

Wiesław PAMUŁA

WŁASNOŚCI TRANSFORMATY FALKOWEJ DLA KOMPRESJI DANYCH Z WIDEOREJESTRATORA ZDARZEŃ DROGOWYCH

Streszczenie. Zaproponowano wykorzystanie transformaty falkowej obliczanej po krzywej Hilberta, poprowadzonej na zarejestrowanym obrazie, dla kompresji danych obrazowych. Bieg krzywej grupuje lokalnie piksele obrazu wspomagając aproksymację pojazdów na obrazie. Obliczenie współczynników jednowymiarowej transformaty falkowej wymaga mniejszej około 30% liczby operacji przetwarzania i nadaje się do implementacji z użyciem mikrokontrolerów.

PROPERTIES OF WAVELET TRANSFORMS FOR COMPRESSING DATA FROM A VIDEORECORDER OF TRAFFIC INCIDENTS

Summary. An algorithm is proposed to incorporate Hilbert scan of image data and a wavelet transform, factorized into lifting steps, to process video traffic data, which is suitable for on site microcontroller implementation. Scanning an image, in space filling curve order, brings together pixels that are highly correlated. This scan - a vector of the image, represents clustered pixels, which may be more effectively transformed. Applying a 1-dimensional wavelet transform requires a smaller number of processing steps.

1. WSTĘP

Poszukiwanie efektywnej metody kompresji danych obrazowych koncentruje się na optymalizacji metod opartych na transformatach falkowych [2]. Aplikacje włączone do systemów przekazu danych z kamer monitorujących ruch drogowy muszą działać w czasie rzeczywistym. Jednocześnie kompresja powinna być zaimplementowana z wykorzystaniem dostępnych na stanowiskach obserwacji środków przetwarzania, którymi są mikrokontrolery sprawnie realizujące jedynie operacje całkowitoliczbowe.

2. KRZYWE WYPEŁNIAJĄCE PRZESTRZEŃ

Krzywe wypełniające przestrzeń są odwzorowaniami odcinka na zadaną 2-wymiarową powierzchnię. Krzywa przebiega wszystkie punkty tej powierzchni. Pierwsze takie krzywe zaproponowali Peano i Hilbert w końcu XIX wieku [7]. Sąsiednie punkty krzywej odpowiadają sąsiednim w sensie otoczenia dwuwymiarowego. Przeglądanie obrazu po takiej krzywej odpowiada zatem analizie kolejnych otoczeń, gdzie wartości pikseli mogą być silnie powiązane ze sobą. Przykładem szeroko znanym wykorzystania korelacji między pikselami może być segmentacja obrazu z wykorzystaniem drzew czwórkowych, będąca podstawą wielu słownikowych metod kompresji danych obrazowych.

Krzywe wypełniające przestrzeń stosowane są w algorytmach cieniowania dla drukarek i wyświetlaczy LCD, dając bezkierunkowe przejścia między odcieniami lub eliminując uciążliwe dla oka schodkowe zmiany kolorów [6].

Znane metody kompresji słownikowej wykorzystują podział obrazu na fragmenty i indywidualne dla każdego fragmentu dopasowanie wpisów ze słownika. Zawartość fragmentów analizowana jest po krzywych Hilberta [2].

Proponuje się odwzorowanie całego obrazu w krzywą wypełniającą przestrzeń i przetwarzanie pełnego obrazu zamiast jego fragmentów. Kolejne kroki algorytmu operować będą na jednowymiarowej reprezentacji obrazu.

Wektor obrazu konstruuje się przez przepisanie zawartości bufora obrazu w kolejności punktów krzywej wypełniającej przestrzeń.

$$p(i) = P(a(1, i), a(0, i)) \quad (1)$$

gdzie:

$P(i, j)$ – bufor obrazu,

$p(i)$ – wektor obrazu,

$a(j, i)$ – tablica adresów.

Kolejność określona jest przez tabelę adresów. Dla ułatwienia implementacji zastosowano krzywą Hilberta i przyjęto do przetwarzania kwadratowe fragmenty obrazów z kamery o rozmiarach 2^n pikseli.

W przypadku prostokątnych obrazów można zastosować uogólnioną krzywą *ghil*, *mur-ray* lub zmodyfikowaną Peano, które posiadają podobne własności, różniąc się sposobem generacji kolejnych punktów pokrycia powierzchni [3].

3. CAŁKOWITOLICZBOWE TRANSFORMATY FALKOWE

Obliczanie współczynników całkowitoliczbowej transformaty falkowej jest oparte na tzw. „podnoszeniu” (lifting), które polega na iteracyjnym wyliczaniu wartości współczynników aktualnego pasma rozdzielczości na podstawie kombinacji liniowej współczynników niższych rozdzielczości [1]. Wyniki obliczeń zaokrągla się. Pierwszymi współczynnikami są wartości pikseli obrazu. Transformata jest odwracalna, co oznacza, że można odtworzyć obraz bez strat [4].

Przeprowadzono badanie własności kilku transformat całkowitoliczbowych. Wybrano transformaty z ułamkowymi współczynnikami, które łatwo można obliczać z użyciem sprzętu zawierającego mikrokontrolery.

Obrazy zdarzeń drogowych zawierają obiekty o ograniczonym zakresie rozmiarów. Jednocześnie obiekty te mogą być aproksymowane przez prostopadłościany lub ich złożenia. Wystarczającą reprezentację falkową takich obiektów uzyskuje się z wykorzystaniem symetrycznych falek o niewielkiej liczbie momentów znikających.

Komitet normalizacyjny JPEG2000 po wieloletnich badaniach przyjął falekę (5,3) jako podstawę bezstratnych metod kompresji w nowym standardzie [9]. Jest to kolejną wskazówką, która pozwala dalej ograniczyć zbiór transformat falkowych uwzględnianych w badaniach.

Dla ilustracji zasady obliczania transformatę falkową Haara można przedstawić za pomocą następujących kroków “podnoszenia” (lifting steps):

$$\begin{aligned} d_{1,i} &= s_{0,2i+1} - s_{0,2i} \\ s_{1,i} &= s_{0,2i} - \left\lfloor \frac{d_{1,i}}{2} \right\rfloor \end{aligned} \quad (2)$$

gdzie:

$d_{n,i}$ – współczynniki różnicowe n -tego poziomu rozdzielczości (subpasma),
 $s_{n,i}$ – współczynniki n -tego poziomu rozdzielczości (subpasma).

Inne konstrukcje wymieniono w tablicy 1.

Tablica 1

Całkowitoliczbowe transformaty falkowe

Nazwa	Lifting steps	Uwagi
D(2,2)	$d_{1,i} = s_{0,2i+1} - \lfloor (s_{0,2i} + s_{0,2i+2})/2 + 1/2 \rfloor$ $s_{1,i} = s_{0,2i} - \lfloor (d_{1,i-1} + d_{1,i})/4 + 1/2 \rfloor$	symetryczne biortogonalne
D(4,4)	$d_{1,i} = s_{0,2i+1} - \lfloor 9(s_{0,2i} + s_{0,2i+2})/16 - (s_{0,2i-2} + s_{0,2i+4})/16 + 1/2 \rfloor$ $s_{1,i} = s_{0,2i} - \lfloor 9(d_{1,i-1} + d_{1,i})/32 - (d_{1,i-2} + d_{1,i+1})/32 + 1/2 \rfloor$	symetryczne biortogonalne

4. KOMPRESJA DANYCH OBRAZOWYCH

Zaimplementowano algorytm transformaty falkowej jednowymiarowej po krzywej Hilberta i klasycznej dwuwymiarowej za pomocą pakietu matematycznego Matlab. Przeprowadzono badania sprawności kompresji dla zbioru reprezentatywnych obrazów zdarzeń drogowych.

Nie podjęto prób realizacji kompletnego algorytmu kompresji, uznając cechy zbioru otrzymanych współczynników za miarę zdolności kompresowania. Przyjęto średnią ważoną entropię E zapisu współczynników dla przeprowadzania porównań ilościowych. Większa wartość entropii oznacza, że zapis wymaga większej liczby bitów, stąd współczynnik kompresji będzie mniejszy.

$$E = -\sum_s \frac{1}{s} \sum P_i \log_2 P_i \quad (3)$$

gdzie:

P_i – znormalizowana wartość współczynnika i -tego subpasma,
 s – rozmiar subpasma.

Entropia uśredniana jest dla każdego subpasma współczynników, aby uwzględnić ich różne własności statystyczne.

Rysunek 1 prezentuje wybrane reprezentatywne klatki obrazów poddanych kompresji. Użyto kilkaset obrazów przedstawiających trzy grupy sytuacji drogowych:

- ruch uliczny,
- ruch autostradowy,
- zjazdy.

Zbadano dwie metody redukcji entropii zapisu transformaty obrazu przez usuwanie współczynników oraz przez zmniejszenie liczby różnych wartości współczynników. Wykorzystano tylko operacje oparte na odcinaniu progowym i redukcji liczby poziomów kwantowania. Obie operacje łatwo realizowane są przez mikrokontrolery.



Rys.1. Obrazy sytuacji drogowych poddawane kompresji

Fig.1. Road traffic test images

Wykonano usunięcie współczynników d_i dla pasm o największych rozdzielczościach. Współczynniki te reprezentują szczegóły obrazu. Obiekty na zrekonstruowanych obrazach posiadały zamazane brzegi. Dyskwalifikuje to tę metodę w zastosowaniach rejestrowania zdarzeń drogowych.

Następnie usunięto współczynniki, których wartości nie mieściły się w zakresie odchylenia standardowego od wartości średniej. Podobnie jak w poprzedniej próbie zamazane zostały kontury obiektów, a dodatkowo odtworzone obiekty były zniekształcone.

Przeprowadzono próby z kombinacjami usuwania szczegółów w różnych subpasmach i zawężania zakresu wartości współczynników, jednak żadna z prób nie przyniosła wyróżniającego ją efektów redukcji entropii zapisu.

Zbadano różne sposoby kwantowania wartości współczynników. Niewielka liczba dopuszczalnych wartości uzyskanych przez procedury próbkowania wymaga mniejszej liczby bitów dla rozróżnienia. Zmniejszanie liczby poziomów kwantowania prowadzi do powstawania łata na zrekonstruowanym obrazie. Zachowane zostają kontury obiektów.

Każda redukcja lub modyfikacja wartości współczynników prowadzi do zniekształceń w rekonstruowanych obrazach. Dla ułatwienia przeprowadzenia porównań przyjęto jako odniesienie poziom zniekształceń wyrażany miarą PSNR równy 30dB. Tablica 2. podsumowuje wyniki kompresowania.

Tablica 2

Ważona entropia zapisu współczynników transformat obrazów zdarzeń drogowych

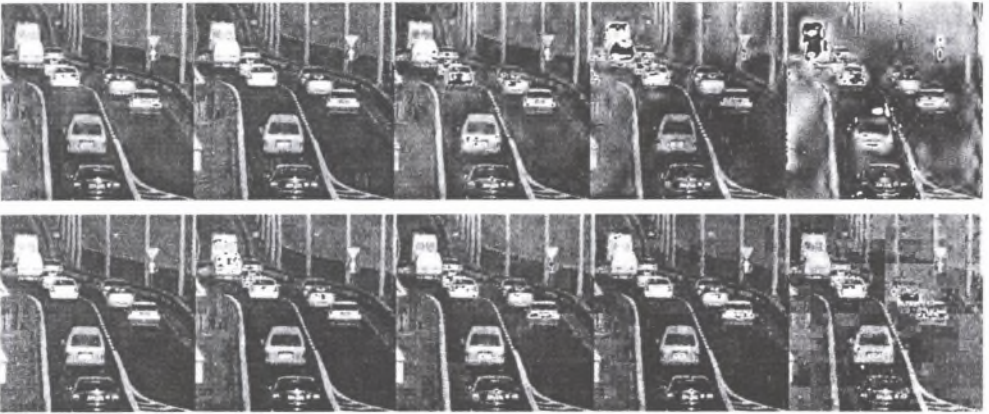
	Sytuacja drogowa	Rekonstrukcja bezstratna						PSNR = 30dB					
		Haar		D(2,2)		D(4,4)		haar		D(2,2)		D(4,4)	
		Hilbert	2-wym	Hilbert	2-wym	Hilbert	2-wym	Hilbert	2-wym	Hilbert	2-wym	Hilbert	2-wym
1.	ulica1 ^[5]	5,34	5,21	5,30	4,71	5,29	4,54	1,82	1,85	1,66	0,77	1,75	0,96
2.	ulica2 ^[5]	5,53	5,47	5,51	5,25	5,51	5,20	2,13	2,06	2,19	1,49	2,16	1,48
3.	ulica3 ^[5]	5,07	4,88	4,90	3,82	4,90	3,82	1,97	1,59	1,87	1,41	1,44	0,68
4.	aut1	5,20	5,11	5,12	4,80	5,11	4,79	1,66	1,57	1,77	1,11	1,72	1,09
5.	aut2	5,00	4,88	4,94	4,51	4,93	4,48	1,47	1,35	1,60	0,67	1,38	0,80
6.	aut3	5,26	5,12	5,18	4,69	5,16	4,64	1,86	1,80	1,83	1,00	1,77	0,96
7.	zjazd1	5,30	5,16	5,34	4,87	5,34	4,84	1,93	1,80	1,77	1,17	1,78	1,12
8.	zjazd2	5,48	5,35	5,48	5,00	5,47	4,96	1,93	2,00	1,93	1,12	1,92	1,23
9.	zjazd3	5,16	5,05	5,19	4,76	5,19	4,72	1,63	1,72	1,60	1,05	1,60	1,03
	wartości średnie	5,26	5,14	5,22	4,71	5,21	4,67	1,82	1,75	1,8	1,09	1,72	1,04

W przypadku bezstratnej rekonstrukcji dwuwymiarowe transformaty mają średnio o 2-12% mniejszą entropię zapisu od poprowadzonych po krzywej Hilberta. Wyniki bardzo zależą od zawartości obrazu, transformata D(4,4) daje najkrótszy zapis.

Rekonstrukcje o gorszej jakości, chociaż oceniane przez użytkownika jako bardzo dobre, mają entropię zapisu transformaty około 4-krotnie mniejszą i przewaga dwuwymiarowej transformaty wzrasta w podobnym stopniu, szczególnie dla obrazów o niewielkiej liczbie obiektów.

Obliczanie transformaty po krzywej wypełniającej przestrzeń daje dwukrotnie mniejsze rozrzuty wartości w porównaniu z dwuwymiarową transformacją. Doświadczenia z redukcją liczby poziomów kwantowania wskazują na kolejną zaletę wektorowej metody liczenia transformaty. Gdy liczba poziomów kwantowania współczynników maleje, odtwarzane obrazy na podstawie transformaty dwuwymiarowej zawierają rozmyte obiekty, a w przypadku jednowymiarowej transformaty obiekty są aproksymowane prostokątami.

Kolejne klatki na rysunku 2 przedstawiają zrekonstruowane obrazy o malejącej jakości od 24 dB do 12 dB co 3dB PSNR. Górny wiersz zawiera obrazy zrekonstruowane na podstawie transformaty dwuwymiarowej.



Rys.2. Pogorszenie jakości rekonstruowanych obrazów w funkcji zmniejszającej się liczby poziomów kwantowania wartości współczynników transformaty

Fig.2. Deterioration of image quality with the lowering of the number of quantization levels of transform coefficients

5. WYMAGANIA OBLICZENIOWE

Sprzętowa, z użyciem mikrokontrolerów, implementacja obliczeń transformat wymaga optymalizacji sposobu wyliczania i liczby operacji. W miejsce mnożeń, dzieleni i liczb zmiennoprzecinkowych należy zastosować przesunięcia, sumowania i liczby całkowite.

W przypadku kwadratowego obrazu o rozmiarze $2^n \cdot 2^n$ pikseli dla każdego subpasma jednowymiarowej transformaty liczba współczynników maleje dwukrotnie. Liczba operacji wymaganych do wyliczenia wszystkich współczynników wynosi:

$$N_v = p \sum_{i=1}^{2n} 2^i = p \cdot 2^{2n} \left(1 + \frac{1}{2} + \dots + \frac{1}{2^{2n-1}}\right) \quad (4)$$

$$N_v \approx p \cdot 2^{2n+1} \quad (5)$$

p – liczba operacji arytmetyczno-logicznych dla wyliczenia współczynnika,
 $2n$ – liczba subpasma.

W przypadku dwuwymiarowej transformaty dane obrazowe przetwarzane są wierszami, a następnie kolumnami. Uzyskuje się ciąg ćwiartek, które podlegają przeliczaniu. W sumie liczba operacji wynosi:

$$N_{2D} = p_r \sum_{i=1}^n 2^i \cdot 2^i + p_c \sum_{i=1}^n 2^i \cdot 2^i = 2 \cdot p \cdot 4^n \left(1 + \frac{1}{4} + \dots + \frac{1}{4^{n-1}}\right) \quad (6)$$

$$N_{2D} \approx \frac{4}{3} p \cdot 2^{2n+1} \quad (7)$$

p_r, p_c, p - liczba operacji arytmetyczno-logicznych dla wyliczenia współczynnika,
 n - liczba subpasm.

6. WNIOSKI

Przedstawione wyniki badań świadczą o korzystnym dla zwiększenia stopnia kompresji zastosowaniu obliczania transformaty falkowej po krzywej wypełniającej przestrzeń w porównaniu do tradycyjnej dwuwymiarowej transformaty falkowej.

Współczynniki obliczone po krzywej zachowują kontury obiektów. Obliczenie transformaty całego obrazu wymaga 1/3 mniej operacji, co jest cenną cechą w przypadku realizacji kompresji w czasie rzeczywistym za pomocą mikrokontrolera.

Można sądzić, że uzupełnienie algorytmu obliczania transformaty jednowymiarowej o kodowanie słownikowe, uzyskanych współczynników, pozwoli dodatkowo zwiększyć stopień kompresji. Wartości punktów obrazu klasy „zdarzenia drogowe” przetwarzane w trybie otoczenie po otoczeniu powinny zostać odwzorowane przez niewielką liczbę zmiennych reprezentujących skupiska pikseli o różnych cechach odpowiadających obiektom na obrazie, w większości przypadków pojazdom.

Literatura

1. Calderbank A. R., Daubechies I., Sweldens W., Yeo B. -L.: Wavelet Transforms that Map Integers to Integers. Technical Report. Department of Mathematics, Princeton University 1996.
2. Salomon R.: Data Compression. The Complete Reference. 2nd Edition. Springer-Verlag, Berlin, Heidelberg, New York 2000.
3. Skarbek W.: Metody reprezentacji obrazów cyfrowych. Akademicka Oficyna Wydawnicza PLJ, Warszawa 1993.
4. Daubechies I., Sweldens W.: Factoring wavelet transforms into lifting steps. Technical Report. Bell Laboratories, Lucent Technologies 1996.
5. Reference traffic image sequences, www.n-nagel.de.
6. Daubechies I., Sweldens W.: Factoring wavelet transforms into lifting steps. Technical Report. Bell Laboratories, Lucent Technologies 1996.
7. Watkins C.D., Sadun A., Marenka S.: Modern Image Processing. Academic Press Inc. 1993.

8. Peitgen H. -O., Jurgens H., Saupe D.: Fractals for the Classroom. Part 1. Introduction to Fractals and Chaos. Springer-Verlag, Berlin, Heidelberg, New York 1992.
9. JPEG2000 standard IEC15444.

Recenzent: Dr hab. inż. Dariusz.Badura, prof. UŚ

Abstract

An algorithm is proposed to incorporate Hilbert scan of image data and a wavelet transform, factorized into lifting steps, to process video traffic data, which is suitable for on site microcontroller implementation. Scanning an image, in space filling curve order, brings together pixels that are highly correlated. This scan - a vector of the image, represents clustered pixels, which may be more effectively transformed. Applying a 1-dimensional wavelet transform requires a smaller number of processing steps.

Presented research results show that wavelet transforms on space filling curves offer an advantage over traditional two dimensional wavelet computations. High compression ratios in the case of hilbert scan preserves object boundaries.

Vector scan attains lower dispersion of coefficient values for a given class of images and 1/3 less number of computing operations.

No coding of the results was carried out. Coding may reveal further advantages as Hilbert scan tends to cluster spacially correlated pixels and thus their transform coefficients. These clusters may constitute template code entries in some dictionary based coder.