II Konferencja Naukowa KNWS'05
"Informatyka- sztuka czy rzemiosło"
15-18 czerwca 2005, Złotniki Lubańskie

NIEOPTYMALNA TECHNIKA DEKORELACJI W CYFROWYM PRZETWARZANIU OBRAZU

Wojciech Zając

Instytut Informatyki i Elektroniki, Uniwersytet Zielonogórski
65-246 Zielona Góra, ul. Podgórna 50

e-mail: W.Zajac@iie.uz.zgora.pl

STRESZCZENIE

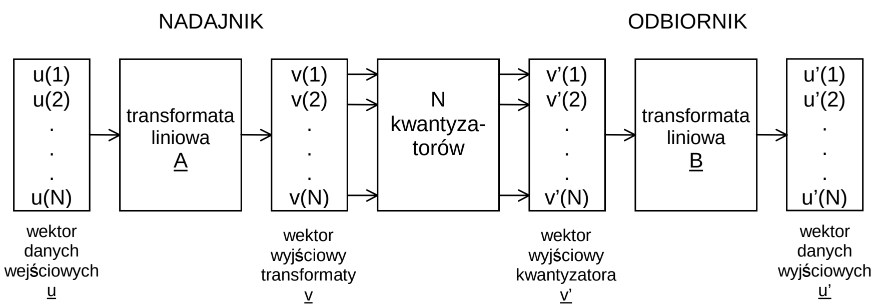
Dekorelacja danych jest jednym z krytycznych etapów przetwarzania danych cyfrowych. Wśród wielu technik można wyróżnić optymalne, wnoszące najmniejsze możliwe błędy przetwarzania, oraz nieoptymalne, które z kolei charakteryzuje korzystny stosunek kosztu numerycznego do poziomu wprowadzanych błędów. Artykuł ma na celu przybliżenie zasad realizacji operacji nieoptymalnej dekorelacji, ukazano na przykładzie jej zalety, wady oraz perspektywy rozwoju.

1. WPROWADZENIE

W technikach cyfrowego przetwarzania danych wizyjnych jednym z istotniejszych jest zagadnienie dekorelacji danych. Techniki dekorelacji wymagają znacznych nakładów numerycznych, mogą też wnosić istotne zakłócenia do sygnału. Dyskretna Transformata Kosinusowa DCT jako narzędzie dekorelacji należy do rodziny transformat nieoptymalnych z punktu widzenia minimalizacji błędu średniokwadratowego. Ich stosowanie uzasadnione jest korzystnym stosunkiem kosztu numerycznego względem poziomu wprowadzanych błędów, przy stosowaniu do założonej klasy sygnałów.

2. TRANSFORMATY DEKORELUJĄCE

Kodowanie z wykorzystaniem transformaty [1, 2] jest alternatywą do kodowania z przewidywaniem [6]. Stanowi ono szczególny przypadek kwantyzacji blokowej, operującej na wektorach o rozmiarze N próbek. W technice kodowania z przewidywaniem kolejne dane wejściowe poddawane są dekorelacji za pomocą nieliniowego filtru rekurencyjnego. Technika kodowania za pomocą transformaty jednowymiarowej kwantyzuje i dekoreluje wszystkie próbki wektora danych jednocześnie (rys. 1.). Transformacja dwuwymiarowa składa się z dwóch operacji transformacji jednowymiarowej, przeprowadzonych w kierunkach prostopadłych.



Rys. 1. Schemat jednowymiarowego kodowania za pomocą transformaty

W cyfrowym przetwarzaniu obrazu używany jest cały szereg transformat: Karhunen'a-Loeve'go (ang. KLT - Karhunen-Loeve Transform), dyskretna transformata kosinusowa (ang. DCT - Discrete Cosine Transform), transformata Walsh'a - Hadamard'a (ang. WHT - Walsh-Hadamard Transform), transformata Haar'a (ang. HT - Haar Transform), dyskretna transformata Fourier'a (ang. DFT - Discrete Fourier Transform), dyskretna transformata sinusowa (ang. DST - Discrete Sine Transform).

Optymalna transformatą dekorelującą, która minimalizuje zakłócenia średniokwadratowe w sygnale odtworzonym jest transformata KLT [3]. W praktyce zamiast niej stosuje się nieoptymalne, lecz szybkie transformaty unitarne. Wykazano [4], że dla losowych sekwencji danych istnieje wiele transformat jednostkowych o zdolnościach upakowywania energii zbliżonych do transformaty KLT, np. transformata kosinusowa, Fouriera i sinusowa. Ich sprawność jest równa sprawności transformaty KLT przy rozmiarze bloku N dążącym do nieskończoności.

Dla stacjonarnego procesu Markowa pierwszego rzędu macierz transformaty DCT, o elementach DCTi,j zdefiniowanych jak w równaniu (1), osiąga sprawność (wydajność dekorelacji w stosunku do kosztu numerycznego) zbliżoną do sprawności transformaty KLT, nawet gdy rozmiar bloku N jest niewielki [4].

$$DCT_{i,j} = \begin{cases} \dfrac{1}{\sqrt{N}} & dla \quad i=1,\ 1 \le j \le N \\ \sqrt{\dfrac{2}{N}} \cos\left[(2j-1)(i-1)\dfrac{\pi}{2N} \right] & dla \quad 1 < i \le N,\ 1 \le j \le N. \end{cases} \tag{1}$$

Jak wspomniano, największą sprawność ma transformata KLT, jednak jej koszt numeryczny jest największy ($2N^3$ operacji mnożenia i $2N^3$ operacji dodawania) [5]. Dodatkowymi jej mankamentami są: fakt, że nie istnieje szybki algorytm KLT oraz fakt, że do operacji potrzebna jest znajomość modelu kowariancji sygnału źródłowego, co stanowi złożone zadanie numeryczne.

Transformata kosinusowa należy do grupy transformat deterministycznych (opiera się na założeniach co do kowariancji sygnału) i w tej grupie jest najbardziej wydajną [4]. Jej koszt numeryczny to $2N^2 \log_2 N$ operacji mnożenia i $2N^2 \log_2 N$ operacji dodawania. Dla sygnałów o dużej korelacji danych sprawność DCT jest zbliżona do sprawności KLT. Z tego powodu transformata DCT jest powszechnie stosowana jako metoda dekorelacji danych w standardowych systemach przetwarzania obrazu stałego i ruchomego (JPEG, MPEG1, MPEG2, H.261, H.263).

3. OBLICZENIA Z WYKORZYSTANIEM TRANSFORMATY DCT

Przetwarzanie za pomocą transformat wykorzystuje fakt, że dowolny sygnał s może być przedstawiony za pomocą liniowej kombinacji pewnych funkcji elementarnych, zwanych funkcjami bazowymi f_i - równanie (2) [3].

$$s = \sum_i C_i f_i \qquad \text{gdzie } C_i \text{ jest stałą.} \tag{2}$$

Dokonanie operacji transformacji danych wymaga podzielenia ich na macierze o rozmiarze zgodnym z wielkością macierzy bazowej transformaty, przy czym jej wielkość równa np. 8×8 implikuje podział danych na wektory 1×8 punktów (w przypadku transformacji jednowymiarowej) lub 8×8 (w przypadku transformacji dwuwymiarowej).

Przeprowadzenie transformaty DCT na macierzy danych oznacza dokonanie splotu bloków macierzy z funkcjami podstawowymi transformaty. Innymi słowy, „ważony” jest udział składowych sygnału, analizowanych przez zespół filtrów czułych na częstotliwość odpowiadającą częstotliwości danej funkcji składowej. Ponieważ analiza częstotliwościowa, realizowana przez transformatę operuje na wektorze danych wejściowych, dając wektor danych wyjściowych, w celu dokonania dwuwymiarowego splotu macierzy m×m operację realizuje się dwukrotnie, w kierunku pionowym i poziomym. Dla wygody obliczeń, zmianę kierunku transformacji osiąga się najczęściej na drodze transpozycji macierzy bazowej DCT.

Większość współczesnych standardów, wykorzystujących transformatę kosinusową jako stopień dekorelacji danych (JPEG, MPEG1, MPEG2, H.261, H.263), stosuje transformatę o rozmiarze macierzy bazowej równym 8×8 punktów. Powodem jest relatywnie duża wydajność dekorelacji połączona z akceptowalnym kosztem numerycznym operacji. Prawdopodobnie w przyszłości wzrost dostępności procesorów o dużej mocy obliczeniowej przyczyni się do wprowadzenia nowych standardów i zastosowania większych rozmiarów bloku, np. 16×16 punktów.

Obraz, który ma zostać poddany dekorelacji przez transformatę, dzielony jest na kwadratowe bloki o rozmiarze 8×8 punktów. Bloki te są przetwarzane odrębnie. Dwuwymiarowa transformacja pojedynczego bloku X polega na dokonaniu operacji opisanej równaniem (3), gdzie DCT jest macierzą bazową transformaty (5), a jej elementy $DCT_{i,j}$ opisane są równaniem (1). Odwrotną operację transformacji przedstawia równanie (4).

$$Y = DCT * X * DCT^{T} \tag{3}$$

$$X' = DCT^{T} * Y * DCT \tag{4}$$

gdzie: X - macierz danych sygnału wejściowego (sygnał przestrzenny),

Y - macierz danych sygnału po transformacie (sygnał częstotliwościowy),

DCT - macierz bazowa transformaty,

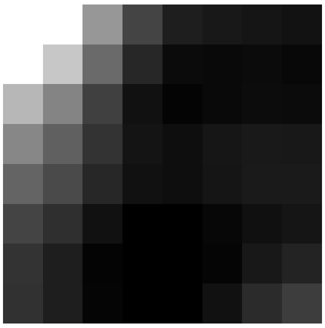
DCT^{T} - transponowana macierz bazowa transformaty,

$X‘$ - macierz odtworzonych danych wyjściowych (sygnał przestrzenny).

Macierz bazowa transformaty kosinusowej DCT, wyznaczona z (1) przy N=8 jest równa

$$DCT_{8\times8} = \begin{bmatrix} 0,3536 & 0,3536 & 0,3536 & 0,3536 & 0,3536 & 0,3536 & 0,3536 & 0,3536 \\ 0,4904 & 0,4157 & 0,2778 & 0,0975 & -0,0975 & -0,2778 & -0,4157 & -0,4904 \\ 0,4619 & 0,1913 & -0,1913 & -0,4619 & -0,4619 & -0,1913 & 0,1913 & 0,4619 \\ 0,4517 & -0,0975 & -0,4904 & -0,2778 & 0,2778 & 0,4904 & 0,0975 & -0,4157 \\ 0,3536 & -0,3536 & -0,3536 & 0,3536 & 0,3536 & -0,3536 & -0,3536 & 0,3536 \\ 0,2778 & -0,4904 & 0,0975 & 0,4157 & -0,4157 & -0,0975 & 0,4904 & -0,2775 \\ 0,1913 & -0,4619 & 0,4619 & -0,1913 & -0,1913 & 0,4619 & -0,4619 & 0,1913 \\ 0,0975 & -0,2778 & 0,4157 & -0,4904 & 0,4904 & -0,4157 & 0,2778 & -0,0975 \end{bmatrix} \tag{5}$$

Po operacji dwuwymiarowej transformacji obrazu otrzymuje się zbiór współczynników DCT, równoliczny ze zbiorem punktów obrazu. Ponieważ obrazy naturalne tj. otrzymane przez sfotografowanie rzeczywistych obiektów charakteryzuje względnie niewielka dynamika jasności, analiza częstotliwościowa przez transformatę daje znaczące współczynniki w zakresie niskich częstotliwości, zaś pozostałe są zbliżone lub równe zeru. Odpowiada to grupowaniu niezerowych współczynników w porządku malejącym w lewym górnym rogu macierzy. Przykładowo fragmentowi obrazu z rys. 2. odpowiada macierz Y zakodowanego bloku 8×8, przedstawiona równaniem (6).



Rys. 2. Przykładowy fragment obrazu, będący źródłem danych X do macierzy Y z równania (6)

$$Y_{8\times 8}=\begin{bmatrix} -36 & 12 & 9 & 2 & 0 & 0 & 0 & 0 \\ 9 & 11 & 3 & 1 & 0 & 0 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6}$$

4. ARTEFAKTY TRANSFORMATY DCT

Zniekształcenia, do jakich prowadzi zgrubne próbkowanie sprawiają, że na etapie odtwarzania sygnału do postaci analogowej wprowadzane są zakłócenia harmoniczne wyższej częstotliwości, czasem o dość istotnym znaczeniu. W sygnałach wizyjnych przyjmują one postać artefaktów w postaci zakłócenia treści bloków sąsiednich względem silnie kontrastowych elementów obrazu - ostrych krawędzi i drobnych detali o większej dynamice. Zakłócenia te mają formę tzw. wzorców błędów transformaty, charakterystycznych dla równań funkcji bazowych transformaty. Rys. 3. przedstawia fragment obrazu testowego CAMERAMAN, z widocznymi artefaktami pochodzącymi od próbkowania z małą rozdzielczością częstotliwościową.

Rys. 3. Fragment obrazu testowego CAMERAMAN; widoczne artefakty pochodzące od harmonicznych, powstałe na skutek niedokładnej analizy częstotliwościowej

5. PODSUMOWANIE

Kodowanie z wykorzystaniem transformat nie znajduje alternatywy. Prowadzi się badania nad wykorzystaniem innych technik transformacji, na szczególną uwagę zasługuje wykorzystanie transformacji DWT (ang. Discrete Wavelet Transform, Dyskretna Transformata Falkowa). Daje ona obiecujące efekty (m.in. zwiększenie odporności sygnału na zakłócenia transmisyjne), jednak posiada też wady – główna to znaczny koszt numeryczny. Istniejące standardy wciąż znajdują na tyle szerokie zastosowanie, że efektywne i powszechne wykorzystanie DWT należy jeszcze do przyszłości.

LITERATURA

[1] N. Ahmed: *Discrete Cosine Transform*. IEEE Transactions on Computers, 1974

[2] R. Clarke: *Transform Image Coding*. Academic Press, London 1985

[3] A. K. Jain, et al.: *Image data compression*. In: M. P. Ekstrom (ed.) *Digital Image Processing Techniques*. Academic Press, London 1981

[4] A. K. Jain: *A Sinusoidal Family of Unitary Transforms*. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*. no. 1, 1979

[5] W. D. Ray, R. M. Driver: *Further Decomposition of the Karhunen-Loeve Series Representation of a Stationary Random Process*. *IEEE Transactions on Information Theory*, vol. 16, no. 11, 1970

[6] W. Zając: *Cyfrowe przetwarzane i transmisji obrazów z wysokim współczynnikiem kompresji*, Zeszyty Naukowe WST w Legnicy, Legnica 1999