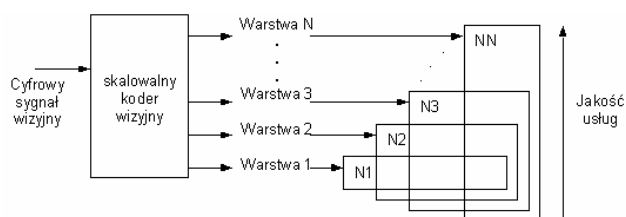


HIERARCHICZNA KOMPRESJA CYFROWYCH SYGNAŁÓW WIZYJNYCH

Streszczenie: Artykuł prezentuje problem możliwości konstrukcji efektywnych algorytmów hierarchicznej kompresji cyfrowych sygnałów wizyjnych dla zastosowań w heterogenicznych sieciach telekomunikacyjnych o wielu poziomach jakości przy założeniu osiągnięcia znacznej kompatybilności ze standardami MPEG. Istniejące i poddane standaryzacji rozwiązania skalowalności nie są zadawalające ze względu na niską efektywność kompresji. 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.



Rys. 1. Warstwowy skalowalny system kodowania sygnału wizyjnego w sieciach heterogenicznych, zawierających podsieci N1,...,NN

1. WPROWADZENIE

Rozwój metod kompresji cyfrowych sygnałów wizyjnych umożliwił powstanie różnorodnych usług multimedialnych. Dlatego istnieje szerokie zainteresowanie transmisją sygnału wizyjnego poprzez heterogeniczne sieci telekomunikacyjne, które charakteryzują się różnym dostępnym poziomem jakości usług. Zróżnicowany poziom jakości usług (QoS) jest często związany z różną dostępną przepływnością binarną kanału transmisyjnego. Z tej przyczyny, transmitowany strumień binarny powinien być podzielony na kilka warstw w taki sposób, aby warstwa podstawowa mogła być zdekodowana niezależnie od pozostałych dostarczając sygnału wizyjnego o zredukowanej jakości lub rozdzielczości. Warstwy 1,...,N odpowiadają podsystemom N1 do NN heterogenicznych sieci komunikacyjnych.

Dotychczas w standardzie telewizji cyfrowej MPEG-2 zaproponowano różne typy skalowalności. Niestety zaproponowane rozwiązania są bardzo nieefektywne. Dlatego na całym świecie odnotowuje się duże zainteresowanie poszukiwaniami efektywnych algorytmów hierarchicznej kompresji sygnałów wizyjnych.

W tym artykule została zaproponowana struktura hybrydowego skalowalnego kodera wizyjnego. Celem jest osiągnięcie całkowitej sumarycznej przepływności binarnej strumieni wizyjnych wszystkich warstw na poziomie przepływności pojedynczego strumienia wizyjnego kodera nieskalowalnego. W proponowanym rozwiązaniu strumień warstwy podstawowej odpowiadający niskiej rozdzielczości jest w pełni zgodny ze standardem MPEG-2.

2. KODOWANIE SKALOWALNE

Standard MPEG-2 [1,2] 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.

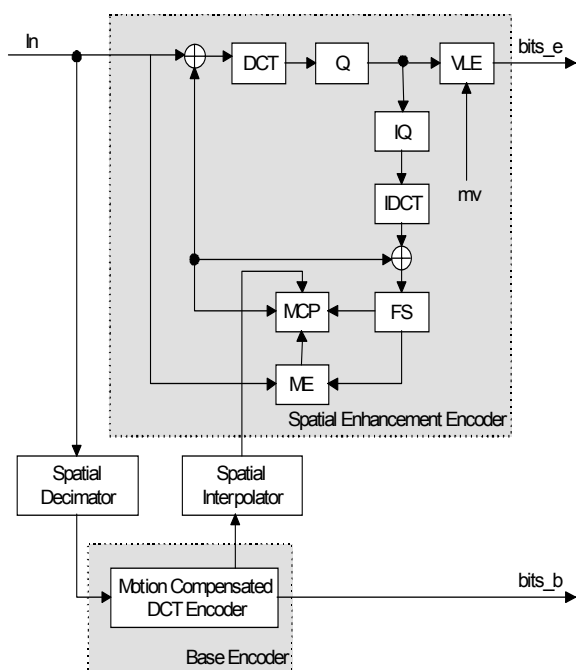
- Podział danych, w której dane zakodowane koderem nieskalowalnym podzielone są między warstwę podstawową i rozszerzającą.

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 perspektywnie 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.

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. Subpasmo odpowiadające niskim częstotliwościom przestrzennym jest umieszczane w warstwie podstawowej podczas, gdy trzy pozostałe subpasma są łącznie transmitowane w warstwie rozszerzającej.

Takie rozwiązanie umożliwiło poprawę



Rys. 2. Schemat blokowy kodera skalowalnego przestrzenno-czasowego dla obrazów typu I oraz P

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

3. SKALOWALNOŚĆ CZASOWO-PRZESTRZENNA

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. Struktura zaproponowanego kodera przedstawiona jest na rysunku 2.

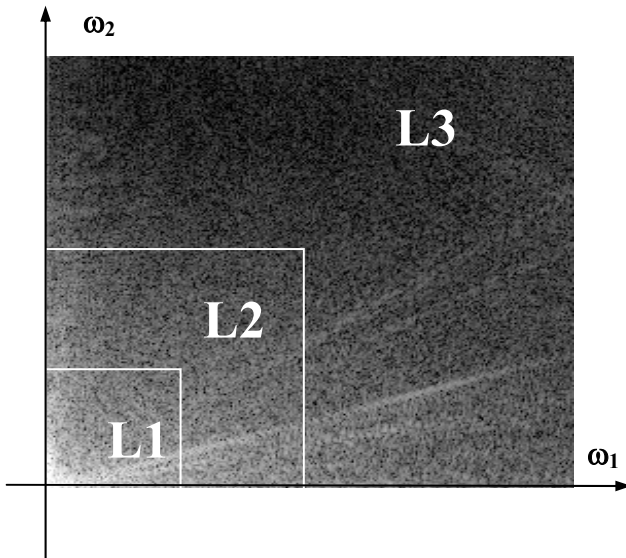
W trakcie takiej kompresji wytwarzane są dwa strumienie binarne: strumień warstwy podstawowej reprezentujący sekwencję wizyjną o odpowiednio zmniejszonej rozdzielczości przestrzennej i czasowej oraz dodatkową informację przesyłaną w warstwie rozszerzającej służącą do zrekonstruowania w dekoderze sekwencji wizyjnej o pełnej rozdzielczości zarówno czasowej jak i przestrzennej.

Wspólną cechą zaproponowanych rozwiązań jest wykorzystanie w warstwie podstawowej kodera hybrydowego zgodnego ze standardem MPEG-2. Koder warstwy rozszerzającej w wyniku analizy oraz wielu eksperymentów badawczych ewoluował do postaci, która jest znacznie efektywniejsza niż rozwiązania zaproponowane w międzynarodowych standardach kompresji cyfrowych sygnałów wizyjnych. Podejście takie ma wiele korzyści ponieważ można wykorzystać obecnie istniejące układy koderów i dekoderów MPEG-2, tylko nieznacznie modyfikując algorytm kompresji, korzystając z gotowych i sprawdzonych rozwiązań. Proponowane modyfikacje bazują na dobrze przebadanych i znanych metodach a wynikają z analizy pracy kodera skalowalnego.

W proponowanym rozwiązaniu zastosowano kilka oryginalnych rozwiązań. Jednym z nich jest zastosowanie skalowalności przestrzenno-czasowej z podziałem danych w obrazach typu B. W celu poprawy efektywności kodowania skalowalnego przestrzennie w koderze wykorzystano kodowanie subpasmowe [8,9].

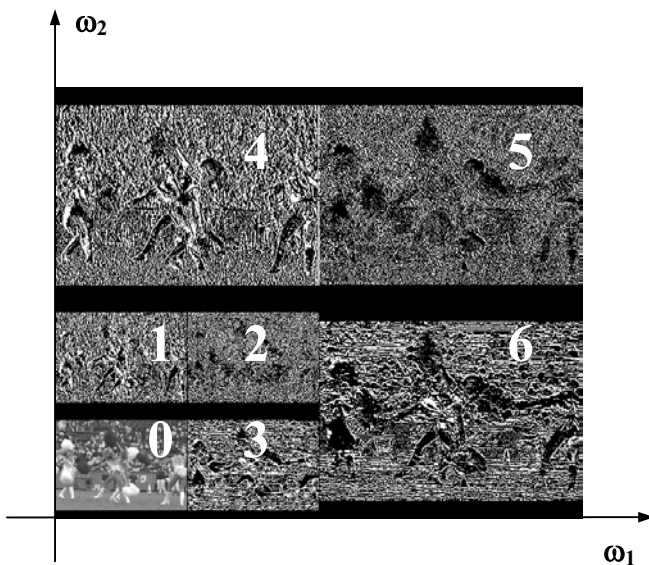
4. KODOWANIE SKALOWALNE WEWNĄTRZOBRAZOWE

Kodowanie wewnątrzobrazowe skalowalne przestrzennie (kodowanie hierarchiczne) może być zaimplementowane z wykorzystaniem dekompozycji subpasmowej. Rysunek 3 przedstawia zależność pomiędzy warstwami odpowiadającymi różnym rozdzielczościom a pasmem częstotliwości przestrzennych obrazu zakodowanego na danej warstwie.



Rys. 3. Warstwy w dziedzinie częstotliwości przestrzennych

W proponowanym rozwiązaniu dwuwarstwowym każdy obraz Intra dzielony jest na cztery subpasma. Subpasmo (0) odpowiadające niskim częstotliwościom przestrzennym jest umieszczane w warstwie podstawowej podczas, gdy trzy pozostałe subpasma (1, 2 i 3) są łącznie transmitowane w warstwie rozszerzającej. W przypadku kodaera wielowarstwowego przyporządkowanie subpasm do warstw przedstawione jest w tabeli 1.



Rys. 4. Subpasma uzyskane w dekompozycji subpasmowej

Tabela 1. Przyporządkowanie subpasm do odpowiadających im warstw w kodowaniu wewnątrzobrazowym

Warstwa	Subpasmo
L1	0
L2	1,2,3
L3	4,5,6

5. KODOWANIE MIĘDZYOBRAZOWE ZE SKALOWALNOŚCIĄ MIESZANĄ PRZESTRZENNO-CZASOWĄ

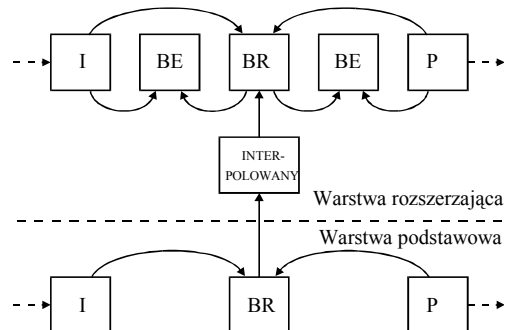
W celu uniknięcia wady związanej z wytworzeniem warstwy podstawowej o zbyt dużej przepływności w stosunku do warstwy rozszerzającej zastosowano redukcję rozdzielczości czasowej w warstwie najniższej.

Redukcja rozdzielczości czasowej jest osiągnięta poprzez podział strumienia obrazów typu B między warstwę podstawową a warstwę rozszerzającą[5]. W tym celu zaproponowano dwa rodzaje obrazów typu B: obrazy BE kodowane w warstwie rozszerzeń, oraz obrazy BR kodowane w obu warstwach (rys.5).

Kolejną zaproponowaną zmianą jest modyfikacja predykcji zdefiniowanej w standardzie MPEG-2. Wykorzystywane są do tego celu obrazy typu BR, które reprezentują obraz typu B w obu warstwach (rys.7). Predykcji każdego z makrobloków obrazu BR o pełnej rozdzielczości można dokonać z następujących makrobloków odniesienia:

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

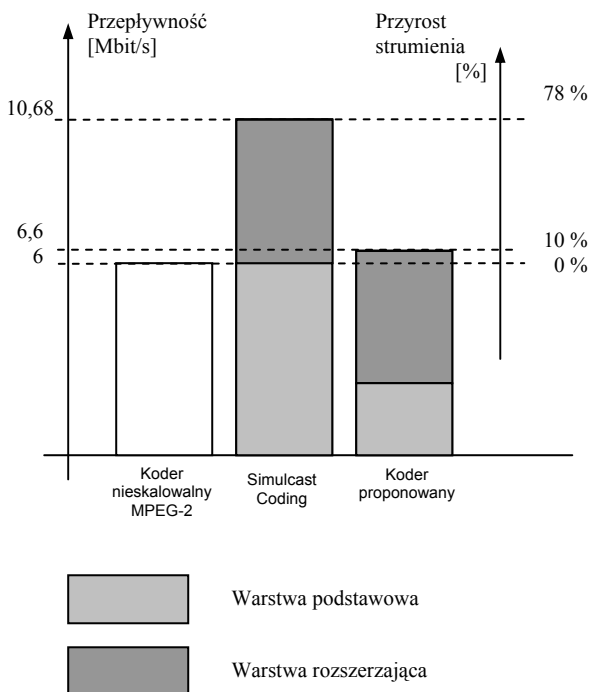
Wyniki eksperymentalne z telewizyjnymi sekwencjami testowymi dowodzą, że w predykcji znacznej części makrobloków wykorzystywany jest aktualny obraz odniesienia BR. Dlatego predykcja obrazów BE w proponowanej wersji jest bardziej dokładna. Dodatkowo obrazy typu BR są używane jako obrazy odniesienia dla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



Rys. 5. Udoskonalona predykcja obrazów typu B

6. WYNIKI EKSPERYMENTALNE

Podczas kodowania równoległego „simulcast” z wykorzystaniem dwóch niezależnych koderów MPEG-2 (jednego dla warstwy podstawowej, drugiego dla warstwy rozszerzeń) uzyskany łączny strumień skalowalny jest większy o $72 \div 88\%$ od strumienia wytworzonego przez koder nieskalowalny MPEG-2. Natomiast proponowany algorytm dla przypadku dwuwarstwowego wykazuje znacznie wyższą efektywność kodowania. Dotychczasowe eksperymenty pokazują, że dla przepływności poniżej 6Mb/s zwiększenie strumienia skalowalnego wynosi około $2 \div 10\%$ w stosunku do kodera nieskalowalnego. Są to obecnie jedne z najlepszych wyników osiągniętych na świecie.



Rys.6 Wyniki eksperymentalne dla sekwencji testowych

7. PODSUMOWANIE

Obecny etap prac można uznać za bardzo udany gdyż już w tym momencie można powiedzieć, że udało się wypracować technikę kompresji skalowalnej, która jest o wiele bardziej efektywna niż techniki przyjęte w międzynarodowych standardach. Jednym z jej największych atutów jest bardzo duża kompatybilność z istniejącymi standardami.

Obecne prace związane są z dalszym udoskonalaniem algorytmu kodowania dla przypadku wielowarstwowego i mają na celu uzyskanie efektywnego podziału bitów między warstwami rozszerzającymi w koderze skalowalnym subpasmowym wykorzystującym skalowalność mieszaną czasowo-przestrzenną.

8. 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] International Organization For Standardization, "MPEG-4 Video Verification Model", ISO/IEC JTC1/SC29/WG11, Stockholm, 1998.
- [4] T. Tsunashima, J. Stampelmann, V. Bove, "A scalable motion-compensated subband image coder," *IEEE Trans. on Communication*, vol. 42, pp. 1894-1901, 1994.
- [5] F. Bosveld, "Hierarchical video compression using SBC," Ph.D. dissertation, Delft University of Technology, Delft 1996.
- [6] H.Gharavi, W.Y.Ng, "H.263 Compatible Video Coding and Transmission," in *Proc. First International Workshop on Wireless Image/Video Communication*, pp. 115-120, Loughborough 1996.
- [7] Senbel, H. Abdel-Wahab, "Scalable and robust image compression using quadtrees," *Signal Processing: Image Communication*, vol. 14, pp. 425-441, 1999.
- [8] Shen, E. Delp, "Wavelet based rate scalable video compression," *IEEE Trans. Circuits Syst. Video Technology*, vol. 9, pp. 109-122, February 1999.
- [9] U. Benzler, "Scalable multi-resolution video coding using a combined subband-DCT approach", *Picture Coding Symposium 1999*, pp. 17-20.
- [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.
- [12] Marek Domański, Adam Łuczak, Sławomir Maćkowiak, "Spatio-Temporal Scalability for MPEG" *IEEE Transactions on Circuits and Systems for Video Technology* Vol. 10 No. 7 October 2000
- [13] Marek Domański, Adam Łuczak, Sławomir Maćkowiak, "On improving MPEG Spatial Scalability", *Proceedings of 2000 International Conference of Image Processing ICIP Vancouver*, Kanada, str II-848 - II-851
- [14] Marek Domański, Adam Łuczak, Sławomir Maćkowiak, "Spatio-Temporal Scalability using modified MPEG-2 predictive video coding", *Proceedings of European Signal Processing X*, Tampere, Finlandia 2000, str. 961-964