

**Paweł TURCZA, Mariusz DUPLAGA**  
KATEDRA METROLOGII (AKADEMIA GÓRNICZO-HUTNICZA)  
COLLEGIUM MEDICUM UNIwersYTET Jagielloński

## Algorytm kompresji obrazów dla bezprzewodowej kapsuły endoskopowej

Dr inż. Paweł TURCZA

Absolwent kierunków Informatyka i Elektronika Wydziału EAIiE Akademii Górniczo-Hutniczej. Związany zawodowo z AGH od 1993 roku, tytuł doktora w dyscyplinie Elektronika uzyskał w roku 2001. Zajmuje się algorytmami przetwarzania sygnałów w zastosowaniach telekomunikacyjnych i pomiarowych oraz metodami ich efektywnej implementacji.

e-mail: turcza@agh.edu.pl



Dr med. Mariusz DUPLAGA

Lekarz medycyny, specjalizacja z chorób wewnętrznych. Obecnie na stanowisku adiunkta w Zakładzie Informacji Naukowej Instytutu Zdrowia Publicznego, część etatu w Klinice Pulmonologii CMUJ. Uczestnik lub lokalny lider szeregu międzynarodowych projektów badawczo-rozwojowych Unii Europejskiej. Główne obszary zainteresowań: systemy komputerowe wspomagające zabiegi endoskopowe, e-zdrowie, telemedycyna, telemonitorowanie.

e-mail: mmduplag@cyf-kr.edu.pl



### Streszczenie

Artykuł prezentuje algorytm kompresji obrazów planowany do zastosowania w bezprzewodowej kapsule endoskopowej. Algorytm przeznaczony do tego typu zastosowań oprócz możliwie wysokiego stopnia kompresji musi cechować się bardzo niskim poborem mocy. To wymaganie wyklucza użycie standardowych metod. Proponowany algorytm oparty jest o całkowitoliczbowe wersje transformacji DCT i transformacji falkowej oraz koder Huffmana. W porównaniu do algorytmów konkurencyjnych proponowany algorytm oferuje znacznie większy stopień kompresji przy nieco większej (głównie pamięciowej) złożoności.

**Słowa kluczowe:** Kompresja obrazów, Kapsuła endoskopowa

### Image compression algorithm for wireless endoscopy capsule

#### Abstract

The paper describes image compression algorithm suitable for wireless capsule endoscopy. Due to power limitation and small size conditions traditional image compression techniques are not appropriate and dedicated ones are necessary. The proposed algorithm is based on integer version of discrete cosine transform (DCT) and wavelet transform (DWT) with Huffman entropy coder. Thanks to integer DCT/wavelet application it has low complexity and power consumption. Additionally, the algorithm can provide lossless compression as well as high-quality lossy compression.

**Keywords:** Image compression, Video endoscopy capsule

## 1. Wprowadzenie

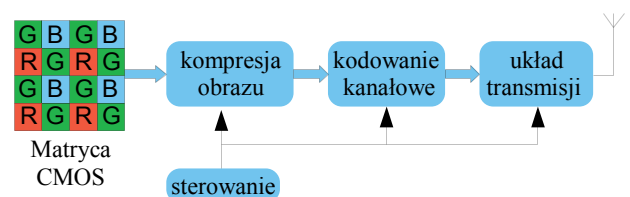
W wielu schorzeniach przewodu pokarmowego, podstawą właściwego rozpoznania jest ocena endoskopowa, często połączona z pobraniem materiału ze zmienionych okolic. Do niedawna badanie endoskopowe było podstawową metodą diagnostyczną pozwalającą określić stan wybranych odcinków przewodu pokarmowego; górnego w przypadku ezofagogastroduodenoskopii lub dolnego w przypadku kolonoskopii. Badanie endoskopowe wiąże się niewątpliwie z nieprzyjemnymi doznaniem dla pacjenta. Postęp jaki dokonał się w mikroelektronice pozwolił na skonstruowanie bezprzewodowej kapsuły endoskopowej umożliwiającej bezinwazyjną ocenę całego przewodu pokarmowego. Pierwsza taka kapsuła została zbudowana przez firmę Given Imaging Ltd. [1]. Kapsuła wyposażona jest w sensor CMOS, system oświetlający i transmisyjny (rys. 1). Po połączniu przez pacjenta przemieszcza się ona wewnątrz przewodu pokarmowego (pod wpływem ruchów perystaltycznych jelit) wykonując zdjęcia. Zdjęcia te są bezprzewodowo transmitowane

do odbiornika noszonego przez pacjenta. Niestety nie jest to jeszcze dojrzała technologia. Choć kapsuła przemieszcza się przez całą długość przewodu pokarmowego brak możliwości sterowania nią uniemożliwia dokładne badania żołądka, który ma stosunkową dużą objętość. Natomiast, dokładne badanie ścian jelita grubego możliwe jest dopiero po jego inflacji, gdyż normalnie jest obkurczone. Jakość obrazu dostarczana przez obecnie produkowane kapsuły pozostawia wiele do życzenia. Również liczba uzyskiwanych obrazów na jednostkę czasu wynosząca obecnie 2 obrazy na sekundę, w porównaniu do standardowych 25, jest zbyt niska. Z przeprowadzonych analiz [2], [3] wynika, że ze względu na ograniczenie pobieranej mocy i znaczne tłumienie fal radiowych przez ciało system transmisyjny kapsuły może osiągnąć przepustowość raczej nie przekraczającą 1Mb/s. Sensor CMOS o rozdzielczości VGA (640 x 480) pixeli dostarcza  $2.45 \times 10^6$  bitów dla każdego obrazu. Widać stąd od razu, że bez zastosowania kompresji danych transmisja jednego obrazu trwałaby by ponad 2 sekundy zużywając przy tym sporo energii.

Ponieważ istniejące algorytmy kompresji obrazów, ze względu na znaczną złożoność obliczeniową, niezbyt nadają się do zastosowania w kapsule poszukuje się algorytmów dedykowanych [3] [4]. Celem niniejszej pracy jest zaproponowanie wydajniejszego algorytmu niż przedstawiony w [3] [4], mogącego znaleźć zastosowanie w kapsule. Proponowany algorytm oparty jest na całkowitoliczbowej wersji transformacji kosinusowej (DCT) i falkowej (DWT), co z jednej strony umożliwia jego efektywną, niskomocową sprzętową implementację z drugiej pozwala by algorytm pracował zarówno w trybie kompresji bezstratnej jak i kompresji stratnej o wysokiej jakości.

## 2. Proponowany algorytm kompresji

Uproszczony schemat przetwarzania sygnałów w kapsule endoskopowej przedstawia rys. 1. W pierwszym etapie przetwarzania, obrazy otrzymywane z matrycy CMOS poddawane są kompresji celem redukcji wymaganego pasma transmisyjnego. Dalej, uzyskany strumień danych poddawany jest operacji kodowania protekcyjnego, celem zabezpieczenia przed



Rys. 1. Schemat przetwarzania sygnałów w kapsule endoskopowej  
Fig. 1. Block diagram of data processing in wireless endoscopy capsule

błędami transmisji. Tak przygotowany strumień danych transmitowany jest drogą radiową do odbiornika noszonego przez pacjenta i tam rejestrowany celem późniejszej analizy. Całość pracy urządzenia jest nadzorowana przez układ mikrokontrolera.

Proponowany algorytm kompresji realizuje sekwencyjnie cztery, niżej opisane, operacje: transformacja przestrzeni kolorów, transformacja danych obrazowych, kwantyzacja współczynników transformacji oraz ich kodowanie entropowe.

### 2.1. Transformacja przestrzeni kolorów

Układ matrycy CMOS dostarcza obraz w przestrzeni barw podstawowych tj. czerwonej (R), zielonej (G) niebieskiej (B) [4]. Ponieważ, jak wiadomo, poszczególne składowe R, B, G są silnie skorelowane korzystna jest zmiana przestrzeni kolorów. W tym celu, zwykle wykorzystywana jest transformacja

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.146 & -0.288 & -0.434 \\ 0.617 & -0.517 & 0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1)$$

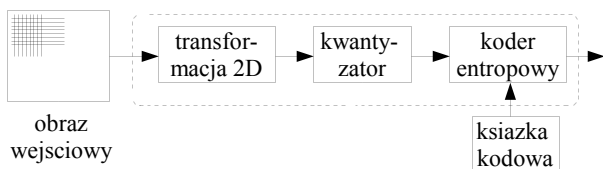
gdzie  $Y$  jest luminancją obrazu, a  $Cb$  i  $Cr$  reprezentują tzw. składowe różnicowe chrominancji: niebieskiej i czerwonej, odpowiednio. Tego typu transformacja wymaga dość wysokiej precyzji obliczeń, a więc skomplikowanego układu elektronicznego i dlatego nie jest najlepszym rozwiązaniem dla kapsuły. W związku z tym proponuje się zastosowanie przekształcenia [6]

$$Y = \frac{1}{2} \left( G + \frac{R+B}{2} \right); \quad C_g = \frac{1}{2} \left( G - \frac{R+B}{2} \right); \quad C_o = \frac{R-B}{2}; \quad (2)$$

gdzie  $Cg$  jest składową różnicową chrominancji zielonej, a  $Co$  pomarańczowej. Jak widać powyższe przekształcenie może być zrealizowane w bardzo prosty sposób.

### 2.2. Transformacja obrazu

Schemat kodera transformacyjnego (dla jednej składowej) przedstawiony jest na rys 2. Pierwszą operacją jest wyznaczenie dwuwymiarowej (2D) transformacji poszczególnych składowych obrazu  $Y, Cg, Co$ . Zadaniem transformacji jest skoncentrowanie energii próbek obrazu w możliwie małej liczbie współczynników transformacji, których wartości następnie będzie można efektywnie zakodować i przesłać. Do tego celu wykorzystuje się zwykle dyskretną transformację kosinusową (DCT) lub falkową (DWT). Ze względu na wspomniane ograniczenia implementacyjne, w prezentowanym algorytmie, wykorzystano ich wersje całkowitoliczbowe.



Rys.2. Schemat blokowy kodera transformacyjnego  
Fig. 2. Block diagram transform based image coder

#### 2.2.1. Transformacja DCT

Transformacja 2D DCT jest transformacją separowalną, tzn. wyznaczana jest jako złożenie dwóch transformacji jednowymiarowych (1D) wykonywanych po wierszach i kolumnach bloku obrazu, tzn. transformacja DCT bloku  $X$  określona jest

$$Y = T X T^T \quad (3)$$

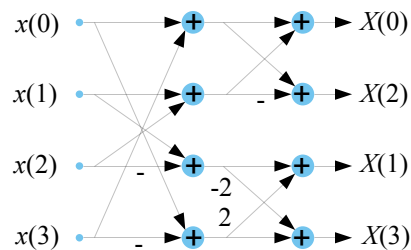
gdzie  $T$  oznacza macierz transformacji DCT, a indeks górny  $T$

transpozycję. W proponowanym algorytmie testowano trzy transformacje DCT: standardową transformację DCT o rozmiarze  $8 \times 8$  i dwie całkowitoliczbowe aproksymacje o macierzach  $T_{4 \times 4}$  i  $T_{8 \times 8}$ . Transformacja DCT oprócz dobrych właściwości dekorrelacyjnych posiada także efektywny algorytm obliczeniowy. Efektywny algorytm obliczeniowy dla 4-punktowej transformacji DCT [5] o macierzy  $T_{4 \times 4}$  przedstawiony jest na rys. 3, a dla transformacji 8-punktowej [6] o macierzy  $T_{8 \times 8}$  na rys 4. Oczywiście jest, że żaden z algorytmów nie wymaga operacji mnożenia. Wymagane są jedynie dodawanie i przesunięcie bitowe (przykładowo algorytm DCT4 wymaga ośmiu operacji dodawania i dwóch przesunięć).

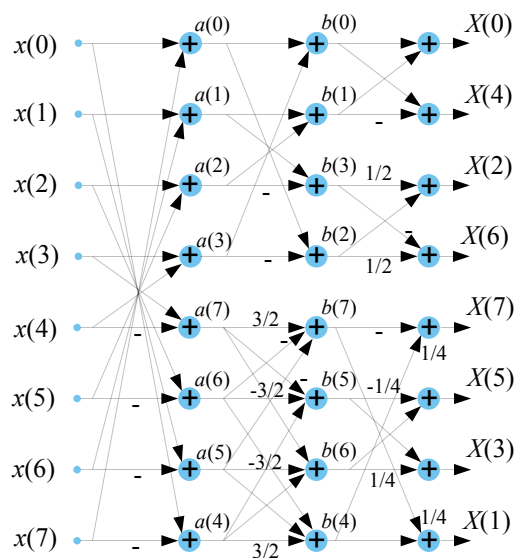
Po wykonaniu transformacji, uzyskane współczynniki ( $Y$ ) podlegają operacji kwantyzacji skalarnej, co z jednej strony zapewnia wysoki stopień kompresji, a z drugiej jest przyczyną bezpowrotnej utraty informacji (jakości) w rekonstruowanym obrazie.

$$T_{4 \times 4} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \quad (4)$$

$$T_{8 \times 8} = \begin{bmatrix} 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ 12 & 10 & 6 & 3 & -3 & -6 & -10 & -12 \\ 8 & 4 & -4 & -8 & -8 & -4 & 4 & 8 \\ 10 & -3 & -12 & -6 & 6 & 12 & 3 & -10 \\ 8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 \\ 6 & -12 & 3 & 10 & -10 & -3 & 12 & -6 \\ 4 & -8 & 8 & -4 & -4 & 8 & -8 & 4 \\ 3 & -6 & 10 & -12 & 12 & -10 & 6 & -3 \end{bmatrix}$$



Rys. 3. Efektywny algorytm obliczeniowy dla 4 punktowej DCT  
Fig. 3. Low complexity 4-point DCT butterfly algorithm

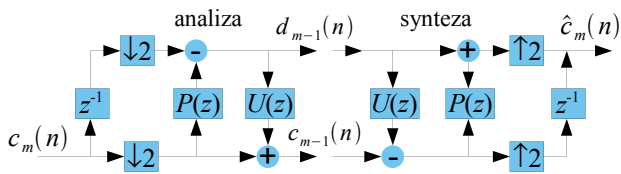


Rys. 4. Efektywny algorytm obliczeniowy dla 8 punktowej DCT  
Fig. 4. Low complexity 8-point DCT [8] butterfly algorithm

## 2.2.2. Transformacja falkowa

Transformacja DCT wyznaczana jest dla rozłącznych bloków obrazu. Oznacza, to występowanie nieciągłości funkcji bazowych na granicach bloków. W przypadku silnej kwantyzacji współczynników transformacji, nieciągłości te uwidaczniają się w rekonstruowanym obrazie jako tzw. efekt blokowy. Wady tej nie posiadają transformacje falkowe, których funkcje bazowe nakładają się i łagodnie gasną na granicach nośnika. Dlatego są one chętnie wykorzystywane w algorytmach kompresji obrazów.

W pracy [7] wykazano, że każdą transformację falkową, ortogonalną bądź biortogonalną, można przedstawić w postaci schematu predykcja-uaktualnienie, który ma formę drabinki. W tym schemacie najprostszą realizację posiadają transformacje interpolacyjne [8] złożone z dwóch filtrów: filtru predykcji  $P$  i uaktualnienia  $U$ . Schemat takiej transformacji przedstawiony jest na rys. 5. Wartości współczynników stosowanych filtrów  $P$  i  $U$  zamieszczone w tablicy 1. Widać, że do realizacji tych filtrów nie potrzebne są układy mnożące. Dokładniejszy opis projektowania transformacji P-U można znaleźć w pracy [7].



Rys. 5. Schemat blokowy jednego poziomu prostej i odwrotnej transformacji falkowej P-U

Fig. 5. Block diagram of single level of forward and inverse P-U wavelet transform

Tab. 1. Wartości współczynników  $P$  i  $U$  wykorzystanych w koderze  
Tab. 1. Coefficients of filters  $P$  and  $U$  used in proposed coder

	$p(n)$	$u(n)$
$N = 2$	$2^{-1}, 2^{-1}$	$2^{-2}, 2^{-2}$
$N = 4$	$-2^{-4}, 2^{-1}+2^{-4}, 2^{-1}+2^{-4}, -2^{-4}$	$-2^{-5}, 2^{-2}+2^{-5}, 2^{-2}+2^{-5}, -2^{-2}$

Dwuwymiarowe (2D) transformacje, zarówno DCT jak i DWT wyznaczane są dla wierszy i kolumn obrazu. Wyznaczanie transformaty dla próbek wierszy można zrealizować bardzo efektywnie podczas ich akwizycji. Realizacja transformacji wzdłuż kolumn obrazu wymaga użycia buforów o znacznych rozmiarach, co w omawianym zastosowaniu (kapsuła), jest poważnym problemem. W proponowanym algorytmie problem ten zminimalizowano poprzez zastosowanie, w transformacji realizowanej po kolumnach obrazu, filtrów  $P$  i  $U$  o bardzo krótkich odpowiedziach impulsowych ( $N=2$ ). Do transformacji próbek w wierszach wykorzystano filtry o dłuższych odpowiedziach impulsowych ( $N=4$ ).

## 2.3. Kodowanie entropowe współczynników

Ostatnim etapem realizowanym przez algorytm kompresji jest kodowanie entropowe skwantowanych współczynników transformacji DCT i DWT. W proponowanym algorytmie współczynniki transformacji falkowej grupowane są w bloki  $8 \times 8$  [9], tak jak przedstawia to rys 6. Tak uzyskane bloki mają analogiczne właściwości jak bloki współczynników uzyskiwanych w wyniku transformacji 2D DCT, tzn. element (L3) reprezentuje składową stałą (DC), a elementy na drugim końcu diagonali (D1) składowe o najwyższych częstotliwościach (AC). Taki sposób uporządkowania współczynników pozwala na wykorzystanie do ich kodowania efektywnej metody znanej ze standardu JPEG [10].

Ponieważ współczynniki DC w przyległych blokach są nadal silnie skorelowane i posiadają duże wartości, kodowane są różnicowo. W wyniku transformacji i kwantyzacji ponad 90% współczynników ma wartość zero. Dlatego, w pierwszym kroku

kodowania entropowego, reszta współczynników (AC) każdego z bloków reprezentowana jest przez parę liczb

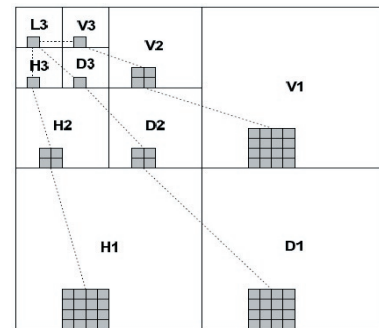
$$(RL, v) \quad (5)$$

gdzie  $RL$  oznacza liczbę współczynników zerowych (zer) poprzedzających współczynnik niezerowy ( $v$ ). W przypadku, gdy od pewnej pozycji w bloku, aż do końca występują same zera reprezentowane są one przez parę  $(0, 0)$ .

W drugim etapie kodowania, przy użyciu algorytmu Huffmiana, utworzone pary  $(RL, v)$  reprezentowane są w postaci kodów o zmiennej długości (VLC - *variable-length code*). Reprezentacja ta składa się z dwóch części. Część pierwsza

$$z = 16RL + |v|_2 \quad (6)$$

koduje liczbę zer ( $RL$ ) oraz długość rozwinięcia dwójkowego ( $|v|_2$ ) modułu współczynnika  $v$ . Wartość ta ( $z$ ) zapisywana jest przy użyciu kodu Huffmiana. Bity stanowiące reprezentację dwójkową współczynnika  $v$  (ich liczba jest już zakodowana w  $z$ ) zapisywane są do strumienia danych bezpośrednio, jeśli  $v > 0$  lub w postaci zanegowanej jeśli  $v < 0$ .



Rys. 6. Relacje przestrzenno-częstotliwościowe pomiędzy współczynnikami transformacji falkowej

Fig. 6. Space-frequency relationships in wavelet coefficients

## 3. Wyniki

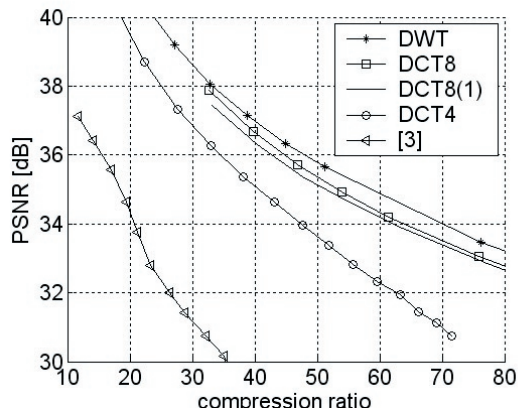
Przedstawiony algorytm został zaimplementowany w postaci programu komputerowego. Wyniki jego działania zostały porównane z wynikami algorytmu konkurencyjnego [3]. W porównaniu użyto przykładowego obrazu pochodzącego z gastroskopii i kolonoskopii. Jako obiektywną miarę jakości rekonstrukcji przyjęto wartość PSNR (*peak signal to noise ratio*) zdefiniowaną w następujący sposób

$$PSNR = 10 \log_{10} \frac{3 \times N \times M \times 255^2}{\sum_{c=1}^3 \sum_{n=1}^N \sum_{m=1}^M (\hat{x}_{n,m,c} - x_{n,m,c})^2} \quad (7)$$

gdzie  $\hat{x}_{n,m,c}$  oznacza wartość pixela rekonstruowanego obrazu o współrzędnych  $(n,m)$  i składowej koloru  $c$ . Zależność jakości rekonstrukcji, określonej wartością PSNR, od stopnia kompresji ( $CR$ ) przedstawiono na rys. 7a i 7b.

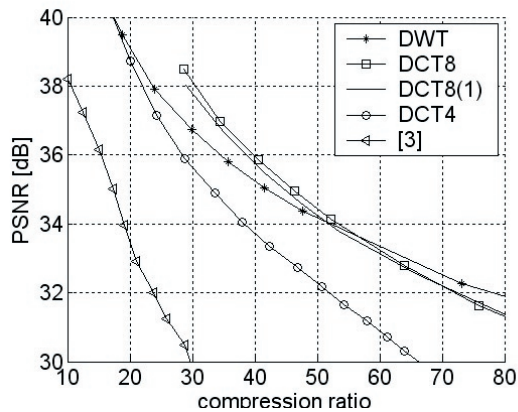
Na rysunkach 7a i 7b poprzez etykiety DCT8 i DCT8(1) oznaczono krzywe zależności PSNR od stopnia kompresji dla algorytmu opartego o standardową 8-punktową transformację DCT i odpowiednio jej aproksymację całkowitoliczbową. Etykietą DCT4 oznaczono wyniki uzyskane przez algorytm oparty o 4-punktową całkowitoliczbową transformację DCT (o macierzy  $T_{4 \times 4}$ ), a przez DWT wyniki transformacji falkowej. Z powodu szczupłości miejsca, na rysunkach 8 i 9 zaprezentowano tylko jeden przykładowy obraz z kompresji metoda opartą o DCT4 (najprostszą) i algorytmem [3], odpowiednio. Biorąc pod uwagę wyniki przedstawione na rysunkach 7a i 7b (kryterium obiektywne) oraz subiektywną jakość rekonstrukcji obrazów kompresowanych poszczególnymi metodami nie ulega wątpliwości, że proponowany algorytm znacznie przewyższa

jakością algorytm [3]. Biorąc pod uwagę złożoność obliczeniową, wymagania pamięciowe i jakość rekonstrukcji uzyskaną przez metody oparte o poszczególne transformacje wydaje się, że najlepszą metodą do zastosowania w kapsule jest metoda oparta o transformację falkową.



Rys. 7a. Wyniki algorytmu dla przykładowego obrazu pochodzącego z kolonoskopii

Fig. 7a. Results for exemplary image from colonoscopy



Rys. 7b. Wyniki algorytmu dla przykładowego obrazu pochodzącego z gastrokopii

Fig. 7b. Results for exemplary image from gastroscopy

#### 4. Wnioski końcowe

W artykule zaproponowano dwie wersje algorytmu kompresji obrazów przeznaczonego dla bezprzewodowej kapsuły endoskopowej. Algorytm bazuje na całkowitoliczbowych transformacjach sygnału i entropowym koderze Huffmana. Dzięki temu możliwa jest jego prosta implementacja sprzętowa o niskim poborze mocy. W trakcie eksperymentów wykazano, że proponowany koder osiąga ponad dwa razy większy stopień kompresji niż koder konkurencyjny przy około dwukrotnie wyższej złożoności implementacji (zajętość pamięci). Dalsze prace będą dotyczyły implementacji sprzętowej kodera (wstępna implementacja w FPGA) i algorytmów kodowania protekcyjnego.

#### ACKNOWLEDGMENT

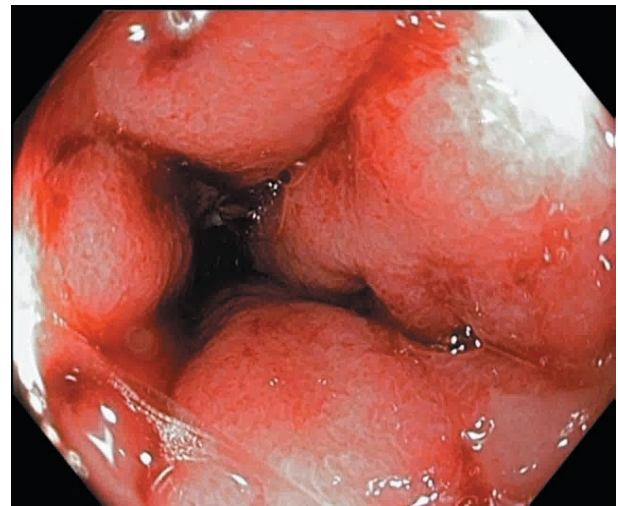
The research activities presented in this paper were conducted under the European Commission R&D Project No 3E061105 (VECTOR).

Artykuł przygotowano przy pomocy pakietu OpenOffice.

#### 5. Literatura

[1] G. Iddan, G. Meron, A. Glukhovskiy, and P. Swain: Wireless capsule endoscopy. *Nature*, no. 6785, 2000, str. 417–418.

- [2] M. Mylonaki, A. Fritscher-Ravens, P. Swain: Wireless capsule endoscopy: a comparison with push enteroscopy in patient with gastroscopy and colonoscopy negative gastrointestinal bleeding. *Gut J.*, 2003, str. 1122–1126.
- [3] D. Turgis, R. Puers: Image compression in video radio transmission for capsule endoscopy. *Sensors and Actuators A*, 2005, str. 129–136.
- [4] Xiang Xie, Guo Lin Li, and Zhi Hua Wang: A Near-Lossless Image Compression Algorithm Suitable for Hardware Design in Wireless Endoscopy System. *EURASIP Journal on Advances in Signal Processing*, Vol. 2007, Article ID 82160.
- [5] H. S. Malvar, A. Hallapuro, M. Karczewicz, L. Kerofsky: Low-Complexity Transform and Quantization in H.264/AVC. *IEEE Trans. on Circuits and Systems for Video Tech.*, no. 7, 2003.
- [6] ITU-T Rec. H.264 / ISO/IEC 11496-10: Advanced Video Coding, Final Committee Draft, Document JVTf100, December 2002.
- [7] I. Daubechies, W. Sweldens: Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl.*, no. 3, 1998, str. 245–267.
- [8] D.L. Donoho: Interpolating wavelet transforms. Preprint, Department of Statistics, Stanford University, 1992.
- [9] Z. Xiong, K. Ramchandran, M. T. Orchard, Y.-Q. Zhang: A Comparative Study of DCT and Wavelet-Based Image Coding. *IEEE Trans. on Circuits and Systems for Video Tech.*, no. 5, 1999, str. 692–695.
- [10] G.K. Wallace: The JPEG still picture compression standard. *Communications of the ACM*, Vol. 34, April 1991, str. 30–44.



Rys. 8. Przykładowy obraz zrekonstruowany metodą opartą o DCT4,  $PSNR=36.49$ ,  $CR=32$

Fig. 8. Exemplary image reconstructed using algorithm based on DCT4,  $PSNR=36.49$ ,  $CR=32$



Rys. 9. Przykładowy, zrekonstruowany obraz metodą [3],  $PSNR=30.74$ ,  $CR=32$

Fig. 9. Exemplary image reconstructed using algorithm [3],  $PSNR=30.74$ ,  $CR=32$