

Transmisja sekwencji wizyjnych w sieciach o zmiennej przepustowości

Andrzej Popławski

Streszczenie: Skalowalność jest niezwykle ważną cechą kodeków sekwencji wizyjnych, wykorzystywanych przy transmisji obrazów poprzez sieci o zmiennej w czasie przepustowości, np. sieci WLAN, GSM. Artykuł przedstawia możliwości zastosowania do takiej transmisji falkowego kodeka sekwencji wizyjnych. W pracy przedstawiono wyniki badań eksperymentalnych dla różnych wartości prędkości transmisji oraz skalowalności przestrzennej i czasowej.

Słowa kluczowe: skalowalność, sekwencje wizyjne, kodek falkowy, strumieniowanie

1. WPROWADZENIE

We współczesnym świecie sprawne przesyłanie informacji ma duże znaczenie w wielu dziedzinach. Nieustanny postęp technologiczny umożliwia na coraz większą skalę odbiór treści zarówno w postaci głosowej jak i wizyjnej. Dziś już nikogo nie dziwi wykorzystanie komputera i sieci Internet do oglądania telewizji czy też transmisja sekwencji wizyjnych poprzez sieci GSM i odtwarzanie ich w telefonie komórkowym. Możliwości przesyłu danych wizyjnych przez różnorodne sieci, charakteryzujące się odmiennymi możliwościami transmisyjnymi powodują konieczność stosowania odpowiednio przygotowanego materiału z wykorzystaniem koderów sekwencji wizyjnych.

W ostatnich latach znacznie wzrosło znaczenie skalowalności koderów sekwencji wizyjnych wykorzystywanych w sieciach komunikacyjnych. Skalowalność jest istotna w wielu zastosowaniach, na przykład 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) [18]. Skalowalność jest to cecha kodeka, która umożliwia odtworzenie z pewnej części strumienia danych prawidłowej sekwencji wizyjnej charakteryzującej się jednak gorszymi parametrami jakościowymi. Zastosowanie kodeków skalowalnych umożliwia emisję programów dla szerokiej grupy odbiorców, wykorzystujących sieci o zróżnicowanych prędkościach transmisji. W wielu zastosowaniach, np. w systemach monitoringu i nadzoru wizyjnego niedopuszczalna jest utrata wizji w przypadku ograniczonej przepustowości sieci np. na skutek awarii czy zatorów w sieci.

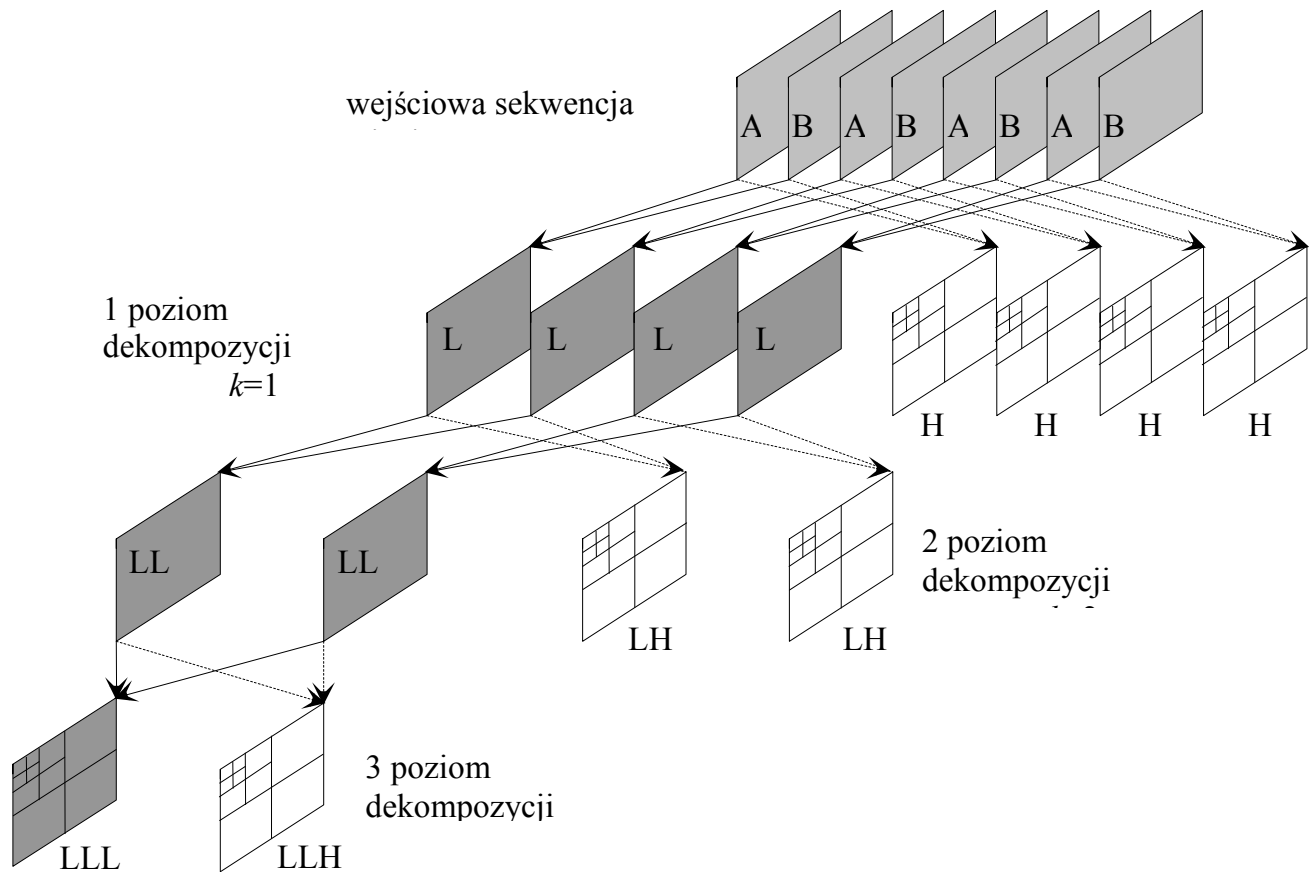
Artykuł przedstawia wyniki badań eksperymentalnych skalowalnego, falkowego koderów sekwencji wizyjnych w warunkach przesyłu informacji

w sieciach o zmiennej przepustowości. Badania przeprowadzono dla reprezentatywnej grupy testowych sekwencji wizyjnych, szeroko stosowanych i akceptowanych w wielu środowiskach naukowych.

2. SKALOWALNOŚĆ KODERÓW SEKWENCJI WIZYJNYCH

Skalowalność (*scalability*) jest to cecha kodeka wizyjnego, umożliwiająca użycie w dekodерze pewnej części strumienia binarnego wytworzonego przez koder, w celu uzyskania obrazu prawidłowego, choć charakteryzującego się gorszymi parametrami jakościowymi. Skalowalność czasowa (*temporal scalability*) umożliwia uzyskanie z części strumienia binarnego, obrazów o zmniejszonej częstotliwości ich występowania. Jeżeli obraz zrekonstruowany z części strumienia wytworzonego przez koder różni się od obrazu oryginalnego zmniejszoną rozdzielczością przestrzenną, to mówi się o skalowalności przestrzennej (*spatial scalability*). W przypadku rekonstrukcji z części strumienia skompresowanych danych obrazów o pełnej rozdzielczości przestrzennej i czasowej lecz o gorszej jakości, mówi się wówczas o skalowalności typu SNR (*SNR scalability*).

Realizacja skalowalności w kodekach hybrydowych (MPEG-2, H.264) następuje poprzez podział strumienia binarnego na kilka warstw. Każda z warstw może zawierać informacje niezbędne do zdekodowania sekwencji dla założonych rozdzielczości i przepływności strumienia danych, jednak z powodu występowania kilku warstw, część danych w strumieniu jest nadmiarowa. Dodatkowo już na etapie kodowania należy określić docelowe parametry zdekodowanej sekwencji wizyjnej, takie jak: prędkość transmisji, rozdzielczość czasowa i przestrzenna. Jakkolwiek istnieją rozwiązania wprowadzające lub udoskonalające skalowalność w najnowszych kodekach hybrydowych (kodek H.264/AVC [6] w pierwszej wersji nie obsługuje skalowalności), to jednak uzyskanie funkcji skalowalności w tych kodekach nie jest rzeczą prostą i wymaga dodatkowych zabiegów [1, 4, 11]. Skalowalność w kodekach falkowych uzyskiwana jest w naturalny sposób, co wynika ze stosownej falkowej analizy w dziedzinie czasu i przestrzeni. Ponieważ w analizie falkowej dokonuje się podziału danych na subpasma, możliwa jest rekonstrukcja obrazu z wykorzystaniem jedynie niektórych subpasm, co prowadzi do uzyskania prawidłowego obrazu o zredukowanej rozdzielczości



Rys. 1. Uproszczony schemat analizy w dziedzinie czasu kodera falkowego

przestrzennej bądź czasowej [13, 16, 17]. Skalowalność typu SNR jest z kolei uzyskiwana z wykorzystaniem technik progresywnego kodowania próbek w subpasmach, np.: EZW [12], EBCOT [14], SPIHT [10], EZBC [5].

3. SKALOWALNY FALKOWY KODER SEKWENCJI WIZYJNYCH

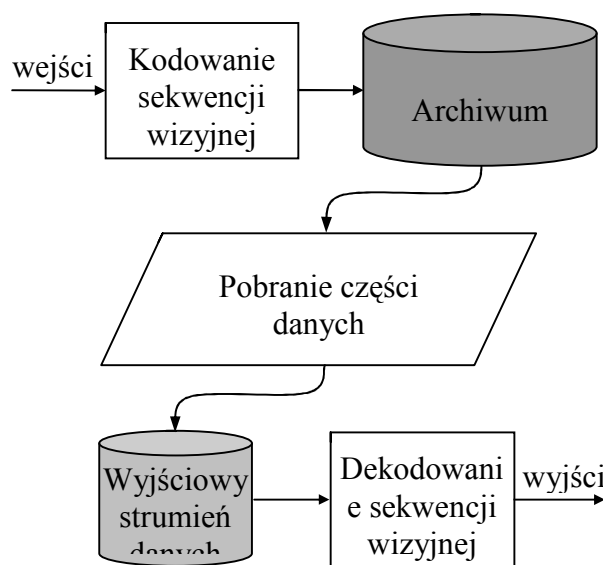
Kluczową rolę w koderach falkowych odgrywa analiza subpasmowa w dziedzinie czasu z kompensacją ruchu (MCTF – *Motion Compensation Time Filtering*). Poprzez zmniejszenie redundancji istniejącej pomiędzy poszczególnymi obrazami, wynikającej z podobieństwa (mniejszego lub większego) sąsiednich obrazów sekwencji wizyjnej możliwe jest wyeliminowanie informacji nadmiarowych i nieistotnych. W celu uzyskania efektywnej kompresji sekwencji wizyjnych stosuje się grupowanie obrazów. Grupa obrazów składająca się np. z ośmiu kolejnych obrazów poddawana jest filtracji z kompensacją ruchu, w wyniku czego powstaje jeden obraz niskoczęstotliwościowy oraz siedem obrazów wysokoczęstotliwościowych [2, 3, 8].

Na rys. 1 przedstawiono poglądowo schemat filtracji w dziedzinie czasu dla grupy składającej się z ośmiu obrazów. Analiza rozpoczyna się na pierwszym poziomie dekompozycji czasowej ($k=1$), poprzez filtrację kolejnych par obrazów wejściowych A i B. W wyniku filtracji z każdej pary powstaje składowa niskoczęstotliwościowa L oraz składowa wysokoczęstotliwościowa H (rys. 1). Na kolejnym poziomie dekompozycji czasowej ($k=2$) przetwarza się wyłącznie obrazy typu L powstałe w wyniku analizy na poprzednim poziomie, w wyniku

czego powstają składowe niskoczęstotliwościowe LL oraz składowe wysokoczęstotliwościowe LH. Analogicznie przeprowadza się analizę na pozostałych poziomach dekompozycji czasowej, która kończy się w momencie otrzymania tylko jednej składowej niskoczęstotliwościowej. Dla grupy obrazów składającej się z ośmiu obrazów, w wyniku filtracji w dziedzinie czasu otrzymuje się jedną składową niskoczęstotliwościową LLL oraz siedem składowych wysokoczęstotliwościowych: cztery H, dwie LH oraz jedną LLH (rys. 1).

Składowe powstałe w wyniku filtracji w czasie są następnie poddawane dwuwymiarowej falkowej analizie przestrzennej, np. przy wykorzystaniu filtrów Daubechies 9/7. Opis analizy falkowej w dziedzinie przestrzennej można znaleźć między innymi w [7, 15, 17].

W koderach falkowych, w przeciwieństwie do koderów hybrydowych, na etapie kodowania sygnału nie ma konieczności deklarowania prędkości transmisji czy też rozdzielczości przestrzennej i czasowej. Zakodowany koderem falkowym strumień bitowy ma charakter archiwum, z którego można odtworzyć prawidłowy obraz z prędkością transmisji i rozdzielczością określoną dopiero w momencie dekodowania strumienia (rys. 2). Z tego wynika, że dane przechowywane w archiwum mogą posłużyć do zapewnienia prawidłowego obrazu dla odbiorców z różnorodną dostępną przepustowością, dla urządzeń o różnej rozdzielczości i mocy obliczeniowej, czy też mogą być wykorzystane w przypadku zmienności dostępnej przepustowości podczas strumieniowania.



Rys. 2. Ogólny schemat kodowania falkowego

4. BADANIA EKSPERYMENTALNE

4.1. METODYKA BADAŃ

Ocena skuteczności oraz efektywności skalowalności kodeka falkowego polegała na wykonaniu eksperymentów z wykorzystaniem testowych sekwencji wizyjnych. W pierwszej kolejności dana sekwencja wizyjna została poddana kompresji, w wyniku czego otrzymano strumień bitowy (archiwum, rys. 2). Następnie poprzez określenie zadanej docelowej prędkości transmisji oraz redukcji rozdzielczości przestrzennej i czasowej dokonywano rekonstrukcji obrazu z części utworzonego wcześniej archiwum. W ostatnim kroku dokonywano porównania sekwencji oryginalnej z sekwencją zdekodowaną w celu oceny jej jakości. Jako miarę jakości wykorzystano impulsowy współczynnik sygnału do szumu PSNR (*Peak Signal-to-Noise Ratio*). Do przeprowadzenia eksperymentów wykorzystano trójwymiarowy falkowy kodek sekwencji wizyjnych MC-EZBC [9] oraz pięć testowych sekwencji wizyjnych (City, Crew, Harbour, Ice, Soccer) w formatach: CIF 30Hz (352x288), CIF 15Hz (352x288), QCIF 15Hz (176x144).

W ramach badań przeprowadzono następujące testy:

a) skalowalności typu SNR poprzez odtworzenie testowych sekwencji wizyjnych (rozdzielczość 352x288, 30Hz) dla zmniejszających się prędkości transmisji w zakresie od 1408 do 256 kb/s,

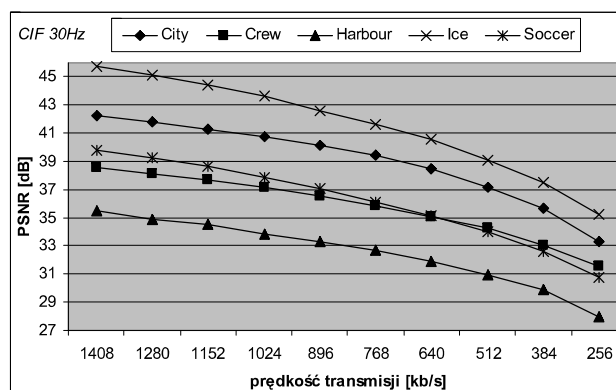
b) skalowalności czasowej i typu SNR poprzez odtworzenie testowych sekwencji wizyjnych ze zmniejszoną do 15Hz częstotliwością występowania obrazów oraz dla zmniejszających się prędkości transmisji w zakresie od 704 do 128 kb/s,

c) skalowalności czasowej, przestrzennej i typu SNR poprzez odtworzenie testowych sekwencji wizyjnych ze zmniejszoną do 15Hz częstotliwością występowania obrazów, zmniejszoną do 176x144 rozdzielczością przestrzenną oraz dla zmniejszających się prędkości transmisji w zakresie od 352 do 64 kb/s.

4.1. WYNIKI BADAŃ

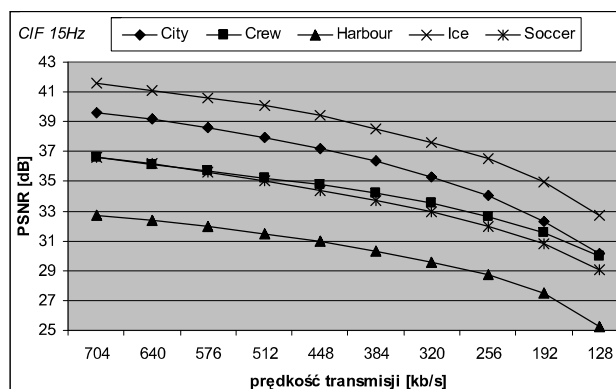
W pierwszej kolejności wykonano eksperyment polegający na symulacji stopniowego ograniczania

dostępnego pasma transmisyjnego od wartości 1408 kb/s do wartości 256kb/s, przy niezmienniej rozdzielczości przestrzennej i czasowej badanych sekwencji. Obserwowano, czy zakodowane sekwencje wizyjne zostaną prawidłowo zdekodowane oraz wyznaczano wartości współczynnika PSNR. Na rys. 3 przedstawiono efektywność kodowania dla poszczególnych sekwencji wizyjnych. Wszystkie badane sekwencje dla przedstawionych prędkości transmisji zostały prawidłowo zdekodowane przy zmniejszającej się wartości współczynnika PSNR.



Rys. 3. Efektywność kodowania w zależności od prędkości transmisji, CIF 30Hz

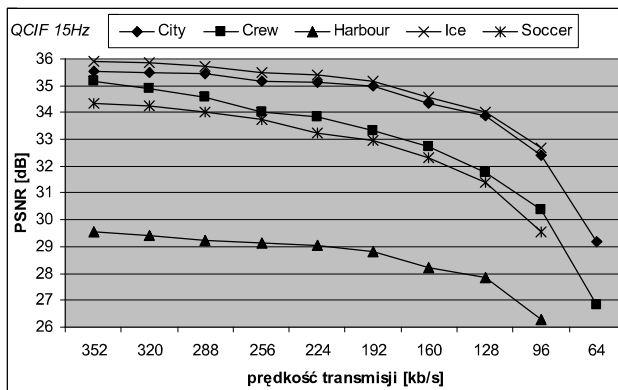
W następnym korku badano zachowanie się kodeka falkowego przy stopniowym ograniczaniu prędkości transmisji od 704kb/s do 128kb/s i jednoczesnym zmniejszeniu o połowę częstotliwości występowania obrazów w sekwencji wizyjnej z 30Hz do 15Hz. Na rys. 4 przedstawiono w formie graficznej wyniki przeprowadzonego eksperymentu. Tutaj również wszystkie badane sekwencje wizyjne zostały prawidłowo zdekodowane.



Rys. 4. Efektywność kodowania w zależności od prędkości transmisji, CIF 15Hz

W dalszej kolejności przebadano sytuację, w której zredukowano o połowę zarówno częstotliwość występowania obrazów w sekwencji wizyjnej z 30Hz do 15Hz, jak również zmniejszono rozmiary przestrzenne sekwencji z 352x288 do 176x144 punktów. Prędkość transmisji wahała się w zakresie od 352kb/s do 64kb/s. Na rys. 5 przedstawiono efektywność kodowania dla badanych sekwencji wizyjnych. W tym przypadku sekwencje Harbour, Ice oraz Soccer nie zostały

prawidłowo zdekodowane przy najmniejszej badanej prędkości transmisji wynoszącej 64kb/s.



Rys. 5. Efektywność kodowania w zależności od prędkości transmisji, CIF 30Hz

5. PODSUMOWANIE

Falkowe kodery sekwencji wizyjnych stanowią interesującą alternatywę w stosunku do klasycznych koderów hybrydowych, gdyż umożliwiają w naturalny sposób uzyskanie pełnej skalowalności. W wielu systemach, np. monitoringu i nadzoru wizyjnego, ciągłość transmisji sygnału wizyjnego jest sprawą kluczową. Możliwość stosowania rozwiązań zapewniających taką transmisję przy ograniczonym paśmie transmisyjnym, spowodowanym awarią czy też zatorami w sieci jest wysoce pożądane.

Przedstawione wyniki badań dowodzą, że kodery falkowe można z powodzeniem stosować w sytuacjach nawet znacznego ograniczenia przepustowości sieci. Warto tu podkreślić, że wszystkie badane sekwencje wizyjne zostały prawidłowo zdekodowane przy prędkości transmisji wynoszącej 96kb/s, podczas gdy przyjęto, że znamionowo prędkość ta wynosiła 1408kb/s. Oznacza to, że pomimo czterynastokrotnie zmniejszonego pasma transmisyjnego możliwe było poprawne odtworzenie oryginalnej sekwencji wizyjnej, choć oczywiście o znacznie gorszych parametrach jakościowych.

Warto również zauważyć, że możliwa jest transmisja prawidłowego sygnału wizyjnego dla różnych parametrów przestrzenno-czasowych sekwencji wizyjnej. Jak wynika z badań, dla prędkości transmisji o wartości 256kb/s można odtworzyć prawidłowy obraz w trzech rozdzielczościach: CIF 30Hz, CIF 15Hz, QCIF 15Hz.

LITERATURA

- [1] Błaszak Ł., Domański M., Lange R., "Modified AVC Codec with SNR Scalability Based on Macroblock Hierarchy", Int. Workshop Systems, Signals and Image Processing IWSSIP'04, 2004, ss. 15–18.
- [2] Chen P., "Fully scalable subband/wavelet coding", Doctoral Thesis, 2003.
- [3] Choi S. J., Woods J. W., "Motion-compensated 3-D subband coding of video", *IEEE Transactions on Image Processing*, vol. 8, 1999, ss. 155–167.
- [4] Domański M., Błaszak Ł., Maćkowiak S., "AVC Video Coders with Spatial and Temporal Scalability", *Picture Coding Symp.*, 2003, ss. 41–46.
- [5] Hsiang S.-T., Woods J. W., "Embedded image coding using zeroblocks of subband/wavelet coefficients and contest modeling", *IEEE International Symposium on Circuits and Systems*, vol. 3, 2000, ss. 662–665.
- [6] ISO/IEC/SC29/WG11/MPEG/N7081, ISO/IEC 14496-10 AVC / ITU-T Rec. H.264, Text for ISO/IEC 14496-10:2005 (AVC 3rd Edition): Information Technology – Coding of audio-visual objects – Part 10: Advanced Video Coding, MPEG 2005.
- [7] Ohm J.-R., "Multimedia Communication Technology", Springer 2004.
- [8] Ohm J.-R., "Three-dimensional subband coding with motion compensation", *IEEE Transactions Image Processing*, vol. 3, 1994, ss. 559–571.
- [9] Ruster T., Hanke K., Wien M., "Optimization for locally adaptive MCTF based on 5/3 lifting", in *Proc. Picture Coding Symposium*, 2004.
- [10] Said A., Pearlman W. A., "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees", *IEEE Transactions on Circuits and Systems for Video Tech.*, vol. 3, 1996., ss. 243–250.
- [11] Schwarz H., Marpe D., Wiegand T., "Scalable Extension of H.264/AVC", ISO/IEC JTC1/SC29/WG11, Doc. M10569/S03, MPEG 2004.
- [12] Shapiro J. M., "Embedded image coding using zerotrees of wavelet coefficients", *IEEE Transactions Signal Processing*, vol. 41, 1993, ss. 3445–3462.
- [13] Strang G., Nguyen T., "Wavelets and Filter Banks", Wellesley-Cambridge Press, 1996.
- [14] Taubman D., "High performance scalable image compression with EBCOT", *IEEE Transactions on Image Processing*, vol. 9, nr7, 2000, ss. 1158–1170.
- [15] Topiwala P. N., "Wavelet image and video compression", Kluwer Academic Publishers, 1998.
- [16] Vetterli M. and Kovačević J., "Wavelets and Subband Coding", Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [17] Woods J. W., "Subband Image Coding", Kluwer Academic Publishers, 1991.
- [18] Wu D., Hou Y., Zhang Y., "Scalable video coding and transport over broad-band wireless networks", *Proc. of the IEEE*, vol. 89, 2001, ss. 6–20.



dr inż. Andrzej Popławski
 Uniwersytet Zielonogórski
 Wydział Elektrotechniki, Informatyki
 i Telekomunikacji
 Instytut Informatyki i Elektroniki
 ul. prof. Z. Szafrana 2
 65-246 Zielona Góra
 tel.: 68 328 22 19
 e-mail: A.Poplawski@iie.uz.zgora.pl