

## MODEL WERYFIKACYJNY TRÓJWARSTWOWEGO SKALOWALNEGO KODERA WIZYJNEGO WYKORZYSTUJĄCEGO STRUKTURY KODERÓW MPEG-2

**Streszczenie:** Artykuł dotyczy zrealizowanego modelu weryfikacyjnego trójwarstwowego skalowalnego kodera wizyjnego przeznaczonego do zastosowań multimedialnych związanych z przesyłaniem obrazów telewizyjnych. Model weryfikacyjny jest złożonym pakietem oprogramowania realizującym autorską koncepcję kodera skalowalnego trójwarstwowego. Zaproponowany koder skalowalny jest bardziej efektywny od kodera użytego w standardzie MPEG-2.

### 1. WPROWADZENIE

W systemach multimedialnych cyfrowe sygnały wizyjne przesyła się często przez heterogeniczne sieci telekomunikacyjne charakteryzujące się różnymi przepływnościami binarnymi. Wysyłany przez serwer usługi wizyjnej strumień może być podzielony na kilka warstw. Terminale odbierające sygnał z mniejszą przepływnością otrzymują tylko strumienie niższych warstw pozwalając na odbiór z gorszą jakością, podczas gdy terminale dołączone do fragmentów sieci o pełnej przepływności odbierają wszystkie warstwy, co umożliwia im odbiór sekwencji pełnej jakości. Ta cecha funkcjonalna nosi nazwę skalowalności strumienia danych. Dotychczas w standardach kodowania MPEG-2 [1,2,6] i MPEG-4 zaproponowano różne typy skalowalności:

- Skalowalność przestrzenna. Warstwa podstawowa reprezentuje sekwencje obrazów o mniejszej rozdzielczości przestrzennej, a warstwa rozszerzająca zawiera dodatkowe informacje potrzebne do zrekonstruowania obrazu o pełnej rozdzielczości przestrzennej.
- Skalowalność czasowa. Warstwa podstawowa reprezentuje sekwencje obrazów o zmniejszonej rozdzielczości czasowej, a warstwa rozszerzająca zawiera dodatkowe informacje potrzebne do zrekonstruowania sygnału o pełnej rozdzielczości czasowej.
- Skalowalność SNR. Warstwa rozszerzająca zawiera poprawki do zgrubnie skwantowanych współczynników transformaty błędu predykcji przesyłanych w warstwie podstawowej.

- Podział danych. Zakodowane dane podzielone są między warstwę podstawową i rozszerzającą.

Niestety implementacja skalowalności przestrzennej opisanej w standardzie MPEG-2 [6] jest bardzo nieefektywna, gdyż odbywa się kosztem znacznego zwiększenia całkowitego strumienia binarnego. Dlatego w ostatnich latach poszukuje się bardziej efektywnych rozwiązań. W literaturze opisano szereg koderów wykorzystujących bezpośrednio analizę subpasmową [3,4], jednak bardziej zachęcające wyniki uzyskano przez łączenie różnych typów skalowalności, skalowalności przestrzennej i SNR [5] lub skalowalności przestrzennej i czasowej (skalowalność przestrzenno-czasowa) [7,8].

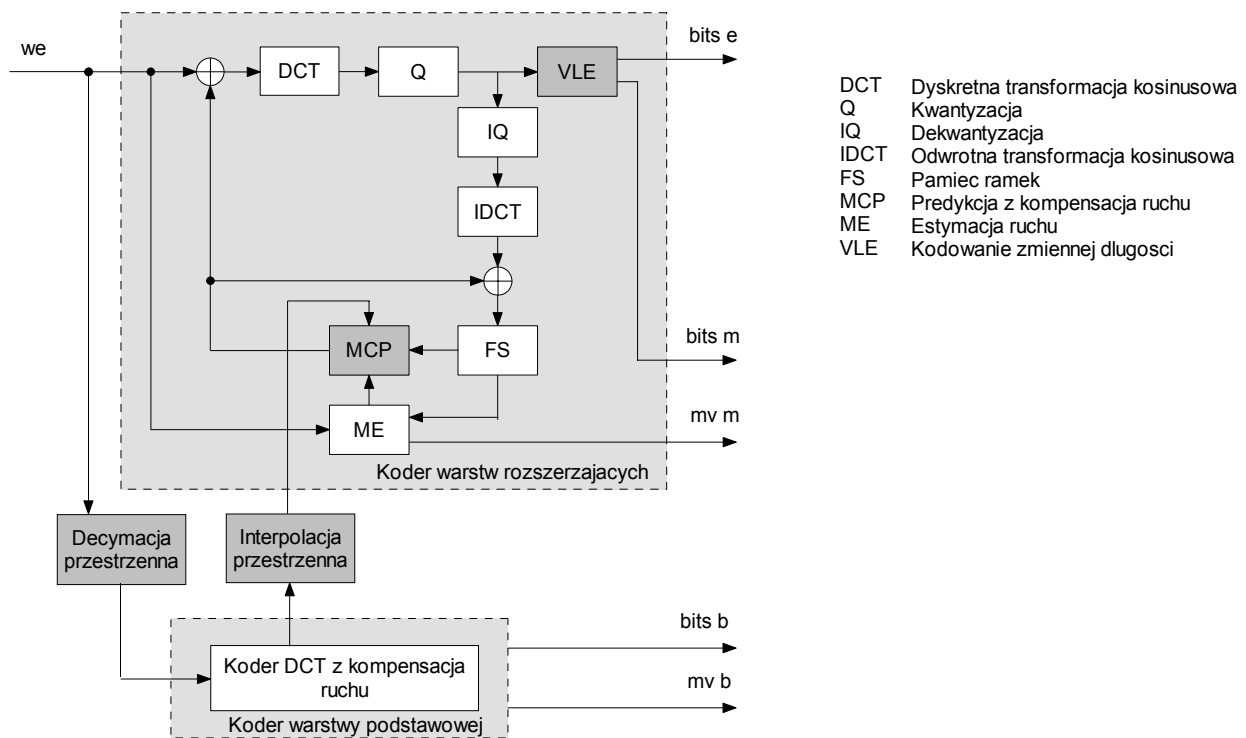
W pracy poddano weryfikacji zaproponowany przez autorów [9,10] trójwarstwowy koder wizyjny oparty na skalowalności przestrzenno-czasowej i podziale danych przeznaczony do zastosowań multimedialnych związanych z przesyłaniem obrazów telewizyjnych.

### 2. STRUKTURA TRÓJWARSTWOWEGO KODERA SKALOWALNEGO

Koder (rys.1.) generuje trzy strumienie danych o podobnych przepływnościach:

- Warstwa podstawowa reprezentuje sekwencję wizyjną o zredukowanej rozdzielczości czasowej i przestrzennej.
- Warstwa pośrednia umożliwia uzyskanie obrazów o pełnej rozdzielczości ale o zredukowanej jakości (o zredukowanej wartości współczynnika PSNR). W warstwie pośredniej przesyłane są współczynniki DCT odpowiadające małym częstotliwościom przestrzennym.
- Warstwa rozszerzająca zawiera pozostałe współczynniki DCT potrzebne do odtworzenia obrazu o pełnej jakości i pełnej rozdzielczości przestrzennej i czasowej.

\*Autor jest stypendystą Fundacji na Rzecz Nauki Polskiej



Rys.1 Schemat blokowy trójwarstwowego koder skalowalnego (dla kodowania międzyobrazowego)

Koder skalowalny złożony jest z:

- Niezależnie działającego koder warstwy podstawowej, w pełni zgodnego ze standardem MPEG-2.
- Koder warstwy pośredniej i rozszerzającej, który wykorzystując informacje z warstwy podstawowej dokonuje kodowania obrazów o pełnej rozdzielczości przestrzennej i czasowej.

Bloki funkcjonalne wprowadzone lub zmodyfikowane w stosunku do koder nieskalowalnego MPEG-2 to zaznaczone na rys.1 bloki predykcji z kompensacją ruchu dla warstw wyższych, decymacji i interpolacji przestrzennej oraz zmodyfikowany koder entropowy warstw pośredniej i rozszerzającej. Koder warstw wyższych dokonuje rozdziału współczynników DCT pomiędzy warstwę pośrednią i rozszerzającą. Strumień warstwy pośredniej zawiera wszystkie nagłówki, wektory ruchu oraz współczynniki DCT odpowiadające małym częstotliwościom. Strumień warstwy rozszerzającej zawiera oprócz pozostałych współczynników DCT, wszystkie nagłówki warstw makrobloków potrzebne do resynchronizacji w przypadku wystąpienia błędów transmisji. Szczegółowy opis koder trójwarstwowego znajduje się w pracach [9,10].

### 3. MODEL WERYFIKACYJNY

Głównym celem pracy było stworzenie bardzo efektywnego narzędzia badawczego umożliwiającego testowanie autorskiej koncepcji trójwarstwowego kodowania skalowalnego.

Założono, że implementacja zostanie zrealizowana w postaci obiektowej. Zrezygnowano z proceduralnego

sposobu pisania programu chcąc uzyskać większą elastyczność oprogramowania, przejrzystość źródła oraz możliwość szybkiej modyfikacji różnych parametrów badanego koder wizyjnego.

W konstrukcji modelu weryfikacyjnego oparto się na uniwersalnym oprogramowaniu [11], służącym do realizacji koderów wizyjnych np. H.263, MPEG-2, stworzonym w Laboratorium Zakładu Komunikacji Multimedialnej i Radiokomunikacji Politechniki Poznańskiej. Wykorzystanie tej biblioteki było możliwe dlatego, że warstwa podstawowa jest w pełni zgodna ze standardem MPEG-2 a cała struktura wykazuje wysoki poziom kompatybilności z poszczególnymi blokami koder MPEG-2.

Ze względu na synchroniczną pracę obu koderów korzystających ze wspólnych danych wejściowych, szczególną uwagę zwrócono na synchronizację pracy koderów, co zostało uzyskane poprzez odpowiednie wywołanie metod danego koder. Fragment kodu źródłowego koder skalowalnego przedstawiono na rysunku 2.

Zalety obiektowej realizacji oprogramowania wynikają z faktu, że schemat blokowy koder skalowalnego w dużej mierze implikuje nam deklaracje obiektów wykorzystywanych w programie. Graf przepływu danych odpowiadający schematowi koder w prosty sposób przekłada się na główną część programu. W tej części brak operacji nie związanych z algorytmem kodowania, co wpływa na przejrzystość kodu źródłowego. Umożliwia to szybkie testowanie i modyfikacje algorytmu.

Pobranie obrazu z sekwencji wejściowej

```
B0.Get (pn, input) ;
```

Decymacja przestrzenna obrazu wejściowego

```
qmfy.Split (B0.Y, LL.Y, LH.Y, HL.Y, HH.Y) ;
qmfc.Split (B0.U, LL.U, LH.U, HL.U, HH.U) ;
qmfc.Split (B0.V, LL.V, LH.V, HL.V, HH.V) ;
```

Kodowanie obrazu w warstwie podstawowej

```
BaseLayer.EncodePFrame (LL, 31) ;
RLL = BaseLayer.Reconstructed;
```

Interpolacja przestrzenna obrazu wyjściowego z kodera warstwy podstawowej

```
LH = 0;
HL = 0;
HH = 0;
```

```
qmfy.Merge (RLL.Y, LH.Y, HL.Y, HH.Y, B1.Y) ;
qmfc.Merge (RLL.U, LH.U, HL.U, HH.U, B1.U) ;
qmfc.Merge (RLL.V, LH.V, HL.V, HH.V, B1.V) ;
```

Zmodyfikowana predykcja i kompensacja ruchu w warstwie rozszerzającej

```
mep.MPEG2PCompensation (B2, B0, T0, 31) ;
mep.MPEG2PCompensation (T2, B0, T1, 31) ;
interpolacja (B0, T0, B1, T1, T2) ;
```

Kodowanie obrazu w warstwie rozszerzającej

```
EnhP.MakeBFrame (pn, B0, T2, T0, T1, PB) ;
```

Obliczenie współczynnika PSNR dla zdekodowanych obrazów wyjściowych warstwy podstawowej i rozszerzającej

```
psnrBY = LL.Y.PSNR (RLL.Y) ;
psnrBU = LL.U.PSNR (RLL.U) ;
psnrBV = LL.V.PSNR (RLL.V) ;
```

```
psnrY = B0.Y.PSNR (PB.Y) ;
psnrU = B0.U.PSNR (PB.U) ;
psnrV = B0.V.PSNR (PB.V) ;
```

Rys.2. Fragment kodu źródłowego trójwarstwowego kodera skalowalnego realizującego kodowanie obrazów w trybie predykcyjnym

Podczas przygotowywania modelu weryfikacyjnego zrezygnowano z optymalizacji prędkości działania algorytmów, ponieważ jest to oprogramowanie eksperymentalne.

Poprawność działania kodera została sprawdzona poprzez wykonanie testów zgodności wybranych fragmentów kodera skalowalnego z modelem referencyjnym kodera MPEG-2. Badano na przykład

wartości przepływności binarnej oraz wartości współczynnika PSNR. Działanie kodera warstwy podstawowej sprawdzono poprzez stwierdzenie poprawności dekodowania wyjściowego strumienia bitów za pomocą oprogramowania referencyjnego MPEG-2.

W celu kontroli parametrów opisujących przebieg procesu kodowania w oprogramowaniu zaimplementowano mechanizm przygotowywania raportów. Informacje odnoszące się do przeprowadzonego kodowania sekwencji wizyjnej są gromadzone w pliku tekstowym. Plik zawiera informacje dotyczące poszczególnych obrazów, jak i sumaryczne podsumowanie wyników kompresji całej sekwencji wizyjnej.

Tabela 1. Wyniki eksperymentalne uśrednione dla całych sekwencji, uzyskane dla trójwarstwowego kodera skalowalnego

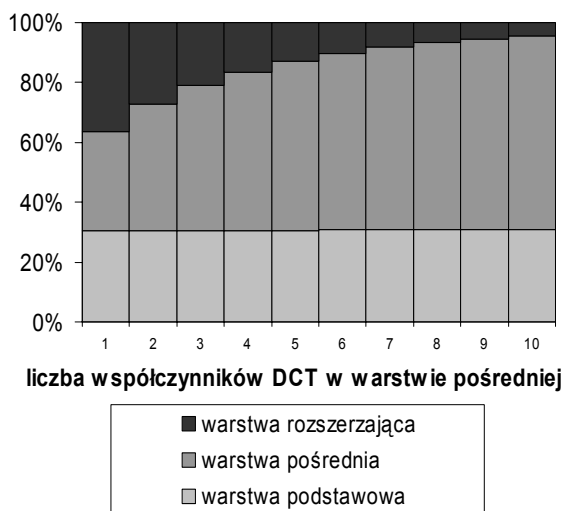
Nazwa sekwencji		Funfair	Stefan	Cheer	
Koder nieskalowalny MPEG-2	Strumień [Mb]	7,0	7,0	7,0	
	PSNR dla składowej luminancji [dB]	33,8	37,0	33,7	
Proponowany koder trójwarstwowy	PSNR dla składowej luminancji [dB]	warstwy podstawowej	34,1	37,5	33,1
		odzyskanej z warstwy podstawowej i pośredniej	29,7	32,9	30,3
		odzyskanej z wszystkich trzech warstw	33,8	37,0	33,7
	Strumień warstwy [Mb]	podstawowej	2,5	2,5	2,5
		pośredniej	2,8	3,57	2,67
		rozszerzającej	2,77	2,84	2,64
	Zwiększenie strumienia [%] (w porównaniu do kodera nieskalowalnego)		15,2	27,3	11,6

#### 4. WYNIKI EKSPERYMENTALNE

Do realizacji modelu weryfikacyjnego wykorzystano kompilator języka Microsoft Visual C++. Oprogramowanie weryfikacyjne przeznaczone jest dla komputerów typu PC z systemem operacyjnym Windows 2000. Do eksperymentów wykorzystywano komputer typu PC wyposażony w dwa procesory Pentium III oraz 1GB pamięci RAM.

Badania zostały przeprowadzone dla trzech, testowych, progresywnych sekwencji telewizyjnych 720x576, 4:2:0, 50Hz: "Cheer", "Stefan" oraz "Fun Fair".

Ekspertyzy miały na celu sprawdzenie czy zaproponowany podział współczynników transformaty DCT między warstwę pośrednią i rozszerzającą jest efektywny. Tabela 1 przedstawia wyniki eksperymentalne uzyskane podczas trójwarstwowego kodowania skalowalnego w odniesieniu do koderów nieskalowalnego MPEG-2 dla przepływności 7Mb/s. Wyniki te ilustrują przypadek umieszczenia w warstwie pośredniej jednego współczynnika DCT na blok. W przypadku takiego podziału danych DCT, wszystkie trzy strumienie danych mają podobne przepływności. Rysunek 3 ilustruje procentowy udział warstw w całym strumieniu danych w przypadku zwiększania liczby współczynników DCT w warstwie pośredniej. Zaproponowane rozwiązanie umożliwia płynną alokację bitów dla warstwy pośredniej i rozszerzającej w zależności od liczby współczynników DCT. Zastosowanie takiego podziału związane jest z wysyłaniem dodatkowych nagłówków warstw makrobloków w warstwie rozszerzającej potrzebnych do resynchronizacji w przypadku wystąpienia błędów transmisji. Ekspertyzy pokazują, że dla przepływności 7Mb/s zwiększenie strumienia skalowalnego związanego z wprowadzeniem skalowalności dwu- i trójwarstwowej wynosi około 10% - 28% w stosunku do koderów nieskalowalnego (Tabela 1). Pomimo przesyłania wymaganej dodatkowej informacji, proponowany przez autorów algorytm wykazuje znacznie wyższą efektywność kodowania niż kodery skalowalne MPEG-2.



Rys.3. Uśredniony dla trzech sekwencji testowych procentowy udział warstw w całym strumieniu danych dla przepływności 7Mb/s

## 5. PODSUMOWANIE

Model weryfikacyjny jest złożonym pakietem oprogramowania, na który składa się około 13 000 linii kodu obiektów źródłowych. Podstawowym zadaniem przygotowanego oprogramowania jest wspomaganie analizy algorytmów kodowania skalowalnego. Zaletą omawianego modelu jest jego elastyczność

umożliwiająca zmianę algorytmu pracy przy testowaniu różnych wersji koderów.

W wyniku przeprowadzonych eksperymentów, pokazano, że możliwy jest efektywny podział informacji między trzy strumienie danych o podobnych przepływnościach. Dla przypadku trójwarstwowego proponowany algorytm wykazuje znacznie wyższą efektywność kodowania niż kodery skalowalne MPEG-2.

Dotychczas badano implementacje koderów skalowalnych dwu- i trójwarstwowych. Obecnie trwają prace nad wprowadzeniem do algorytmu kodowania: skalowalności o wielu poziomach jakości usług i obsługę kodowania sekwencji z wybieraniem międzyliniowym.

## 6. SPIS LITERATURY

- [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] U. Benzler, "Scalable multi-resolution video coding using a combined subband-DCT approach", *Picture Coding Symposium 1999*, pp. 17-20
- [6] International Telecommunication Union, Radiocommunication Study Groups, Document 11-3/103-E, MPEG-2 spatial scalable video coding
- [7] M. Domański, A. Łuczak, S. Maćkowiak, "Spatio-Temporal Scalability for MPEG" *IEEE Transactions on Circuits and Systems for Video Technology* Vol. 10 No. 7 October 2000
- [8] M. Domański, A. Łuczak, S. Maćkowiak, "Spatio-Temporal Scalability using modified MPEG-2 predictive video coding", *Proceedings of European Signal Processing X*, Tampere, Finlandia 2000, str. 961-964
- [9] M. Domański, A. Łuczak, S. Maćkowiak, "Modified MPEG-2 video coders with efficient multi-layer scalability", *Proc. IEEE Int. Conf. Image Proc.*, Tessaloniki, 2001, (w druku)
- [10] M. Domański, S. Maćkowiak, "MPEG 2 - Based Video Coding with Three-Layer Mixed Scalability", *CAIP 2001*, Warszawa 2001, (w druku)
- [11] A. Łuczak, S. Maćkowiak, "Uniwersalne oprogramowanie do realizacji koderów wizyjnych", *Poznańskie Warsztaty Telekomunikacyjne*, Poznań 1998