jest szczególnie interesująca ze względu na perspektywie szerokie zastosowanie. Jednakże wyniki eksperymentalne [11] pokazują, że całkowity strumień danych warstwy podstawowej i rozszerzającej jest porównywalny do sumy dwóch niezależnie zakodowanych strumieni reprezentujących dwie sekwencje o różnych rozdzielczościach przestrzennych. Koder skalowalny zaproponowany przez standard MPEG-2 jest bardziej złożony od kodera nieskalowalnego ze strumieniem wytworzonym w kodowaniu równoległym (*simulcast*), które polega na niezależnym wytworzeniu dwóch strumieni danych i równoczesną ich transmisję. Podstawową jednak wadą takiego podejścia jest nieoptymalne wykorzystanie pasma transmisyjnego dostępnego kanału (tabela 1).

W literaturze [np. 3-6] opisano szereg prób poprawy efektywności kodowania skalowalnego przestrzennie. Najwięcej prac proponowało zastosowanie kodowania subpasmowego (falkowego wavelet) [3-5]. W rozwiązaniu tym każdy obraz dzielony jest na cztery subpasma (rys.1). Subpasmo LL odpowiadające niskim częstotliwościom przestrzennym jest umieszczane w warstwie podstawowej podczas, gdy trzy pozostałe subpasma (LH, HL i HH) są łącznie transmitowane w warstwie rozszerzającej.

Takie rozwiązanie umożliwiło poprawę efektywności kodowania, ale niekiedy prowadziło do wytworzenia warstwy podstawowej o zbyt dużej przepływności w stosunku do warstwy rozszerzającej.



Rys. 1. Dekompozycja subpasmowa

## 2. SKALOWALNOŚĆ PRZESTRZENNO-CZASOWA Z PODZIAŁEM DANYCH W OBRAZACH TYPU B

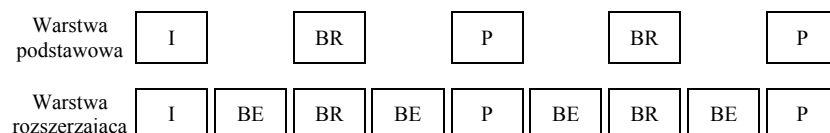
W celu uniknięcia wspomnianych wad i trudności, w Zakładzie Komunikacji Multimedialnej i Radioelektroniki Instytutu Elektroniki i Telekomunikacji Politechniki Poznańskiej zaproponowano [7,8] zastosowanie koderów ze skalowalnością mieszaną, przestrzenną i czasową.

W proponowanym rozwiązaniu warstwa podstawowa odpowiada strumieniowi obrazów o zredukowanej rozdzielczości zarówno czasowej jak i przestrzennej. Natomiast warstwa rozszerzająca jest używana do transmisji informacji potrzebnych do odtworzenia sekwencji wizyjnej o pełnej rozdzielczości czasowej i przestrzennej. Założono wysoki poziom kompatybilności ze standardem MPEG-2. W szczególności przyjęto koder warstwy podstawowej w pełni zgodny ze standardem MPEG-2.

Redukcja rozdzielczości czasowej jest osiągnięta poprzez podział strumienia obrazów typu B między warstwę podstawową a warstwę rozszerzającą. W tym celu zaproponowano dwa rodzaje obrazów typu B:

- obrazy BE kodowane w warstwie rozszerzającej,
- obrazy BR kodowane w obu warstwach (rys.2).

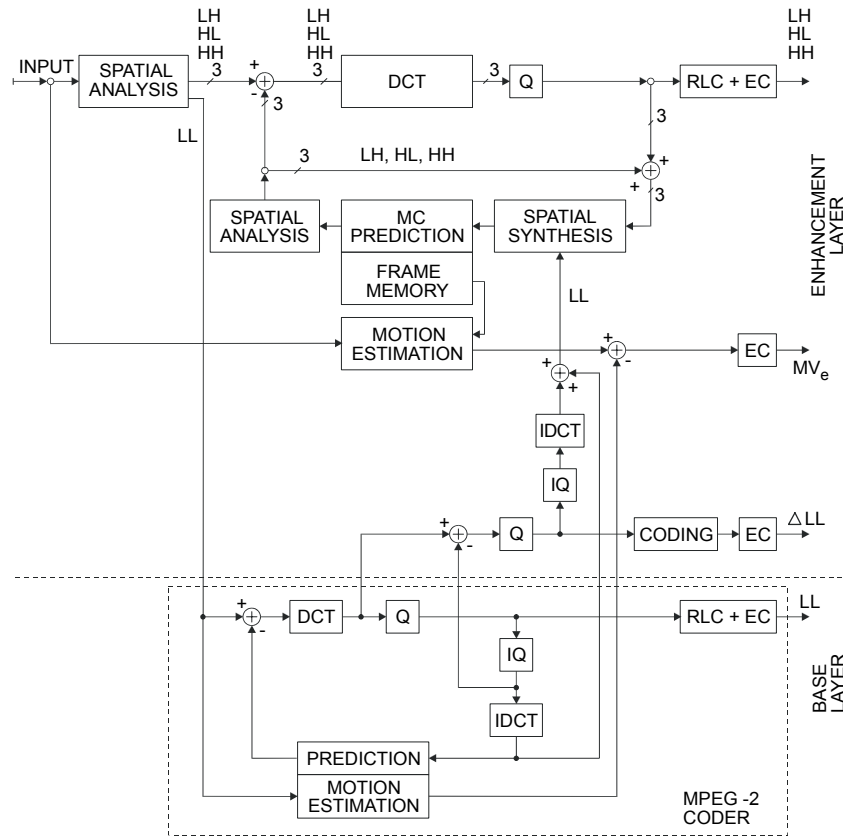
W warstwie podstawowej kodowane są: subpasmo LL obrazów typu I, P oraz BR. Natomiast warstwa rozszerzająca reprezentuje obrazy typu BE uzupełniające rozdzielczość czasową oraz subpasma LH, HL i HH obrazów typu I, P i hierarchiczne rozszerzenie obrazów typu BR służące do osiągnięcia pełnej rozdzielczości przestrzennej.



Rys. 2. Rozmieszczenie obrazów w obu warstwach

Warstwa podstawowa reprezentuje sekwencje obrazów o niższej jakości wynikającej z silniejszej kwantyzacji współczynników DCT w stosunku do warstwy

rozszerzającej. Z drugiej jednak strony niska jakość subpasma LL ogranicza jakość obrazu pełnej rozdzielczości. Dlatego wymagane jest sukcesywne transmitowanie dodatkowej informacji  $\Delta LL$  w warstwie rozszerzającej służącej do uzyskania pełnej jakości obrazów. Struktura kodera dla obrazów typu I oraz P przedstawiona jest na rysunku 3.



Rys. 3. Schemat blokowy kodera dla obrazów typu I oraz P

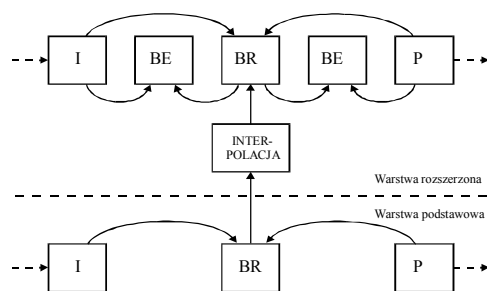
### 3. UDOSKONALONA PREDYKCJA OBRAZÓW TYPU B

Zaproponowano modyfikację predykcji zdefiniowanej w standardzie MPEG-2. Autorzy wykorzystują do tego celu obrazy typu BR, które reprezentują obraz typu B w obu warstwach. Predykcji każdego z makrobloków obrazu BR o pełnej rozdzielczości można dokonać z następujących makrobloków odniesienia (rys.4):

- poprzedniego makrobloku z obrazu typu I lub P,
- następnego makrobloku z obrazu typu I lub P
- aktualnego makrobloku z obrazu BR niskiej rozdzielczości.

Wybór odpowiedniego makrobloku następuje na podstawie kryterium najmniejszego błędu predykcji wśród kandydatów z makrobloku odniesienia, uśrednienia dwóch lub trzech.

Wyniki eksperymentalne z telewizyjnymi sekwencjami testowymi dowodzą, że w predykcji znacznej części makrobloków wykorzystywany jest aktualny obraz odniesienia BR. Dlatego predykcja obrazów BE w proponowanej przez autorów postaci jest bardziej dokładna. Dodatkowo obrazy typu BR są używane jako obrazy odniesienia dla obrazów BE w warstwie rozszerzającej. Rozwiązanie to pozwala zmniejszyć rozmiar strumienia wyjściowego oraz zwiększyć wartość współczynnika PSNR w porównaniu do predykcji określonej w standardzie.



Rys. 4. Udoskonalona predykcja obrazów typu B

#### 4. WYNIKI EKSPERYMENTALNE I WNIOSKI

Do celów eksperymentalnych koder zrealizowano w oparciu o biblioteki w języku C++ służące do realizacji oprogramowania wspomagającego analizę algorytmów kodowania. Cechą tego oprogramowania [9] jest jego elastyczność podczas testów z różnymi wariantami schematu kodowania. Biblioteka zawiera programową implementację koder standardu MPEG-2 MP@ML wykorzystaną w warstwie podstawowej.

W eksperymentach wykorzystywano progresywne sekwencje testowe 720×576, 50Hz: „Flower Garden”, „Mobile & Calendar” oraz „Fun Fair”.

Tabela 1 przedstawia wyniki uzyskane podczas kodowania równoległego „simulcast” z wykorzystaniem dwóch niezależnych koderów MPEG-2 (jednego dla warstwy podstawowej, drugiego dla warstwy rozszerzającej). Uzyskany łączny strumień skalowalny jest większy o  $72 \div 88\%$  od strumienia wytworzonego przez koder nieskalowalny MPEG-2. Natomiast proponowany przez autorów algorytm wykazuje znacznie wyższą efektywność kodowania. Eksperymenty pokazują, że dla przepływności poniżej 6Mb/s zwiększenie strumienia skalowalnego wynosi około  $2 \div 5\%$  w stosunku do koder nieskalowalnego. Dodatkowo proponowany koder spełnia założenie, że rozmiar strumienia warstwy podstawowej nie może przekraczać rozmiaru strumienia warstwy rozszerzającej lub może być niewiele większy. Są to obecnie jedne z najlepszych wyników osiąganych na świecie. W tabeli 2 przedstawione są wyniki eksperymentalne porównania proponowanego koder z koderem nieskalowalnym.

Tabela 1. Wyniki eksperymentalne kodowania równoległego (simulcast)

		Sekwencja testowa	
		<i>Mobile &amp; Calendar</i>	<i>Flower Garden</i>
Koder nieskalowalny MPEG-2 (CIF)	Strumień [Mb/s]	2.93	3.40
	Uśredniona wartość PSNR [dB] dla składowej luminacji	27,63	29,60
Koder nieskalowalny MPEG-2 (ITU-R BT.601)	Strumień [Mb/s]	4.04	3.88
	Uśredniona wartość PSNR [dB] dla składowej luminacji	27,73	29,63
Kodowanie równoległe (simulcast)	Strumień [Mb/s]	6.97	7.28
	Uśredniona wartość PSNR [dB] dla składowej luminacji	27,73	29,63
Zwiększenie strumienia w stosunku do koder nieskalowalnego	Strumień [Mb/s]	2.93	3.40
	Strumień [%]	72,5	87,6

Tabela 2. Porównanie proponowanego kodera z koderem nieskalowalnym standardu MPEG-2

		Sekwencja testowa			
		Flower Garden		Funfair	
Warstwa podstawowa (MPEG-2)	Strumień [Mb/s]	5.24	8.76	5.03	8.48
	Wartość średnia PSNR [dB] dla 60 obrazów składowej luminacji	30.63	33.75	31.38	34.07
Proponowany koder skalowalny	Strumień [Mb/s]	5.34	8.74	5.04	8.48
	Wartość średnia PSNR [dB] dla 60 obrazów składowej luminacji	30.44	33.41	30.81	34.16
	Strumień warstwy podstawowej [Mb/s]	2.21	3.00	2.11	2.70
	Strumień warstwy podstawowej w stosunku do całego strumienia skalowalnego [%]	41	34	41	31

## 5. LITERATURA

- [1] ISO/IEC International Standard 13818, Information Technology - Generic Coding of Moving Pictures and Associated Audio Information.
- [2] B. Haskell, A. Puri, A. N. Netravali, *Digital Video: An Introduction to MPEG-2*, New York: Chapman & Hall, 1997.
- [3] T. Tsunashima, J. Stampleman, V. Bove, *A scalable motion - compensated subband image coder*, IEEE Trans. on Communication, vol. 42, pp. 1894-1901, 1994.
- [4] F. Bosveld, *Hierarchical video compression using SBC*, Ph.D. dissertation, Delft University of Technology, Delft 1996.
- [5] Senbel, H. Abdel-Wahab, *Scalable and robust image compression using quadrees*, Signal Processing: Image Communication, vol. 14, pp. 425-441, 1999.
- [6] U. Benzler, *Scalable multi-resolution video coding using a combined subband-DCT approach*, Picture Coding Symposium 1999, pp. 17-20.
- [7] M. Domański, A. Łuczak, S. Maćkowiak, R. Świerczyński, *Hybrid coding of video with spatio-temporal scalability using subband decomposition*, in Signal Processing IX: Theories and Applications, pp. 53-56, Rhodes, September 1998.
- [8] M. Domański, A. Łuczak, S. Maćkowiak, R. Świerczyński, Ulrich Benzler, *Spatio-Temporal Scalable Video Codecs with MPEG-Compatible Base Layer*, Picture Coding Symposium 1999, pp. 45-48.
- [9] A. Łuczak, S. Maćkowiak, *Uniwersalne oprogramowanie do realizacji koderów wizyjnych*, Poznańskie Warsztaty Telekomunikacyjne, Poznań 1998, str. 3.5-1 3.5-4, 1998
- [10] M. Domański, A. Łuczak, R. Świerczyński, *Skalowalne kodowanie sygnałów wizyjnych*, Poznańskie Warsztaty Telekomunikacyjne, Poznań 1997, 1997.
- [11] International Telecommunication Union, Radiocommunication Study Groups, Document 11-3/103-E, MPEG-2 spatial scalable video coding.