

ROZDZIAŁ 11. KODOWANIE TRANSFORMACYJNE

Metoda kompresji stratnej zwana kodowaniem transformacyjnym (ang. transform encoding - TE) jest najpopularniejszym sposobem kompresji danych, szczególnie obrazowych. Potencjalne możliwości tej metody, zarówno ze względu na efektywność kompresji jak i możliwości aplikacyjne, są znaczące. Znalazło to wyraz w pracach nad standardami JPEG, MPEG-1, MPEG-2, gdzie algorytmy stratnej kompresji obrazów zostały skonstruowane właśnie w oparciu o koncepcję kodowania transformacyjnego. Zastosowanie liniowej transformaty do dekorelacji danych we wstępnej fazie dekompozycji, a następnie zoptymalizowanych schematów kwantyzacji i kodowania danych ma przede wszystkim duże walory praktyczne ze względu na niewielkie koszty obliczeniowe i łatwość realizacji koderów, zarówno programowych jak i sprzętowych. Przedstawione poniżej rozważania dotyczą zarówno podstaw teoretycznych technik kodowania transformacyjnego, jak też praktycznych rozwiązań stosowanych w bardzo szerokiej gamie zastosowań, dotyczących przede wszystkim kompresji różnego typu danych obrazowych.

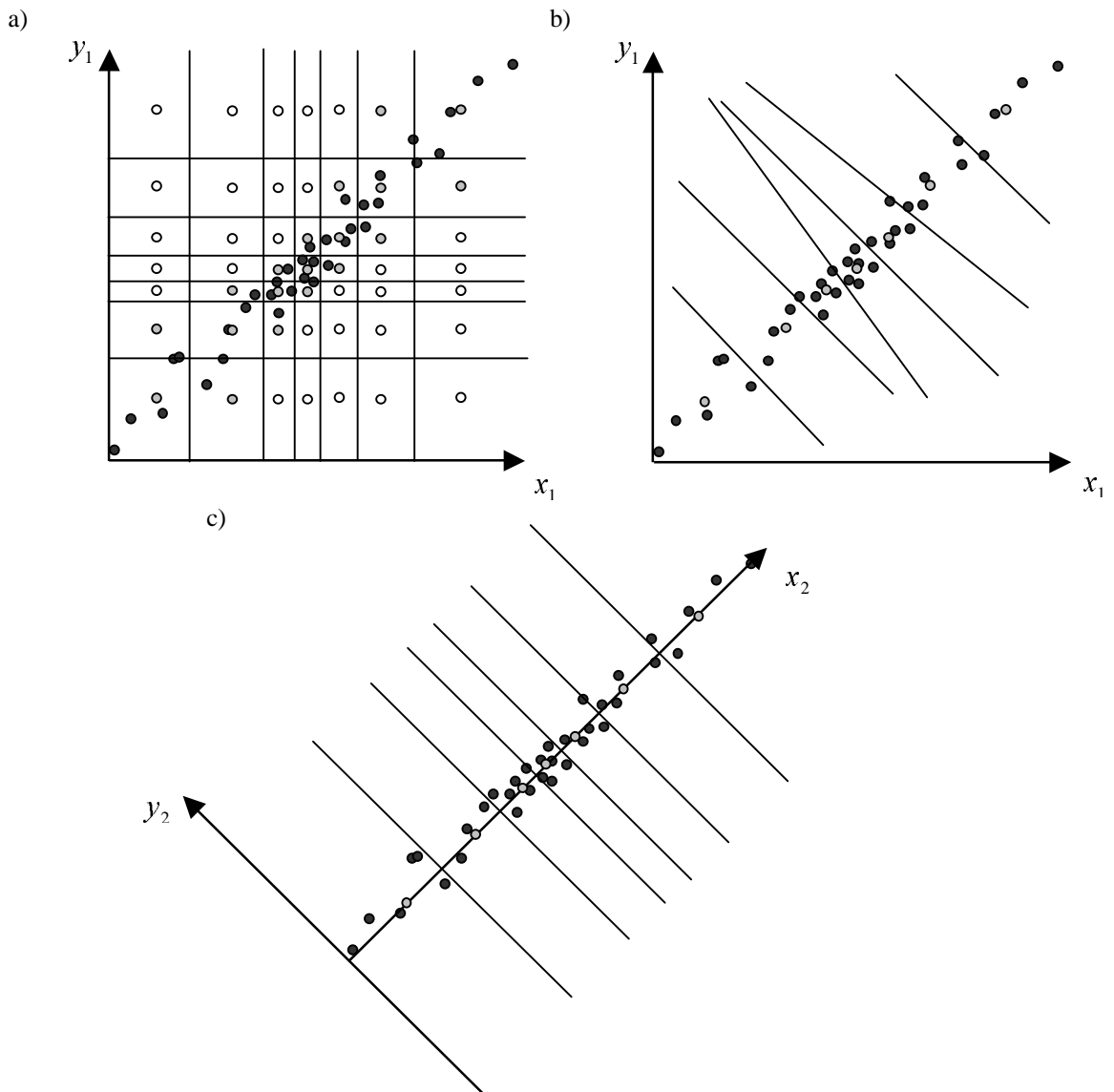
11.1. Wprowadzenie

Zasadniczym celem transformacyjnego kodowania jest zwiększenie skuteczności kwantyzacji skalarnej poprzez lepsze usytuowanie przestrzeni, na której działa operator kwantyzacji. Pokazuje to rys.11.1. Na pierwszym rysunku – rys. 11a) - widać nieskuteczność skalarnej kwantyzacji w oryginalnej, dwuwymiarowej przestrzeni współrzędnych (x_1, y_1) . Punkty danych nie są rozłożone równomiernie w przestrzeni, lecz skupione w pewnym obszarze sugerując w przybliżeniu silną liniową zależność dwu współrzędnych (wymiarów wektorów danych). W skalarnej, niezależnej kwantyzacji po obu współrzędnych nie można uwzględnić tej zależności, a co za tym idzie zredukować w procesie kwantyzacji oczywiście nadmiarowości oryginalnej reprezentacji danych. Nierównomierna skalarne kwantyzacja prowadzi do powstania całego zbioru niepotrzebnych punktów rekonstrukcji - centroidów przedziałów kwantyzacji, do których nie wpada żadna dana. Widać na rysunku, że wśród centrów rekonstrukcji pokrywających całą przestrzeń danych znajduje się liczny zbiór punktów oznaczonych kółkiem z białym wypełnieniem, zupełnie nieprzydatny w procesie kwantyzacji przykładowego zbioru danych..

Kwantyzacja wektorowa tego samego zbioru danych na kolejnym rysunku – rys. 11b) - pozwala skonstruować wektory kodu, które znacznie lepiej, bez nadmiarowości, opisują zbiór danych z minimalnym błędem kwantyzacji dla założonej liczby tychże wektorów. Jednak trudności implementacyjne wektorowej kwantyzacji, występujące przy większej wymiarowości wektorów (o których wspomniano w poprzednim rozdziale), zmuszają do poszukiwania innych rozwiązań o zbliżonej skuteczności, przy jednocześnie znacznie zmniejszonej złożoności obliczeniowej metody. Jedną z możliwości jest zastąpienie VQ wstępną transformacją danych wraz z kwantyzacją skalarne.

Na kolejnym rysunku – rys. 11c) – schematycznie przedstawiono taki sposób. Obrót układu współrzędnych powoduje, że wariancja, która jest powszechnie uznanym wskaźnikiem ilości informacji niesionej przez daną zmienną, wzdłuż jednej składowej wyraźnie maleje, wobec czego następuje koncentracja znaczącej części informacji na drugiej składowej. Tak więc mając na początku $\sigma_{x_1}^2 \cong \sigma_{y_1}^2$ po obrocie uzyskujemy $\sigma_{x_2}^2 \gg \sigma_{y_2}^2$, przy czym $\sigma_{x_1}^2 + \sigma_{y_1}^2 = \sigma_{x_2}^2 + \sigma_{y_2}^2$ (zachowana zostaje energia w nowej przestrzeni (x_2, y_2)),

utworzonej przez obrót układu współrzędnych o 45°). Kwantyzacja może być teraz przeprowadzona jedynie wzdłuż osi x_2 bez większych strat jakości całego procesu, wynikających z pominięcia składowej y_2 . Nastąpiła redukcja wymiarowości przestrzeni, co zwiększyło efektywność procesu kwantyzacji skalarnej, a także potencjalnie całego schematu kompresji. Innymi słowy została silnie zredukowana nadmiarowość danych występująca w oryginalnym układzie współrzędnych poprzez redukcję wymiarowości przestrzeni. W takim rozwiązaniu pozostaje więc wybór optymalnego obrotu, który zminimalizuje poziom korelacji wzajemnej danych opisanych poprzez nowy układ współrzędnych. Jeśli tylko udałoby się znaleźć niezbyt złożone narzędzie, które pozwoli wyznaczyć optymalny obrót poprzez realizację algorytmu o rozsądnych kosztach obliczeniowo-sprzętowych, to wówczas można by zastąpić trudną realizacyjnie technikę wektorowej kwantyzacji równie efektywną techniką kompresji, opartą o transformację układu współrzędnych oraz kwantyzację skalarną w zredukowanej przestrzeni, o stosunkowo niewielkich kosztach aplikacyjnych.



Rys.11.1. Uzasadnienie operacji dekompozycji danych przy pomocy liniowej transformaty w zastosowaniu do kompresji danych. Zamiast nieefektywnej kwantyzacji skalarnej (duża liczba niepotrzebnych poziomów rekonstrukcji) - a) - i trudno implementowalnej efektywnej kwantyzacji wektorowej - b), można obrócić układ współrzędnych i przy pomocy metody

skalarnej uzyskać zbliżoną efektywność kwantyzacji - c); oznaczenia: kółka czarne - punkty danych, kółka z szarym wypełnieniem - reprezentanci regionów decyzyjnych (wartości rekonstruowane) z niezerowym zbiorem przyporządkowanych danych, kółka z białym wypełnieniem - punkty rekonstrukcji w tym przypadku nadmiarowe (nie przybliżające żadnego wektora danych).

Liniowa transformata jest mechanizmem ustawienia układu współrzędnych w n -wymiarowej przestrzeni wejściowych wektorów danych, w pozycji możliwie silnie redukującej poziom korelacji danych. Towarzyszy temu zmniejszenie realnego wymiaru przestrzeni, w której zdecydowana większość energii sygnału skoncentrowana jest wokół kilku najbardziej znaczących współrzędnych. W wyniku transformacji wektora danych do nowej przestrzeni okazuje się więc, że tylko kilka współczynników będących rzutami na najbardziej istotne osie współrzędnych ma znaczące wartości stanowiąc oszczędną reprezentację informacji zawartej w oryginale.

Transformata jest realizowana przy pomocy zbioru funkcji ortogonalnych stanowiących jądro przekształcenia. To one ustalają kierunki rzutowania danych, czyli osie nowego układu współrzędnych. Wybór jądra pozwalającego uzyskać najsilniejszą redukcję nadmiarowości stanowi kluczowe zagadnienie w optymalizacji kodowania transformacyjnego. Ważnym także okazuje się dobór schematu kwantyzacji, najczęściej skalarnej oraz metody odwracalnego kodowania współczynników transformaty.

Zastosowanie optymalnej transformacji unitarnej sygnału oryginalnego zwanej transformatą Karhunen-Loeve'go (KLT), pozwala uzyskać w pełni zdekorelowany zbiór współczynników transformaty (czemu odpowiada najefektywniejsze ustawienie osi układu współrzędnych). Nie jest on niestety zbiorem danych statystycznie niezależnych, więc zastosowanie VQ do kwantyzacji wartości współczynników transformat unitarnych (nawet KLT) może spowodować silniejsze 'uniezależnienie' danych. Próby z wektorową kwantyzacją współczynników transformat, której postać w wyniku redukcji nadmiarowości danych poprzez transformację może być prostsza, nie przyniosły wyraźnej poprawy w stosunku do rozwiązań skalarnych, a ze względu na znacznie większą złożoność obliczeniową algorytmów VQ nie znalazły one szerszego zastosowania.

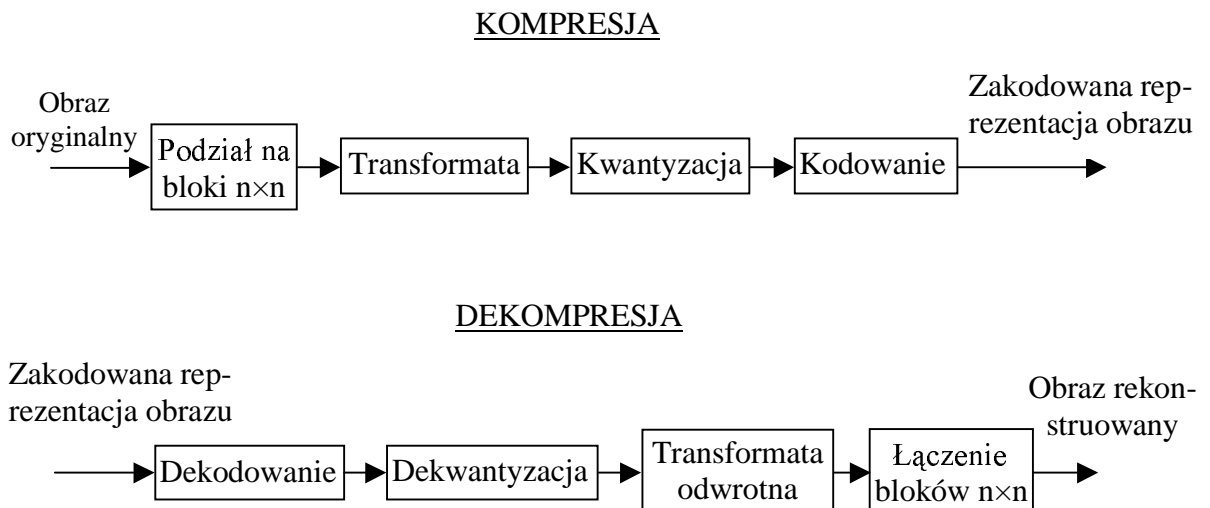
11.2. Schemat podstawowy kodowania transformacyjnego

Metoda ta znajduje najczęściej zastosowanie w kompresji obrazów, gdzie transformowane wektory danych stanowią kolejne bloki podziału obrazu, jakkolwiek może być też wykorzystywana w kompresji mowy i innych zbiorów danych pomiarowych. Ogólny schemat kodowania transformacyjnego zawiera najczęściej podział obrazu o rozmiarach $N \times N$ (dla wygody posłużymy się kwadratową matrycą obrazu) na mniejsze bloki $n \times n$ i wykonanie transformacji unitarnej każdego z tych bloków. Transformacja unitarna jest odwracalną liniową transformatą, której jądro opisuje zbiór zupełny ortonormalnych dyskretnych funkcji bazowych. Zajmują one szczególną pozycję wśród transformat dyskretnych ponieważ są łatwo odwracalne oraz zachowują energię sygnału. Celem transformacji jest usunięcie korelacji danych występującej w obrazie oryginalnym. Dekorelację osiąga się zasadniczo przez fakt, iż po transformacji energia sygnału (2-D) zawarta jest jedynie w wartościach niewielkiej liczby współczynników. Umożliwia to efektywną kompresję wskutek usunięcia wielu bliskich zeru wartości pozostałych współczynników w procesie kwantyzacji, bez większego wpływu na jakość rekonstruowanego obrazu. Odpowiedni dobór współczynników kwantyzacji (uwzględniający funkcję czułości kontrastu) pozwala często uzyskać przy

pomocy technik transformacyjnych wizualnie bezstratną kompresję. Podstawowy schemat blokowy metod transformacyjnego kodowania obrazów przedstawiony jest na rys.11.2.

2-D (dwuwymiarowe) transformacje stosowane do kompresji obrazów są rozłączne, tzn. jądro transformaty może ulec dekompozycji na dwa 1-D (jednowymiarowe) jądra wyszczególniające rozłączne horyzontalne i wertykalne operacje. Stąd też 2-D transformacja na bloku $n \times n$ pikseli może być wykonana w dwu krokach, 1-D n -punktowej transformacji wzdłuż każdego wiersza w bloku, a następnie 1-D n -punktowej transformacji wzdłuż każdej kolumny należącej do danego bloku.

Transformacja n -punktowa wektora danych może być interpretowana w dwojaki sposób. Albo jako obrót n -wymiarowego układu współrzędnych, definiowany przez wartości danych, transformujący wektor oryginalny w równoważny wektor o lepszej podatności na kompresję, albo jako dekompozycja sygnału oryginalnego przy pomocy zbioru n ortonormalnych funkcji bazowych. Współczynniki transformaty kształtowane przez wartości danych są wagami poszczególnych funkcji bazowych, których liniowa kombinacja pozwala odtworzyć wektor danych wejściowych.



Rys.11.2. Blokowy schemat stratnej kompresji i dekompresji obrazów metodą transformacyjnego kodowania.

11.3. Transformaty wykorzystywane w kompresji

Do najbardziej pożądanых cech stosowanych w algorytmach kompresji transformat ortogonalnych należą:

- 1) całkowita dekorelacja danych oryginalnych, co oznacza upakowanie możliwie największej ilości energii w najmniejszej liczbie współczynników, najlepiej skupionych w jednym miejscu przestrzeni transformaty;
- 2) niezależne od przetwarzanego sygnału efektywne funkcje bazowe, których zastosowanie daje optymalne (w znaczeniu wysokiej dekorelacji danych) wyniki bez względu na zmienną statystykę danych wejściowych; jest to szczególnie istotne dla silnie niestacjonarnych źródeł wektorów danych wejściowych;

3) szybka implementacja, czyli szybkie algorytmy zmniejszające liczbę operacji obliczania transformaty z $O(n^2)$ do $O(n \cdot \log n)$, co dla rozłącznych 2-D transformat zmniejsza liczbę operacji z $O(n^4)$ do $O(2n^2 \log n)$; zapis $O(x)$ oznacza złożoność liniową (proporcjonalnie do x).

Rozważmy dowolny rzeczywisty sygnał $x(t)$ w dziedzinie czasu. Jeśli przez $\{\varphi_k\}_{k=1}^{\infty}$ oznaczymy zbiór zupełny ortonormalny funkcji bazowych przekształcenia, określonych w dziedzinie sygnału, to sygnał $x(t)$ można wyrazić następująco:

$$x(t) = \sum_{k=1}^{\infty} X_k \varphi_k(t) \quad (11.1)$$

Jest to odwrotna liniowa transformacja dyskretna sygnału rzeczywistego, polegająca na reprezentacji sygnału poprzez liniową kombinację funkcji bazowych φ_k ze współczynnikami X_k . Jeśli próbkujemy sygnał w regularnych odstępach czasowych Δt w skończonym czasie uzyskując n próbek sygnału x , to wyrażenie na i -tą próbkę sygnału jest następujące:

$$x_i = \sum_{k=1}^n X_k \varphi_{i,k}, \quad i = 1, \dots, n, \quad (11.2)$$

gdzie $x_i = x(i\Delta t)$, $\varphi_{i,k} = \varphi_k(i\Delta t)$, a X_k oznacza k -ty współczynnik tego rozwinięcia (transformaty). W zapisie macierzowym równanie (11.2) upraszcza się do postaci:

$$\mathbf{x} = \boldsymbol{\varphi} \mathbf{X} = \sum_{k=1}^n X_k \boldsymbol{\varphi}_k, \quad (11.3)$$

gdzie $\mathbf{x} = \{x_1, \dots, x_n\}^T$, $\mathbf{X} = \{X_1, \dots, X_n\}^T$ i $\boldsymbol{\varphi}$ jest macierzą postaci:

$$\boldsymbol{\varphi} = \begin{bmatrix} \varphi_{1,1} & \cdots & \varphi_{1,n} \\ \vdots & & \vdots \\ \varphi_{n,1} & \cdots & \varphi_{n,n} \end{bmatrix}, \quad (11.4)$$

przy czym $\boldsymbol{\varphi}_k = (\varphi_{1,k}, \dots, \varphi_{n,k})$ oznacza k -tą kolejną kolumnę macierzy $\boldsymbol{\varphi}$, tj. kolejny wektor bazowy tego przekształcenia.

Postać tej macierzy decyduje o własnościach transformacji. Jeśli spełniony jest warunek:

$$\boldsymbol{\varphi}^{-1} \boldsymbol{\varphi} = \mathbf{I}, \quad (11.5)$$

to macierz $\boldsymbol{\varphi}$ jest nieosobliwa, a transformacja nazywana jest unitarną. W przypadku funkcji zespolonych warunek (11.5) można zapisać w postaci $(\boldsymbol{\varphi}^*)^T \boldsymbol{\varphi} = \mathbf{I}$, a dla rzeczywistych funkcji bazowych warunek ten upraszcza się do wyrażenia $\boldsymbol{\varphi}^T \boldsymbol{\varphi} = \mathbf{I}$. Widać wtedy zalety transformacji unitarnych ze względu na prostotę wyznaczenia $\boldsymbol{\varphi}^{-1} = \boldsymbol{\varphi}^T$. Transformacja dyskretna niekoniecznie jest transformacją unitarną. Wystarczy jedynie, by wiersze macierzy $\boldsymbol{\varphi}$ były liniowo niezależne.

Wektor \mathbf{X} jest transformatą sygnału \mathbf{x} , a obliczany jest poprzez wyrażenie nazywane transformacją prostą:

$$\mathbf{X} = \boldsymbol{\varphi}^{-1} \mathbf{x} \quad (11.6)$$

Wprowadźmy oznaczenie funkcji bazowych transformacji prostej przez $\boldsymbol{\psi}_k$, przy czym $\boldsymbol{\psi} = \boldsymbol{\varphi}^{-1}$, a wówczas dyskretna transformacja prosta przyjmuje postać:

$$X_i = \sum_{k=1}^n x_k \psi_{i,k}, \quad i = 1, \dots, n \quad (11.7)$$

Transformatą unitarną, pozwalającą uzyskać pełną dekokorelację danych, a więc najwyższy poziom redukcji nadmiarowości reprezentacji kompresowanego zbioru danych spośród transformat tej klasy, jest transformata Karhunen-Loeve'go.

Transformata KL

Optymalną transformacją z punktu widzenia pierwszej cechy jest dyskretna transformata Karhunen-Loeve'go – KLT. Wersja ciągła tej transformaty nazywana jest transformatą Hotellinga. Generalnie można powiedzieć, że określona liczba współczynników transformaty KL zawiera większą część całkowitej energii sygnału niż w jakiegokolwiek innej transformacie unitarnej, a średniokwadratowy błąd przybliżenia sygnału na podstawie tych współczynników jest najmniejszy. Sygnał wejściowy ulega całkowitej dekokorelacji, co automatycznie redukuje do minimum nadmiarowość oryginalnej reprezentacji. Niestety zbiór funkcji bazowych KLT zależy od statystyki strumienia danych, algorytm ich wyznaczenia jest czasochłonny, a szybkie algorytmy obliczania KLT na podstawie wyznaczonego jądra transformaty nie istnieją. Ponieważ wektory bazowe są zależne od sygnału, muszą być przekazane do dekodera jako informacja dodatkowa zmniejszając efektywność kompresji. Trudno więc zastosować ten rodzaj transformaty w rozwiązaniach praktycznych, zwłaszcza że, jak zostanie wykazane, inne transformaty są prawie tak samo skuteczne i znacznie prościej aplikowalne.

Pełna dekokorelacja danych w wyniku transformacji unitarnej możliwa jest w sposób następujący. Załóżmy wejściowy zbiór n -elementowych wektorów danych $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(k)}$ w R^n . Przykładowo, w przypadku transformacyjnego kodowania obrazów wektory są zbudowane z bloków zawierających n pikseli (kolejne kolumny obrazu). Ustalamy także zerowy wektor średni danych (wektor składający się z wartości średnich poszczególnych składowych wszystkich wektorów danych), czyli obliczamy wektor średni i odejmujemy go od kolejnych wektorów danych. Otrzymujemy

$$\sum_{i=1}^k \mathbf{x}^{(i)} = \mathbf{0}. \quad (11.8)$$

Macierz współczynników transformaty otrzymujemy poprzez wyrażenie:

$$\mathbf{X} = \boldsymbol{\Psi} \mathbf{x}, \quad (11.9)$$

gdzie wektory danych tworzą kolejne kolumny macierzy danych wejściowych \mathbf{x} o wymiarach $n \times k$:

$$\mathbf{x} = (\mathbf{x}^{(1)} \mathbf{x}^{(2)} \dots \mathbf{x}^{(k)}), \quad (11.10)$$

a transformowane wektory

$$\mathbf{X}^{(i)} = \boldsymbol{\Psi} \mathbf{x}^{(i)}. \quad (11.11)$$

stanowią kolumny macierzy współczynników \mathbf{X} . Średni wektor zbioru wektorów $\mathbf{X}^{(i)}$ jest także równy $\mathbf{0}$, a więc średnia kwadratów wartości wszystkich elementów j -tego wiersza macierzy \mathbf{X} jest równa wariancji j -tego elementu wektora współczynników transformaty. Inaczej, wariancja współczynnika j jest równa elementowi o współrzędnych (j,j) (podzielonemu przez k), leżącemu na diagonalu macierzy kowariancji współczynników transformaty

$$\mathbf{R}_{\mathbf{xx}} = \mathbf{XX}^T. \quad (11.12)$$

Pozostałe elementy macierzy $\mathbf{R}_{\mathbf{xx}}$ to kowariancje pomiędzy wektorami s i t (współzrędnymi s i t macierzy $\mathbf{R}_{\mathbf{xx}}$ o rozmiarach $n \times n$):

$$(\mathbf{R}_{\mathbf{xx}})_{s,t} = \sum_{i=1}^k \mathbf{x}_s^{(i)} \mathbf{x}_t^{(i)T} \quad (11.13)$$

Dwie współzrędnne są zdekorelowane, jeśli ich kowariancja wynosi 0. Warunek całkowitej dekorrelacji danych wejściowych w nowej przestrzeni transformaty unitarnej sprowadza się więc do diagonalnej postaci macierzy $\mathbf{R}_{\mathbf{xx}}$.

Skonstruujmy więc tak macierz transformaty prostej $\boldsymbol{\psi}$, aby otrzymać diagonalną postać $\mathbf{R}_{\mathbf{xx}}$. Macierzowe wyrażenie na $\mathbf{R}_{\mathbf{xx}}$ wygląda następująco:

$$\mathbf{R}_{\mathbf{xx}} = \mathbf{XX}^T = (\boldsymbol{\psi}\mathbf{x})(\boldsymbol{\psi}\mathbf{x}^T) = \boldsymbol{\psi}(\mathbf{xx}^T)\boldsymbol{\psi}^T. \quad (11.14)$$

Drogą do diagonalizacji macierzy kowariancji współczynników transformaty jest diagonalizacja macierzy kowariancji wejściowych wektorów danych $\mathbf{R}_{\mathbf{xx}} = \mathbf{xx}^T$, której elementy wyznaczone są analogicznie jak w (11.13):

$$(\mathbf{R}_{\mathbf{xx}})_{s,t} = \sum_{i=1}^k \mathbf{x}_s^{(i)} \mathbf{x}_t^{(i)T}. \quad (11.15)$$

Macierz $\mathbf{R}_{\mathbf{xx}}$ jest więc symetryczna, stąd jest diagonalizowana przez jej wektory własne, które na dodatek są ortogonalne, a wartości własne tej macierzy (rzeczywiste i nieujemne) stanowią kolejne elementy diagonalni nowej macierzy. Tak więc wybierając wektory własne $\mathbf{R}_{\mathbf{xx}}$ jako bazowe wektory transformaty opisanej macierzą $\boldsymbol{\psi}$ (stanowią wiersze tej macierzy) otrzymujemy:

$$\mathbf{R}_{\mathbf{xx}} = \mathbf{XX}^T = \boldsymbol{\psi}(\mathbf{xx}^T)\boldsymbol{\psi}^T = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n), \quad (11.16)$$

gdzie $\text{diag}(a_1, a_2, \dots, a_n)$ - operator tworzący macierz diagonalną o wartościach a_1, a_2, \dots, a_n na diagonalnej, a $\lambda_1, \lambda_2, \dots, \lambda_n$ - wartości własne macierzy $\mathbf{R}_{\mathbf{xx}}$.

Uzyskaliśmy więc pełną dekorrelację wektorów współczynników transformaty, a wariancja kolejnych współczynników jest równa tymże wartościom własnym macierzy kowariancji wejściowych wektorów danych. Ortonormalne wektory własne $\mathbf{R}_{\mathbf{xx}}$ stanowią bazowe wektory KLT.

Wykazaliśmy, że KLT dekorreluje wektory współczynników nowej dziedziny, co jest bardzo cenne z punktu widzenia kompresji danych (usunięta zostaje nadmiarowość korelujących wartości wejściowych wektorów danych). Okazuje się też, że dekorrelacja współzrędnnych macierzy kowariancji $\mathbf{R}_{\mathbf{xx}}$ jest zbieżna z optymalnym upakowaniem energii w kolejnych wartościach możliwie najmniejszej liczby współczynników. Jest to szczególnie istotne w zastosowaniach stratnej kompresji danych, gdzie w uproszczonym modelu kwantyzację można sprowadzić do wyboru kilku pierwszych współczynników transformaty z pominięciem pozostałych. Ważne przy tym jest, aby współczynniki były ustawione w porządku malejącym (dla KLT wartości własne, a więc skojarzone z nimi wektory własne tworzące bazę winny być więc odpowiednio uporządkowane). Obcinanie pozostałych współczynników domaga się koncentracji możliwie największej ilości energii sygnału oryginalnego w wartościach zachowywanych współczynników. Pozwoli to wierniej

zrekonstruować sygnał w procesie dekompresji (z mniejszym błędem). Sformułujmy więc istotne twierdzenie:

TWIERDZENIE 11.1. Niech \mathbf{R} będzie symetryczną, kwadratową macierzą $n \times n$ o nieujemnych wartościach, a liczby rzeczywiste i nieujemne $\{\lambda_i\}_{i=1}^n$ będą wartościami własnymi tej macierzy w porządku malejącym, tj. $\lambda_i \geq \lambda_{i+1}$ dla $i=1, \dots, n-1$. Niech $\boldsymbol{\psi}$ będzie dowolną macierzą ortonormalną $n \times n$ oraz niech $\{d_i\}_{i=1}^n$ będą elementami diagonalnej macierzy

$$\mathbf{D} = \boldsymbol{\psi} \mathbf{R} \boldsymbol{\psi}^T \quad (11.17)$$

w porządku malejącym (one są także nieujemne).

Wtedy zachodzi nierówność

$$\sum_{i=1}^m d_i \leq \sum_{i=1}^m \lambda_i \quad (11.18)$$

dla każdego $m \leq n$.

Dowód tego twierdzenia można sprowadzić do wykazania faktu, iż to właśnie KLT daje najmniejszy błąd średniokwadratowy w przypadku rekonstrukcji sygnału oryginalnego na podstawie m współczynników transformaty, czyli że średniokwadratowa energia, rozumiana jako $\sum_{i=1}^m E(X_i^2)$ równa de facto średniej wariancji m współczynników, jest dla KLT największa.

Powróćmy do ciągu n próbek wejściowych i wyrażenia (11.3) na transformatę odwrotną. Jeśli reprezentacja oryginalna danych wejściowych zawiera pewną nadmiarowość, to znaczy, że można próbować ograniczyć wymiarowość nowej przestrzeni określonej przez bazę funkcji $\boldsymbol{\varphi}^{(i)}$ (kolumny macierzy $\boldsymbol{\varphi}$) do wartości m , gdzie $m < n$ (liczba wierszy macierzy $\boldsymbol{\varphi}$ zostaje zmniejszona do m). Pozwoli to przy pomocy zredukowanego zbioru $\{X_i\}_{i=1}^m$ elementów wektora współczynników transformaty odtworzyć rzeczywistą informację zawartą w sygnale x redukując nadmiarowość reprezentacji oryginalnej. Prowadzi to do następującego przybliżenia (aproksymacji) sygnału x :

$$\hat{x} = \sum_{i=1}^m X_i \boldsymbol{\varphi}^{(i)}. \quad (11.19)$$

Błąd średniokwadratowy, jako miara tego przybliżenia, wygląda następująco (dla rzeczywistych funkcji bazowych $\boldsymbol{\varphi}_k$):

$$\varepsilon = E \left[\sum_{i=m+1}^n \boldsymbol{\varphi}^{(i)T} [\mathbf{xx}^T] \boldsymbol{\varphi}^{(i)} \right] = \sum_{i=m+1}^n \boldsymbol{\varphi}^{(i)T} E[\mathbf{xx}^T] \boldsymbol{\varphi}^{(i)}. \quad (11.20)$$

W równaniu (11.20) wyrażenie charakteryzujące dane wejściowe przy zerowej średniej wektora \mathbf{x} jest macierzą kowariancji:

$$\mathbf{R}_{\mathbf{xx}} = E[\mathbf{xx}^T] \quad (11.21)$$

Problem minimalizacji błędu ε z równania (11.20) poprzez dobór optymalnych funkcji bazowych (różniczkowanie względem $\boldsymbol{\varphi}^{(i)}$ i zerowanie) sprowadza się do rozwiązania problemu wyznaczenia wartości własnych:

$$(\mathbf{R}_{xx} - \lambda_i \mathbf{I})\boldsymbol{\varphi}^{(i)} = 0 \quad (11.22)$$

Prowadzi to do wektorów własnych macierzy \mathbf{R}_{xx} jako wektorów bazowych transformaty opisanej macierzą $\boldsymbol{\varphi}$, czyli do KLT. Aby błąd przybliżenia ε był najmniejszy, trzeba uszeregować wartości własne (oraz skojarzone z nimi wektory własne) w kolejności malejących wartości, bowiem:

$$\varepsilon = \sum_{i=m+1}^n \lambda_i \quad (11.23)$$

PRZYKŁAD 11.1. Rozważmy prosty przykład liczenia transformaty KL dla zbioru pięciu dwuwymiarowych wektorów danych postaci:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} -1 \\ -2 \end{bmatrix}, \mathbf{x}_5 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}. \quad (11.24)$$

Aby wyznaczyć jądro KLT wyznaczamy po pierwsze macierz kowariancji wektorów danych wejściowych. Liczymy w tym celu wektor średni $\boldsymbol{\mu}_x = \frac{1}{5} \sum_{i=1}^5 \mathbf{x}_i = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, który odejmujemy od każdego z wektorów formułując następującą postać macierzy \mathbf{x} :

$$\mathbf{x} = \begin{bmatrix} 0 & -1 & 1 & -2 & 2 \\ 2 & 1 & -1 & -2 & 0 \end{bmatrix}, \quad (11.25)$$

gdzie $\mathbf{x}^{(i)} = \mathbf{x}_i - \boldsymbol{\mu}_x$, $i = 1, \dots, 5$.

Macierzy kowariancji w takim przypadku ma postać:

$$\mathbf{R}_{xx} = \mathbf{X}\mathbf{X}^T = \begin{bmatrix} 10 & 2 \\ 2 & 10 \end{bmatrix}. \quad (11.26)$$

Wartości własne tej macierzy po diagonalizacji wynoszą $\lambda_1 = 8$ i $\lambda_2 = 12$, a odpowiadające im ortogonalne wektory własne wynoszą odpowiednio $\boldsymbol{\varphi}_1 = [1, -1]$ i $\boldsymbol{\varphi}_2 = [1, 1]$, tak aby spełnione było równanie $\mathbf{R}_{xx}\boldsymbol{\varphi} = \boldsymbol{\varphi}\boldsymbol{\lambda}$, tj.:

$$\begin{bmatrix} 10 & 2 \\ 2 & 10 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 12 \end{bmatrix}. \quad (11.27)$$

Wariancje składowych wektorów oryginalnych są równe 10, czyli energia jest rozłożona równomiernie na obie współrzędne. Wartości wariancji wektora współczynników wykazują już zróżnicowanie rozkładu energii względem współrzędnych: 8 i 12 (przy zerowej kowariancji współrzędnych). Poziom zróżnicowania jest stosunkowo niewielki ze względu na małą wartość kowariancji pomiędzy obiema współrzędnymi wektorów danych, która wynosi 2.

Wektory własne aby stanowić ortonormalną bazę KLT muszą zostać znormalizowane, a ponieważ ustawienie wartości własnych w porządku malejącym wymaga odwrócenia kolejności także odpowiednich wektorów własnych, postać wektorów bazowych KLT jest następująca: $\boldsymbol{\varphi}_1 = [\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]$ i $\boldsymbol{\varphi}_2 = [-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}]$. Postać macierzy $\boldsymbol{\varphi}$, której kolumny stanowią wektory bazowe, jest następująca:

$$\boldsymbol{\varphi} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad (11.28)$$

co daje następującą postać transformaty macierzy \mathbf{x} :

$$\mathbf{X} = \boldsymbol{\varphi} \mathbf{x} = \begin{bmatrix} \frac{2}{\sqrt{2}} & 0 & 0 & -\frac{4}{\sqrt{2}} & \frac{2}{\sqrt{2}} \\ -\frac{2}{\sqrt{2}} & -\frac{2}{\sqrt{2}} & \frac{2}{\sqrt{2}} & 0 & \frac{2}{\sqrt{2}} \end{bmatrix} \quad (11.29)$$

Tak więc wyznaczenie transformaty KL obrazu (zbioru wektorów danych) lub jego części składa się z kilku etapów, a mianowicie:

- formowanie zbioru wektorów $n \times 1$ z kolejnych kolumn lub wierszy obrazu (bloku) lub kolejnych wektorów sygnału wejściowego,
- ustalenie zerowej wartości średniej wektorów,
- wyznaczenie macierzy kowariancji $n \times n$ na podstawie zbioru wektorów,
- wyznaczenie wartości własnych i znormalizowanych wektorów własnych macierzy kowariancji,
- transformata wektorów danych przy pomocy wektorów własnych jako funkcji bazowych przekształcenia (tj. jądra).

Zależność od sygnału funkcji bazowych przekształcenia KLT i związane z tym problemy aplikacyjne skłaniają do wykorzystania transformat o zdeterminowanych bazach, niezależnych od sygnału wejściowego. Bazy te nie są niestety zbiorem wektorów własnych macierzy kowariancji sygnału, ale przynajmniej niektóre z nich z dużą skutecznością przybliżają optymalne bazy KLT przy znacznych oszczędnościach kosztów aplikacji. Dominują tutaj bazy wykorzystujące funkcje sinusoidalne.

Transformata DC

Dyskretna kosinusowa transformata (DCT) w zastosowaniu do typowych obrazów o dużej korelacji wartości pikseli daje efektywność zbliżoną do KLT. Bazę tego przekształcenia stanowi zbiór ortonormalny funkcji kosinusoidalnych (zobacz rys. 11.3), będący nieco zmodyfikowaną częścią rzeczywistej transformaty Fouriera. Wyróżnia się cztery podstawowe rodzaje funkcji bazowych transformaty DC (dla n - próbek sygnału wejściowego):

$$\text{DCT rodzaj I:} \quad \varphi_{i,k} = C_{i,k}^{\text{I}} = \sqrt{\frac{2}{n}} \left[c_i c_k \cos\left(\frac{ik\pi}{n}\right) \right], \quad i, k = 0, \dots, n$$

$$\text{DCT rodzaj II:} \quad \varphi_{i,k} = C_{i,k}^{\text{II}} = \sqrt{\frac{2}{n}} \left[c_i \cos\left(\frac{i(k + \frac{1}{2})\pi}{n}\right) \right], \quad i, k = 0, \dots, n-1$$

$$\text{DCT rodzaj III:} \quad \varphi_{i,k} = C_{i,k}^{\text{III}} = \sqrt{\frac{2}{n}} \left[c_k \cos\left(\frac{k(i + \frac{1}{2})\pi}{n}\right) \right], \quad i, k = 0, \dots, n-1$$

$$\text{DCT rodzaj IV: } \varphi_{i,k} = C_{i,k}^{\text{IV}} = \sqrt{\frac{2}{n}} \left[\cos \left(\left(i + \frac{1}{2} \right) \frac{\left(k + \frac{1}{2} \right) \pi}{n} \right) \right], \quad i, k = 0, \dots, n-1 \quad (11.30)$$

$$\text{gdzie } c_i = \begin{cases} \frac{1}{\sqrt{2}} & \text{dla } i = 0 \text{ lub } n \\ 1 & \text{w p. p.} \end{cases}$$

Funkcje bazowe rodzaju II i III znalazły najczęstsze zastosowanie w algorytmach kompresji metodą transformacyjnego kodowania. Funkcje bazowe DCT pokrywają się ze zbiorem funkcji bazowych KLT danych obrazowych przy tym samym modelu obrazu – źródle Markowa pierwszego rzędu. Ponadto istnieją szybkie algorytmy bezpośredniego obliczania DCT, jak też z wykorzystaniem algorytmów FFT [1],[2,rozdz.6]. Wykorzystuje się tutaj zależność

$$X_i = \sqrt{\frac{2}{n}} c_i \operatorname{Re} \left[e^{-\frac{j\pi i}{2n}} \sum_{k=0}^{2n-1} x_k e^{\frac{-j2\pi i k}{2n}} \right] = \sqrt{\frac{2}{n}} c_i \operatorname{Re} \left[e^{-\frac{j\pi i}{2n}} FFT_{2n} \right], \quad i = 0, \dots, n-1 \quad (11.31)$$

gdzie $j = \sqrt{-1}$ i $x_k = 0$ dla $k = n, n+1, \dots, 2n-1$. Algorytmy liczenia DCT optymalizują szybkość algorytmu dla konkretnego rozmiaru wektorów wykorzystując arytmetykę liczb całkowitych i tablice wartości funkcji sinusoidalnych [3].

Blokowa DCT

Aby policzyć dwuwymiarową transformatę obrazu o wymiarach $N \times N$, pamiętając że jest to transformata rozłączna, można skorzystać ze zmodyfikowanego równania (11.7) postaci (pamiętając, że dla bazy transformaty unitarnej o funkcjach rzeczywistych $\psi = \varphi^T$):

$$X_{i,j} = \sum_{k=1}^N \sum_{l=1}^N x_{k,l} \varphi_{i,j,k,l} = \sum_{k=1}^N \sum_{l=1}^N x_{k,l} \varphi_{i,k} \varphi_{j,l}, \quad i = 1, \dots, N, \quad j = 1, \dots, N \quad (11.32)$$

gdzie $x_{k,l}$ oznacza wartość funkcji jasności piksela o współrzędnych (k,l) .

Korzystając z równań (11.30) oraz (11.32) 2-D dyskretną transformację kosinusową bloku pikseli o rozmiarach $n \times n$ i wartościach funkcji jasności $f(x,y)$ definiujemy więc w sposób następujący:

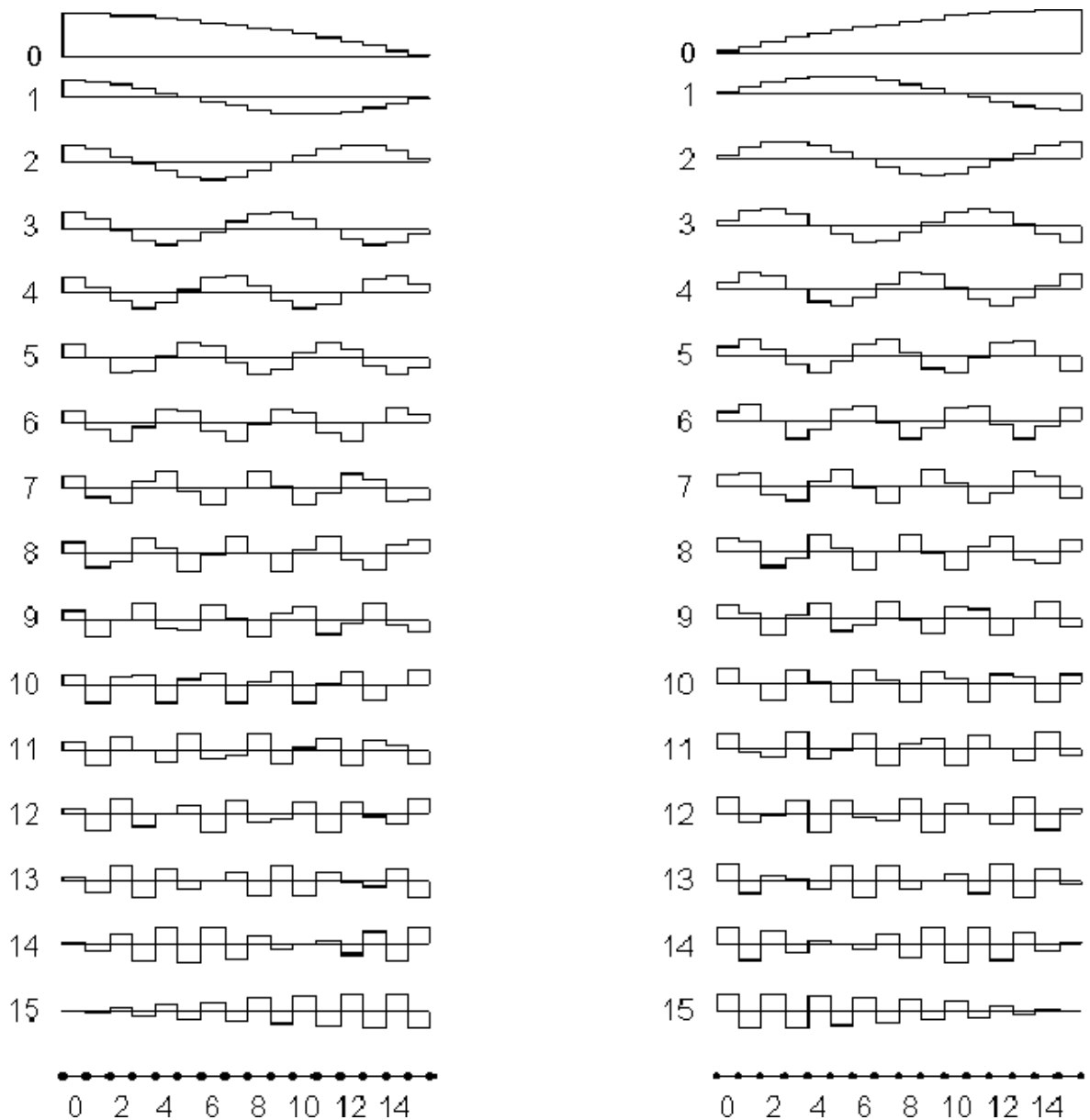
$$k(u,v) = \frac{1}{\sqrt{2n}} c_u c_v \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} f(x,y) \cos \frac{(x + \frac{1}{2})u\pi}{n} \cos \frac{(y + \frac{1}{2})v\pi}{n}, \quad (11.33)$$

gdzie $c_u, c_v = \frac{1}{\sqrt{2}}$ dla $u, v = 0$; $c_u, c_v = 1$ w p.p.

$k(u,v)$ są wartościami współczynników DCT w jej dziedzinie (u,v) . Transformacja odwrotna, będąca syntezą wejściowego zbioru danych przy pomocy liniowej kombinacji zbioru funkcji bazowych wygląda jak niżej:

$$f(x,y) = \frac{1}{\sqrt{2n}} \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} c_u c_v k(u,v) \cos \frac{(x + \frac{1}{2})u\pi}{n} \cos \frac{(y + \frac{1}{2})v\pi}{n}, \quad (11.34)$$

Takie zależności na transformatę DC są wykorzystywane w standardach kompresji JPEG i MPEG.



Rys. 11.3. Porównanie kształtu funkcji bazowych dyskretnej transformaty kosinusowej (z lewej) oraz sinusowej (z prawej) - w obu przypadkach rodzaj IV, $n=16$.

Inne transformaty

W pewnych aplikacjach znajdują zastosowanie także inne transformaty, takie jak chociażby dyskretna transformata sinusowa (DST), która w pewnych modelowych sytuacjach jest identyczna z KLT. DST jest używana w technice rekursywnego kodowania bloków (ang. recursive block coding) [4] zmniejszającej efekt artefaktów powstających na granicy bloków. Analogicznie jak w przypadku DCT definiowane są cztery rodzaje funkcji bazowych transformaty DS:

$$\text{DST rodzaj I: } \varphi_{i,k} = S_{i,k}^I = \sqrt{\frac{2}{n+1}} \left[\sin \left(\frac{ik\pi}{n+1} \right) \right], \quad i, k = 1, \dots, n$$

$$\begin{aligned}
\text{DST rodzaj II:} \quad \varphi_{i,k} = S_{i,k}^{\text{II}} &= \sqrt{\frac{2}{n}} \left[c_i \sin \left(\frac{i(k - \frac{1}{2})\pi}{n} \right) \right], & i, k = 1, \dots, n \\
\text{DST rodzaj III:} \quad \varphi_{i,k} = S_{i,k}^{\text{III}} &= \sqrt{\frac{2}{n}} \left[c_k \sin \left(\frac{k(i - \frac{1}{2})\pi}{n} \right) \right], & i, k = 1, \dots, n \\
\text{DST rodzaj IV:} \quad \varphi_{i,k} = S_{i,k}^{\text{IV}} &= \sqrt{\frac{2}{n}} \left[\sin \left((i - \frac{1}{2}) \frac{(k - \frac{1}{2})\pi}{n} \right) \right], & i, k = 1, \dots, n \quad (11.35)
\end{aligned}$$

gdzie $c_i = \begin{cases} \frac{1}{\sqrt{2}} & \text{dla } i = 0 \text{ lub } n \\ 1 & \text{w p. p.} \end{cases}$

Przykładowe funkcje bazowe transformaty sinusowej przedstawiono na rys. 11.3.

W technice rekursywnego kodowania bloków wykorzystywana jest DST rodzaju I. Podobnie jak w przypadku DCT, istnieją także szybkie algorytmy liczenia transformaty DS [2, rozdz.6]. Ponieważ jądra transformat DC i DS mają cechy ortogonalności i symetrii wyznaczenie macierzy funkcji bazowych transformat odwrotnych jest bardzo proste:

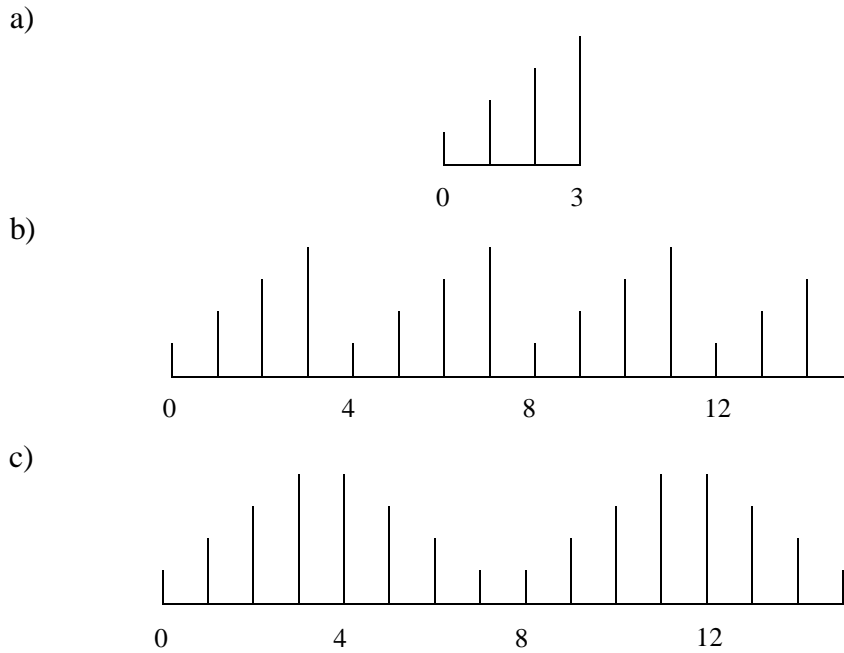
$$\begin{aligned}
[C^{\text{I}}]^{-1} &= [C^{\text{I}}]; & [C^{\text{II}}]^{-1} &= [C^{\text{III}}]; & [C^{\text{IV}}]^{-1} &= [C^{\text{IV}}]; \\
[S^{\text{I}}]^{-1} &= [S^{\text{I}}]; & [S^{\text{II}}]^{-1} &= [S^{\text{III}}]; & [S^{\text{IV}}]^{-1} &= [S^{\text{IV}}]. \quad (11.36)
\end{aligned}$$

Ponadto wykorzystywana jest też dyskretna transformata Fouriera (DFT), głównie ze względu na powszechność zastosowań szybkich algorytmów transformacji, jak też możliwość wykorzystania w pewnych szczególnych przypadkach kompresji danych [5]. Generalnie jednak potencjalne możliwości DFT w zastosowaniach kompresji danych są mniejsze niż DCT oraz DST. Świadczą o tym porównania przy pomocy różnych wskaźników potencjału kompresji przedstawione poniżej. Widać to także wyraźnie przy porównaniu periodyczności transformat DF i DC - patrz rys. 11.4. Powstanie wysokoczęstotliwościowych składowych transformaty o znaczących wartościach powoduje ograniczenie efektywności kodowania takiej reprezentacji, zmniejsza się ilość energii upakowana w pierwszych kilku wartościach współczynników i przy zachowaniu takiej samej długości kodu wyjściowego pojawiają się większe zniekształcenia.

Z transformat niesinusoidalnych na szczególną uwagę zasługuje transformata Walsh-Hadamard'a (WHT) [6]. Jakkolwiek skromna w możliwościach dekorelacji, ze względu na prostotę i możliwość szybkiej implementacji (nie wymaga realizacji mnożeń, wartości funkcji bazowych wynoszą ± 1) stała się popularna w zastosowaniach sprzętowych.

Porównanie efektywności transformat unitarnych

Aby zbadać użyteczność różnego typu transformat w algorytmach transformacyjnego kodowania wykorzystanych zostanie kilka kryteriów, a także prosty model danych wejściowych opisanych źródłem Markowa 1-go rzędu. Korelację sąsiednich wartości próbek w strumieniu danych wejściowych opisuje współczynnik korelacji ρ . Wartość tego współczynnika dla obrazów zaszumionych lub też z dużą ilością drobnych, wyraźnych szczegółów wynosi około 0.5, natomiast dla obrazów z małą liczbą szczegółów jest bliska jedności.



Rys. 11.4. Porównanie okresowości transformat DC i DF z punktu widzenia zastosowań w kompresji. Na rys. a) przedstawiony jest cztero-próbkowy sygnał wejściowy. Okresowe rozwinięcie sygnału dla DFT na rys. b) wynika z założenia o okresowości transformowanego sygnału. Widoczna jest silna nieciągłość pomiędzy próbkami 3 i 4, 7 i 8 oraz 11 i 12. Ponieważ DCT (n wymiarowa) jest częścią rzeczywistą $2n$ DFT sygnału powstałego poprzez lustrzane odbicie n pierwszych próbek (pomnożonego jeszcze przez pewien czynnik - zobacz równanie (11.31)), uzyskujemy rozszerzenie sygnału jak na rys. c). W tym przypadku nie ma niewygodnych nieciągłości dla sygnałów, w których wyraźnie różnią się wartości próbek granicznych (cyklicznie próbek $n-1$ oraz 0).

Jako średnia wartość współczynnika korelacji charakteryzującego obraz naturalny wykorzystywana jest najczęściej liczba 0.9. Postać macierzy kowariancji dla źródeł Markowa 1-szego rzędu jest następująca:

$$\mathbf{R}_\rho = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{bmatrix} \quad (11.37)$$

Na początek przyjrzyjmy się kształtowi funkcji bazowych transformat unitarnych. Zależna od sygnału baza KLT w przykładzie 11.1 nabiera charakteru jądra transformaty Walsh-Hadamarda. Jeśli natomiast do opisu sygnału użyjemy jednowymiarowego modelu źródła Markowa rzędu 1 ze współczynnikiem korelacji ρ , wówczas baza KLT nabiera charakteru sinusoidalnego. Otóż w takim przypadku wektory własne oraz wartości własne macierzy \mathbf{R}_ρ określone są równaniami (według [1]):

$$\varphi_{i,k} = \left[\frac{2}{n + \lambda_i^2} \right]^{1/2} \sin \left[w_i \left(k - \frac{n-1}{2} \right) + \frac{(i+1)\pi}{2} \right] \quad (11.38)$$

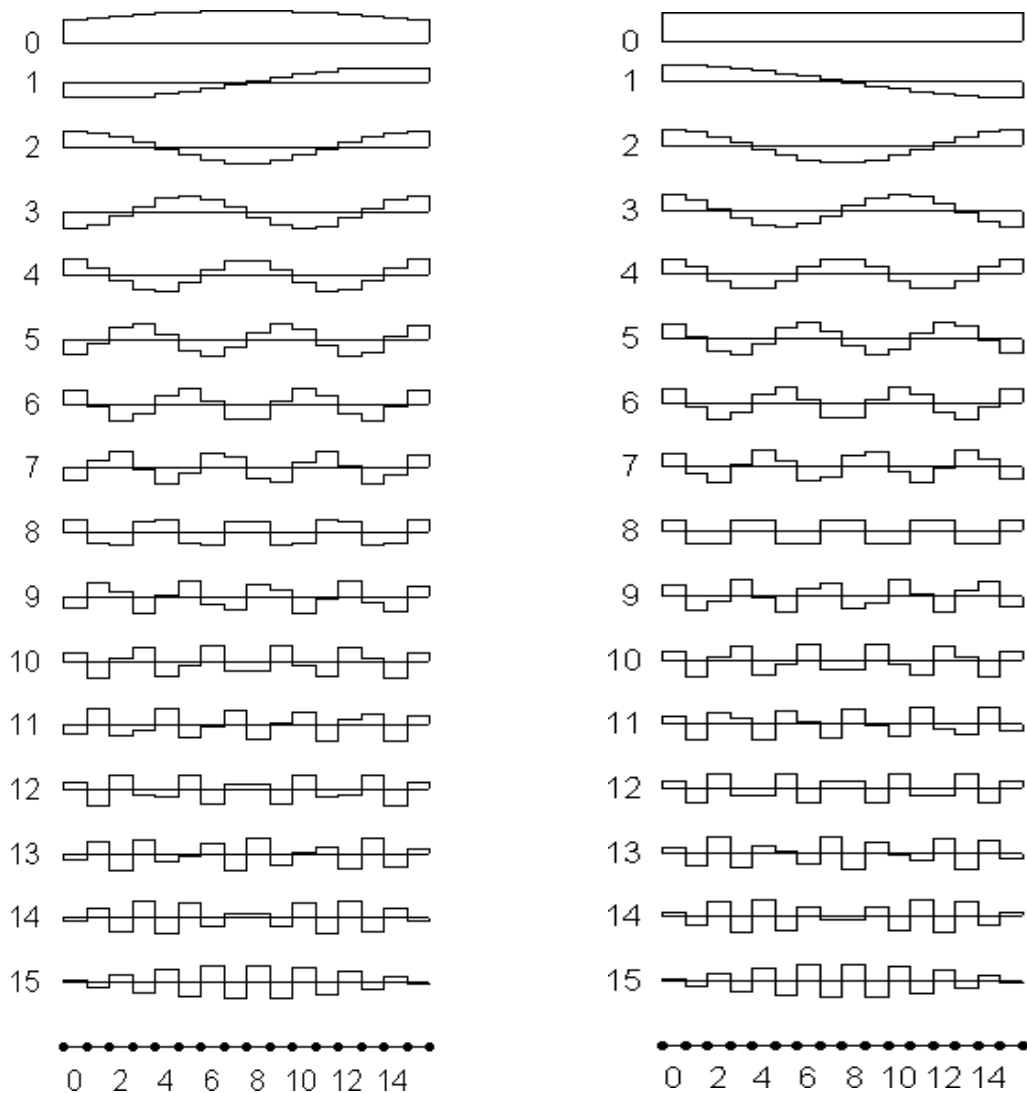
oraz

$$\lambda_i = \frac{1 - \rho^2}{1 - 2\rho \cos[w_i] + \rho^2} \quad (11.39)$$

dla $0 \leq i, k \leq n-1$, gdzie w_i jest pierwiastkiem równania przestępnego postaci:

$$\tan(nw) = \frac{(1 - \rho^2) \sin(w)}{\cos(w) - 2\rho + \rho^2 \cos(w)} \quad (11.40)$$

Na rys. 11.5 przedstawiono funkcje bazowe KLT oraz, dla porównania, funkcje bazowe DCT-III. Podobieństwo jest bardzo widoczne, co wydaje się już na wstępie potwierdzać dużą skuteczność przybliżania KLT przez DCT. Sprawdźmy to poprzez bardziej obiektywne kryteria.



Rys.11.5. Z lewej przedstawione są funkcje bazowe KLT (dla procesu Markowa 1-szego rzędu, $\rho=0.9$), z prawej zaś funkcje bazowe DCT-III. Widoczne jest duże podobieństwo obu baz.

Pierwszym ważnym kryterium jest wariancja poszczególnych współczynników transformaty w n -wymiarowym wektorze danych transformowanych. Jest ona miarą energii sygnału

zawartej w wartościach odpowiadającego danemu współczynnikowi wektora bazowego (utworzonego z funkcji bazowej jądra transformaty). Większa wartość wariancji oznacza więc większą ilość informacji zawartą w wartościach tego wektora bazowego, czyli lepszą reprezentację sygnału wejściowego przez dany element jądra transformaty. Jeśli niemal cała informacja zawarta w sygnale będzie wyrażona w wartościach jedynie kilku wektorów bazowych, wówczas otrzymujemy bardzo oszczędną reprezentację danych w przestrzeni transformaty, podatną na kompresję. Jeśli przez Φ oznaczmy macierz transformaty, to wektor wariancji otrzymujemy poprzez operację:

$$\sigma_i^2 = (\Phi \mathbf{R}_{xx} \Phi^T)_{i,i}, \quad i = 1, \dots, n \quad (11.41)$$

Rozkład wartości wariancji współczynników różnych transformat został przedstawiony w tabeli 11.1. Umożliwia on porównanie ilości informacji przenoszanej przez kolejne współczynniki transformat n - wymiarowego wektora. Bardziej wygodną do oceny skuteczności transformat w kompresji danych wydaje się całkowita miara upakowania energii, np. suma wartości wariancji m kolejnych współczynników.

Tabela 11.1. Rozkład wariancji współczynników transformat $n=16$ elementowego wektora danych wejściowych przybliżonych modelem Markowa rzędu 1 o współczynniku korelacji $\rho = 0.9$. Literka i oznacza numer współczynnika transformaty w wektorze.

i	Transformata										
	KLT	DFT	WHT	DCT				DST			
				I	II	III	IV	I	II	III	IV
1	9.9268	9.8346	9.8346	9.8678	9.8346	8.9324	8.6494	9.2183	8.9142	8.9324	8.6494
2	2.9487	1.8342	2.5360	2.9244	2.9327	3.3212	3.3195	2.6418	2.4576	3.3212	3.3195
3	1.1278	1.8342	1.0200	1.1416	1.2108	1.2934	1.2788	1.4678	1.4326	1.2934	1.2788
4	0.5682	0.5189	0.7060	0.5587	0.5815	0.6984	0.7256	0.7091	0.6895	0.6984	0.7256
5	0.3411	0.5189	0.3070	0.3361	0.3483	0.4333	0.4418	0.5312	0.5607	0.4333	0.4418
6	0.2294	0.2502	0.3030	0.2283	0.2314	0.2879	0.3166	0.3139	0.3320	0.2879	0.3166
7	0.1671	0.2502	0.2830	0.1708	0.1685	0.2161	0.2318	0.2628	0.3123	0.2161	0.2318
8	0.1290	0.1553	0.1060	0.1329	0.1295	0.1581	0.1862	0.1738	0.2060	0.1581	0.1862
9	0.1043	0.1553	0.1050	0.1126	0.1047	0.1319	0.1510	0.1526	0.2111	0.1319	0.1510
10	0.0875	0.1126	0.1050	0.0942	0.0877	0.1029	0.1302	0.1102	0.1493	0.1029	0.1302
11	0.0758	0.1126	0.1040	0.0867	0.0759	0.0920	0.1132	0.0988	0.1624	0.0920	0.1132
12	0.0675	0.0913	0.1040	0.0760	0.0675	0.0763	0.1028	0.0777	0.1211	0.0763	0.1028
13	0.0616	0.0913	0.1030	0.0743	0.0616	0.0716	0.0942	0.0707	0.1381	0.0716	0.0942
14	0.0574	0.0811	0.1020	0.0675	0.0574	0.0635	0.0892	0.0608	0.1075	0.0635	0.0892
15	0.0547	0.0811	0.0980	0.0692	0.0547	0.0618	0.0855	0.0569	0.1277	0.0618	0.0855
16	0.0531	0.0780	0.0780	0.0588	0.0531	0.0593	0.0840	0.0535	0.0780	0.0593	0.0840

Nieco inne kryterium efektywności upakowania energii (ang. energy packing efficiency – *EPE*) zostało zaproponowane przez Kitajima [7]. Jest ono wyrażone stosunkiem sumy wariancji m -pierwszych współczynników do sumy wariancji wszystkich n -współczynników, zdefiniowanym w sposób następujący:

$$EPE = \frac{\sum_{i=1}^m \sigma_i^2}{\sum_{i=1}^n \sigma_i^2} \quad (11.42)$$

Wartości miary EPE , analogiczne do prezentowanych w tabeli 11.1, zawiera tabela 11.2. Na podstawie tabeli 11.2. widać dużą efektywność transformaty DC (szczególnie I i II) ze względu na zdolność upakowania energii w wartościach kolejnych współczynników.

Tabela 11.2. Wartości miary upakowania energii EPE dla procesu Markowa 1-szego rzędu o współczynniku korelacji $\rho=0.9$, transformowanego w $n=16$ elementowych wektorach danych. Literka m (z równania 11.42) oznacza liczbę współczynników transformaty, w których zawarta jest względna ilość energii sygnału opisana przez EPE .

m	EPE										
	KLT	DFT	WHT	DCT				DST			
				I	II	III	IV	I	II	III	IV
1	0.6204	0.6147	0.6147	0.6167	0.6147	0.5583	0.5406	0.5761	0.5571	0.5583	0.5406
2	0.8047	0.7293	0.7732	0.7995	0.7980	0.7658	0.7481	0.7413	0.7107	0.7658	0.7481
3	0.8752	0.8439	0.8369	0.8709	0.8736	0.8467	0.8280	0.8330	0.8003	0.8467	0.8280
4	0.9107	0.8764	0.8810	0.9058	0.9100	0.8903	0.8733	0.8773	0.8434	0.8903	0.8733
5	0.9320	0.9088	0.9002	0.9268	0.9317	0.9174	0.9009	0.9105	0.8784	0.9174	0.9009
6	0.9464	0.9244	0.9192	0.9411	0.9462	0.9354	0.9207	0.9301	0.8992	0.9354	0.9207
7	0.9568	0.9401	0.9368	0.9517	0.9567	0.9489	0.9352	0.9466	0.9187	0.9489	0.9352
8	0.9649	0.9498	0.9435	0.9600	0.9648	0.9588	0.9469	0.9574	0.9316	0.9588	0.9469
9	0.9714	0.9595	0.9500	0.9671	0.9714	0.9670	0.9563	0.9670	0.9448	0.9670	0.9563
10	0.9769	0.9665	0.9566	0.9730	0.9769	0.9735	0.9644	0.9738	0.9541	0.9735	0.9644
11	0.9816	0.9736	0.9631	0.9784	0.9816	0.9792	0.9715	0.9800	0.9642	0.9792	0.9715
12	0.9858	0.9793	0.9696	0.9831	0.9858	0.9840	0.9779	0.9849	0.9718	0.9840	0.9779
13	0.9897	0.9850	0.9760	0.9878	0.9897	0.9885	0.9838	0.9893	0.9804	0.9885	0.9838
14	0.9933	0.9900	0.9824	0.9920	0.9933	0.9924	0.9894	0.9931	0.9872	0.9924	0.9894
15	0.9967	0.9951	0.9885	0.9963	0.9967	0.9963	0.9947	0.9967	0.9951	0.9963	0.9947
16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Innym ciekawym kryterium jest korelacja resztkowa (ang. residual correlation), która pokazuje procentowo poziom korelacji danych, jaki pozostał w dziedzinie transformaty. Innymi słowy jest to miara nadmiarowości reprezentacji danej transformaty, która zmniejsza efektywność kompresji. Definiowana jest ona w sposób następujący: niech σ oznacza macierz diagonalną z wartościami wariancji obliczonymi według (11.41):

$$\sigma = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix}, \quad (11.43)$$

a $\tilde{\mathbf{R}}_\rho$ będzie aproksymacją macierzy kowariancji w dziedzinie danych oryginalnych przez odwrotną transformatę σ według równania:

$$\tilde{\mathbf{R}}_\rho = \Phi^T \sigma \Phi. \quad (11.44)$$

Korelację resztkową definiuje się wówczas według równania:

$$RC = \frac{|\mathbf{R}_\rho - \tilde{\mathbf{R}}_\rho|^2}{|\mathbf{R}_\rho - \mathbf{I}|^2}, \quad (11.45)$$

gdzie $|\mathbf{M}|^2$ oznacza słabą normę Schmidta zdefiniowaną według:

$$|\mathbf{M}| = \frac{1}{n} \sum_{i,j=1}^n |(M)_{i,j}|^2. \quad (11.46)$$

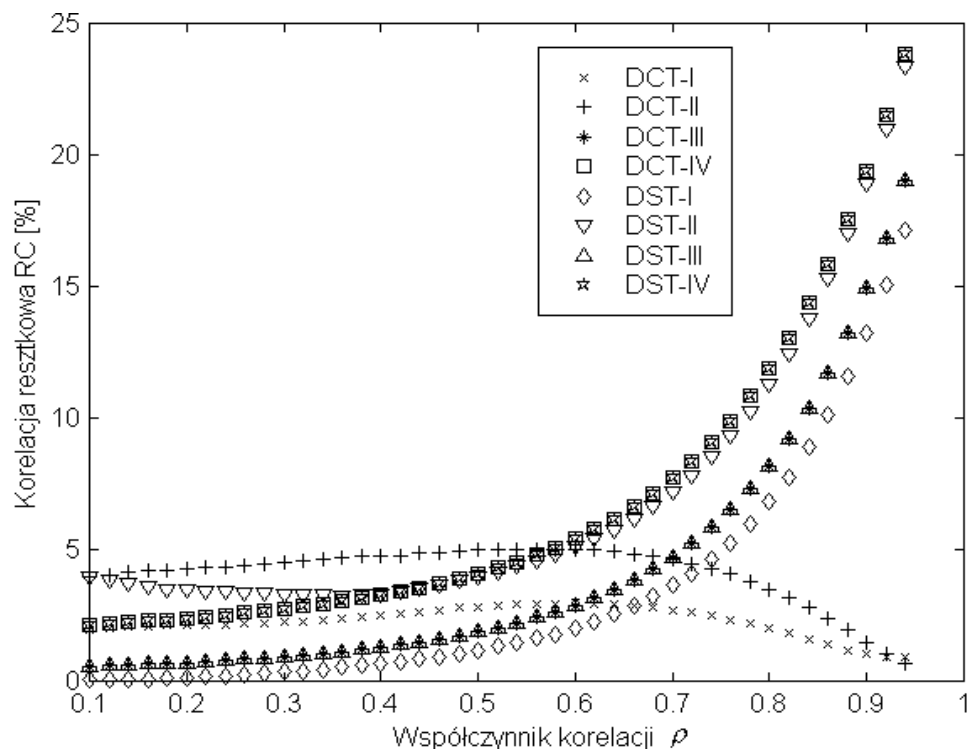
Zależność korelacji resztkowej od współczynnika korelacji dla różnych transformat przedstawiono na rys. 11.6. Oczywiście resztkowa korelacja dla KLT wynosi zero bez względu na wartość współczynnika ρ .

Jeszcze bardziej przydatne z punktu widzenia kompresji kryteria oceny transformat to miara stopnia zniekształceń R-D oraz maksymalna zdolność redukcji bitów (ang. maximum reducible bits – MRB). Funkcja stopnia zniekształceń $R(T,D)$ może być rozumiana jako ilość informacji (w bitach) przypadająca na współczynnik transformaty T , która musi być zakodowana przy założonym poziomie zniekształceń D . Im większa to wartość dla danej wartości zniekształcenia, tym mniejsza skuteczność kompresji. Miara stopnia zniekształceń musi zakładać pewien schemat kwantyzacji. Najprościej można przyjąć prosty mechanizm przydziału bitów według funkcji stopnia zniekształceń kwantyzatora Shannona (zobacz następny punkt tego podrozdziału). Wtedy funkcję stopnia zniekształceń można zdefiniować następująco:

$$R(T,D) = \frac{1}{2n} \sum_{i=1}^n \max \left\{ 0, \log_2 \frac{\sigma_i^2}{\theta} \right\}, \quad (11.47)$$

gdzie θ determinuje wartość zniekształceń D :

$$D(\theta) = \frac{1}{n} \sum_{i=1}^n \min(\theta, \sigma_i^2) \quad (11.48)$$



Rys. 11.6. Zależność resztkowej korelacji od wartości współczynnika korelacji modelu Markowa 1-szego rzędu, modelującego źródło informacji, dla transformat DC i DS.

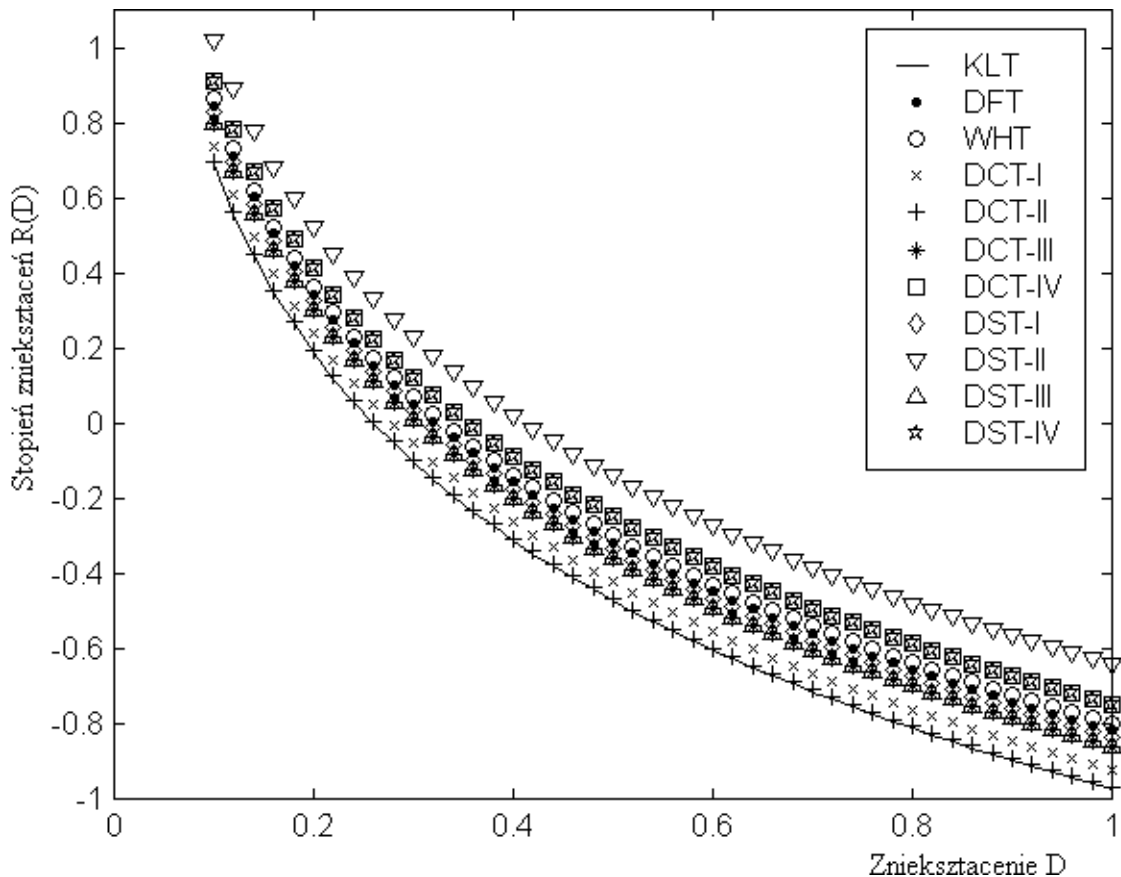
Przy założeniu małego poziomu zniekształceń (niewielkich stopni kompresji) i spełnieniu warunku $\theta < \sigma_i^2$, dla $i = 1, \dots, n$ mamy $D = \theta$ i równania (11.47) i (11.48) można zapisać w postaci jednego wyrażenia:

$$R(T, D) = \frac{1}{2n} \sum_{i=1}^n \log_2 \sigma_i^2 - \frac{1}{2} \log_2 D \quad (11.49)$$

Na rys. 11.7. wykreślono funkcje stopnia zniekształceń na podstawie zależności (11.49) dla różnych transformat unitarnych. Na szczególną uwagę zasługuje fakt praktycznie pokrywania się tej funkcji dla optymalnej KLT oraz DCT II, co szczególnie pokazuje przydatność transformaty kosinusowej w algorytmach stratnej kompresji obrazów, zawierających mechanizm kwantyzacji współczynników transformaty. Pierwszy składnik wyrażenia (11.49) zależy od rodzaju transformaty, gdyż $\sigma_i^2 = \sigma_i^2(\varphi) = \sigma_i^2(T)$ zgodnie z równaniem (11.41). Jego negacja jest definiowana jako maksymalna zdolność redukcji bitów *MRB* w sposób następujący:

$$MRB = -\frac{1}{2n} \sum_{i=1}^n \log_2 \sigma_i^2. \quad (11.50)$$

Większa wartość *MRB* dla współczynników danej transformaty oznacza większą redukcję bitów potrzebnych do zapisu informacji. Rozkład wariancji poszczególnych współczynników umożliwia oszczędniejszy zapis wartości współczynników. Wartości parametru *MRB* dla różnych transformat przedstawiono w tabeli 11.3.



Rys. 11.7. Funkcja stopnia zniekształceń dla wybranych transformat unitarnych.

TABELA 11.3. Wartości maksymalnej zdolności redukcji bitów MRB dla wybranych transformat unitarnych; $n = 16$, na wejściu proces Markowa 1-szego rzędu z $\rho = 0.9$.

Transformata	MRB
KLT	1.1231
DFT	0.9485
WHT	0.9374
DCT-I	1.0734
DCT-II	1.1172
DCT-III	1.0113
DCT-IV	0.8983
DST-I	0.9848
DST-II	0.7915
DST-III	1.0113
DST-IV	0.8983

Zamodelowanie statystycznych zależności pomiędzy wartościami sąsiednich pikseli w obrazie za pomocą źródła Markowa pierwszego rzędu i sklasyfikowanie transformat ortogonalnych według takich kryteriów jak rozkład wariancji współczynników (wprost proporcjonalnej do zawartej informacji) i upakowanie energii, resztkowa korelacja (miara korelacji jaka pozostała w przestrzeni transformaty), stopień zniekształceń oraz maksymalna zdolność redukcji bitów doprowadziło do stwierdzenia, iż generalnie, a także według większości z tych kryteriów, DCT jest transformacją optymalną z punktu widzenia zastosowań transformacyjnego kodowania. Bardziej szczegółowa analiza prezentowanych wyników pokazuje, że w upakowaniu energii DCT prawie dorównuje KLT (praktycznie zrównuje się zawartość energii dla sześciu i więcej współczynników). Jeszcze wyraźniejsza dominacja DCT w stosunku do pozostałych transformat unitarnych, które by mogły zastąpić KLT w algorytmach kompresji, jest widoczna w przypadku funkcji stopnia zniekształceń oraz zdolności redukcji bitów kodu. Przy czym najlepsze wyniki osiągnęte są dla II i I rodzaju DCT. Nieco bardziej złożone wnioski wynikają z analizy wartości korelacji resztkowej dla różnych rodzajów danych wejściowych symulowanych współczynnikiem korelacji w modelu Markowa. Należy tutaj wyróżnić dwa rodzaje źródeł danych: o słabej korelacji (ρ poniżej 0.7), dla których najlepsze wyniki daje DST (I rodzaj), oraz o silnej korelacji (ρ powyżej 0.7), gdzie najlepszą skuteczność daje DCT, szczególnie-II rodzaju.

11.4. Optymalizacja algorytmu kodowania transformacyjnego

Aby uczynić kodowanie transformacyjne metodą kompresji o dużych walorach użytkowych w szerokiej gamie zastosowań konieczne jest skuteczne rozwiązanie dodatkowo kilku zagadnień wpływających w sposób decydujący na jakość kompresji, tj. kwantyzacji i odwracalnego kodowania współczynników transformat.

Porównanie technik pełnokadrowych i blokowych

Metody transformacyjnego kodowania zakładają wykonanie w etapie wstępnym dekompozycji oryginalnych danych obrazowych przy pomocy transformaty unitarnej, w zdecydowanej większości przypadków DCT. Przekształcenie to może być realizowane w skali całego obrazu (techniki pełnokadrowe) lub też niezależnie w blokach, którymi pokrywana jest przestrzeń obrazu. Jedne i drugie rozwiązanie znalazło zastosowanie w praktyce, zasadniczo jednak blokowa transformata jest bardziej użyteczna ze względu na zalety istotne w praktycznej implementacji oraz często lepszą jakość rekonstruowanych obrazów. Przykładowe rozwiązania wraz z uwypukleniem zalet i wad obu sposobów realizacji transformaty przedstawiono poniżej.

Techniki pełnokadrowe DCT

Lo i Huang [8] zastosowali pełnokadrową 2-D DCT do kompresji obrazów ($512 \times 512 \times 8b$) radiograficznych płuc, subtrakcyjnej angiografii, arteriografii, CT, USG oraz obrazu fantomu. Jako kryteria oceny jakości obrazów rekonstruowanych zastosowano NMSE, obserwację obrazów różnicowych oraz subiektywną klasyfikację obrazów o różnym stopniu kompresji (zobacz rozdz. 8), według których dopuszczalny stopień kompresji wynosi 16. Wykorzystany algorytm kompresji zawiera realizację kwantyzacji i kodowania współczynników DCT przy pomocy techniki przydziału bitów, polegającej na tworzeniu pełnokadrowej tablicy przydziału bitów i odpowiedniej kwantyzacji. Tablica ta musi być dopisana do zbioru danych skompresowanych, by umożliwić rekonstrukcję obrazu. Bardzo podobne wyniki uzyskano także w badaniach przeprowadzonych na Uniwersytecie Kalifornijskim [9].

Templeton [10] zastosował algorytm pełnokadrowej kompresji DCT oceniając subiektywnie jakość rekonstruowanych obrazów CT, głównie płuc, 10 i 12 bitowych o rozdzielczości 2048×2048 . Zauważono, iż obrazy o niższej dynamice (większa korelacja) i słabym kontraście (bardziej nierównomierny histogram) kompresują się lepiej. Podjęto także ciekawą próbę wielokrotnej kompresji obrazów przy pomocy tej samej techniki. Oczywiście, skutek był taki, iż znacznie wzrósł poziom szumów w tych obrazach, nie dając oczekiwanych, coraz większych stopni kompresji.

Metody wykorzystujące blokową DCT

Zalety blokowej transformacji DCT w stosunku do pełnokadrowej są znaczące, powodując w wielu zastosowaniach większą przydatność metod blokowych. Doprowadziło to m.in. do powstania standardu JPEG opartego na blokowej DCT. Wady metod blokowych są jednak także dobrze znane, w związku z czym podejmowanych jest szereg prób minimalizacji niepożądanych efektów [11][12][13].

Do zalet metod blokowych w pierwszej kolejności należy ich adaptacyjność, pozwalająca przy wyższych stopniach kompresji lepiej zachować ostre krawędzie w ważnych obszarach, czy też skuteczniej wyeliminować nieistotną informację (szumy, artefakty metody obrazowania) w innych częściach obrazu. Jest to szczególnie istotne przy kompresji obrazów medycznych, czy też wszystkich innych o silnie zróżnicowanej wadze prezentowanej informacji. Inną zaletą są oszczędności czasowe (mniejsza ilość koniecznych obliczeń), a także znacznie mniejsze wymagania pamięciowe (rozmiar bufora niezbędnego do przechowania wartości współczynników transformaty). Przy zastosowaniu szybkich algorytmów liczenia transformat, wymagających wykonania liczby operacji rzędu $n \log_2 n$ przy transformacji wektora $n \times 1$, koszt obliczeń w porównaniu z technikami pełnokadrowymi

redukowany jest w stosunku $\log_2 NM / \log_2 nm$ (dla obrazu $N \times M$ i rozmiarach bloku $n \times m$), natomiast wymagania pamięciowe są zmniejszone w stosunku NM / nm .

Ponadto, blokowe przetwarzanie umożliwia większe zrównoleglenie obliczeń, tak bardzo korzystne w implementacjach sprzętowych, oraz ogranicza występowanie charakterystycznych dla kompresji pełnokadrowej artefaktów Gibbsa. Są to widoczne pierścienie pojawiające się w obszarach wokół ostrych krawędzi, związane z eliminacją składowych wysokoczęstotliwościowych przy wyższych stopniach kompresji. W technice blokowej (np. bloki 8×8) niewielki rozmiar bloku ogranicza propagację tych artefaktów, a ponadto umożliwia lepsze zachowanie składowych o wyższej częstotliwości.

Kwantyzacja współczynników transformat

Obok przeprowadzenia efektywnej transformacji na blokach obrazu oryginalnego o odpowiednio dobranych rozmiarach w technikach TE istotny jest także proces kwantyzacji współczynników transformaty, które są następnie odwracalnie kodowane. W przeważającej większości przypadków aplikowane są schematy kwantyzacji skalarnej, przy czym nierzadko wykorzystywane są techniki przydziału bitów (ang. bit allocation), pozwalające uzyskać dokładnie założoną *a priori* długość reprezentacji współczynników. I tak wykorzystywane są różne metody klasyfikacji bloków, używane różne schematy kwantyzacji równomiernej i nierównomiernej, ustalana jest liczba poziomów kwantyzacji dla poszczególnych klas wektorów danych, utworzonych przez współczynniki z kolejnych bloków obrazu, itp. Z przypisanej ilości poziomów wynika liczba bitów konieczna do zapisu skwantowanych wartości współczynników danego wektora. Zasadnicza idea technik przydziału bitów polega na wierniejszym opisie współczynników niosących istotną informację poprzez znaczący przydział bitów do zapisu ich wartości oraz eliminacji wektorów danych mało znaczących poprzez zerową alokację bitów (usuwanie ich wartości).

Przydział bitów

Najczęściej miarą znaczenia elementu bloku czy innego wektora danych przy niezależnej jego kwantyzacji jest wariancja zmiennej losowej (σ_i^2) modelującej dany element, charakteryzująca liczbę informacji przenoszonej przez wartości tej zmiennej. Zastosowanie optymalnej statystycznie KLT lub innej transformaty aproksymującej jej zdolności dekorelacji powoduje upakowanie energii sygnału w wartościach zaledwie kilku współczynników, podczas gdy zmienność wartości pozostałych współczynników jest bliska zeru. Oznacza to, że występuje silne zróżnicowanie wartości wariancji σ_i^2 , a więc przydział liczby bitów potrzebny do reprezentacji tych wartości powinien być również nierównomierny.

Załóżmy, że mamy n -wymiarowy wektor losowy \mathbf{X} , którego składowymi są zmienne losowe X_1, \dots, X_n (o zerowej wartości średniej i wariancji σ_i^2). Kwantyzacja wektora jest skalarna, tzn. wartości każdej zmiennej losowej kwantowane są niezależnie. Oznacza to, że należy zdefiniować n niezależnych kwantyzatorów tak, aby jak najlepiej wykorzystać założony budżet bitowy R (z minimalną wartością zniekształcenia). Uzyskanie możliwie najmniejszych zniekształceń przy danym R możliwe jest poprzez przydział odpowiedniej liczby bitów dla poszczególnych kwantyzatorów składowych wektora. Liczbę przydzieloną poszczególnym zmiennym losowym oznaczmy przez r_i . Miara zniekształceń kwantyzacji powinna dawać liniowy wzrost zniekształcenia w funkcji wariancji sygnału wejściowego oraz funkcji określającej jednostkowe zniekształcenia wywołane przydziałem określonej liczby bitów r_i . Można ją zapisać następująco:

$$D = \sum_{i=1}^n \sigma_i^2 g(r_i), \quad (11.51)$$

gdzie funkcja $g(r)$ oznacza średniokwadratowe zniekształcenie powstające na r -bitowym kwantyzatorze, jeśli na wejście podawana jest zmienna losowa o jednostkowej wariancji. Nazwijmy ją funkcją stopnia zniekształceń kwantyzatora. Zależy ona od statystycznego opisu sygnału wejściowego (najczęściej jest to funkcja gęstości prawdopodobieństwa) oraz rodzaju kwantyzatora (równomierny, Lloyd-Maxa, itd.). Nie ma niestety ogólnej postaci wyrażenia na $g(\cdot)$, stąd też wiele prób optymalnego rozwiązania problemu optymalnego przydziału bitów dla różnych zagadnień. Intuicyjnie można przyjąć, że funkcja $g(\cdot)$ jest wypukła, ma pierwszą pochodną, która zeruje się w (granicy do) nieskończoności. Przy takich założeniach można dowieść [14], że wyrażenie na optymalny przydział bitów, minimalizujący wyrażenie (11.51), ma postać

$$r_i = \begin{cases} h\left(\frac{\theta g'(0)}{\sigma_i^2}\right) & \text{dla } 0 < \theta < \sigma_i^2, \\ 0 & \text{dla } \theta \geq \sigma_i^2 \end{cases}, \quad (11.52)$$

gdzie $h(\cdot)$ jest funkcją odwrotną pochodnej $g'(\cdot)$, czyli liczba bitów r_i jest dobierana tak, że $g'(r_i) = \frac{\theta}{\sigma_i^2} g'(0)$. Wartość θ jest dobierana ze względu na założoną średnią długość kodu i jest pierwiastkiem równania

$$\sum_{i:\sigma_i^2 > \theta} r_i = \sum_{i:\sigma_i^2 > \theta} h\left(\frac{\theta g'(0)}{\sigma_i^2}\right) = nR. \quad (11.53)$$

Wartość ta pełni funkcję progu dla wartości wariancji współczynników. Zerowane są bowiem wartości tych współczynników, których wariancja nie przekracza wartości progu. Minimalna wartość zniekształcenia w takim schemacie kwantyzacji składa się z dwu elementów: zniekształcenia zmiennych o znaczącej wariancji (większej od wartości θ) oraz zniekształcenia zmiennych o małej wariancji (równej sumie wariancji usuniętych elementów wektora losowego poprzez zerowy przydział bitów) i wynosi

$$D_{\min} = \sum_{i:\sigma_i^2 > \theta} \sigma_i^2 g(r_i) + \sum_{i:\sigma_i^2 \leq \theta} \sigma_i^2 \quad (11.54)$$

Jakkolwiek wyrażenie (11.52) wydaje się nieco skomplikowane, to w praktyce schemat przydziału bitów może być bardzo prosty. Pokazuje to przykład kwantyzatora Shannona z funkcją stopnia zniekształceń postaci

$$g(r) = 2^{-2r} \quad (11.55)$$

Taka postać funkcji zapewnia 'idealny' optymalny kwantyzator blokowy dla Gaussowskich źródeł bez pamięci. Wyznacza on teoretyczną granicę dla praktycznych koderów, służąc do oceny ich efektywności. Postać funkcji $g(\cdot)$ kwantyzatora Shannona pokazuje rys. 11.8.

Pochodna funkcji Shannona (11.55) oraz jej odwrotność mają następującą postać:

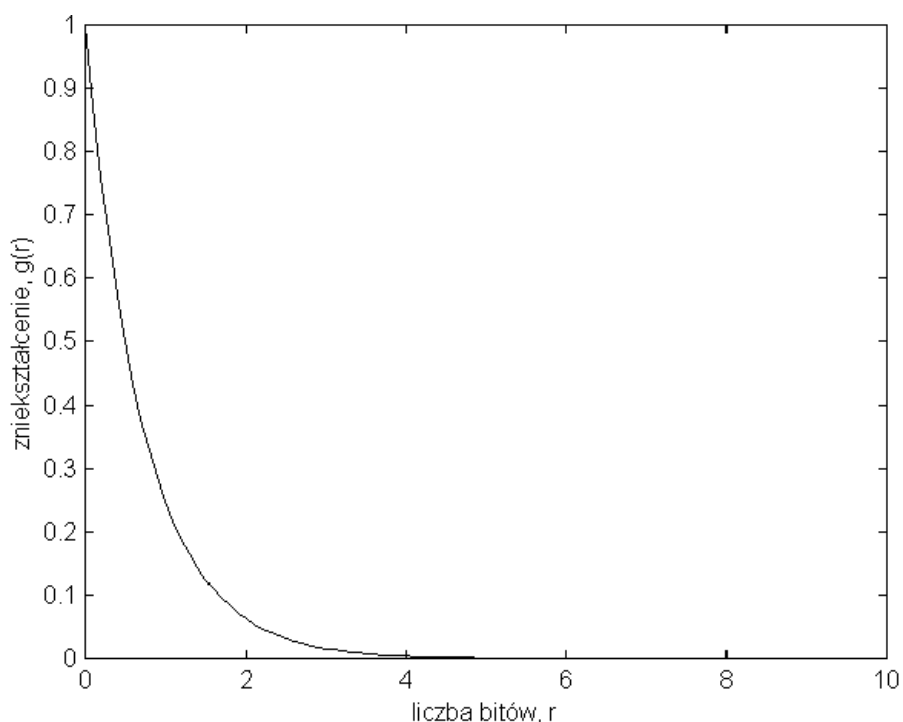
$$g'(r) = -(2 \ln 2) \cdot 2^{-2r}, \quad (11.56)$$

$$h(x) = \frac{1}{2} \log_2 \frac{-2 \ln 2}{x}, \quad (11.57)$$

co daje równanie określające przydział bitów w postaci

$$r_i = \max\left(0, \frac{1}{2} \log_2 \frac{\sigma_i^2}{\theta}\right), \quad (11.58)$$

gdzie wartość θ jest tak dobrana, by średnia bitowa reprezentacji kodo wej była równa R .



Rys. 11.8. Shannon'owska funkcja stopnia zniekształceń kwantyzatora.

Ciekawą cechą tego modelu przydziału bitów jest fakt, iż dla dużych wartości założonej średniej bitowej strumienia wartości skwantowanych R , kiedy dla wszystkich elementów wektora wariancja jest większa od parametru θ , błąd kwantyzacji każdej zmiennej d_i jest jednakowy, niezależny od σ_i^2 . Wynika to z równań (11.51), (11.55) i (11.57):

$$d_i = \sigma_i^2 g(r_i) = \sigma_i^2 \cdot 2^{-2 \log_2 \frac{\sigma_i^2}{\theta}} = \theta. \quad (11.59)$$

Postać optymalnego błędu kwantyzacji z funkcją Shannona jest więc następująca:

$$D_{\min} = \sum_{i=1}^n \min(\theta, \sigma_i^2). \quad (11.60)$$

Rozważmy nieco bardziej złożony model funkcji stopnia zniekształceń kwantyzatora. Zakładając znaną postać funkcji gęstości prawdopodobieństwa dla każdej zmiennej losowej

$f_{X_i}(x)$ poszukujemy n optymalnych kwantyzatorów, określonych przez liczbę przydzielonych bitów r_i . Przyjmując model wysokorozdzielczy, gdzie poziom zniekształcenia jest dużo mniejszy od wariancji sygnału ($D_i \ll \sigma_i^2$), można wykazać [15], że postać funkcji $g(\cdot)$ dla każdego kwantyzatora jest następująca

$$g(r_i) = a_i 2^{-2r_i}, \quad (11.61)$$

gdzie $a_i = \frac{1}{12} \left\{ \int_{-\infty}^{\infty} [f_{X_i}(x)]^{1/3} dx \right\}^3$. Wobec tego na podstawie równań (11.51) oraz (11.61)

zniekształcenie kwantyzatora o 2^{r_i} poziomach wynosi

$$D_i(B_i) \cong a_i \sigma_i^2 2^{-2r_i}, \quad (11.62)$$

Optymalny przydział bitów sprowadza się więc do minimalizacji wyrażenia

$$D = \sum_{i=1}^n a_i \sigma_i^2 2^{-2r_i} \quad (11.63)$$

ze względu na wartości przydzielonych bitów r_i , przy czym $\sum_{i=1}^n r_i = nR$.

Huang i Schultheiss [16] podali następującą postać optymalnego przydziału bitów przy takich założeniach

$$r_i = R + \frac{1}{2} \log_2 \frac{\sigma_i^2}{\rho^2} + \frac{1}{2} \log_2 \frac{a_i}{A}, \quad (11.64)$$

gdzie $\rho^2 = \left(\prod_{i=1}^n \sigma_i^2 \right)^{\frac{1}{n}}$ jest średnią geometryczną wartości wariancji, a $A = \left(\prod_{i=1}^n a_i \right)^{\frac{1}{n}}$ średnią geometryczną wartości a_i .

Taki schemat przydziału bitów daje optymalną (minimalną) wartość zniekształcenia

$$D_{\min} = NH \rho^2 2^{-2R} \quad (11.65)$$

Dla Gaussowskich rozkładów kwantowanych zmiennych (określających wartość współczynników a_i), przy założeniu optymalnego kwantyzatora Lloyd-Maxa (dla każdej zmiennej kwantyzacja do 2^{r_i} poziomów jest nierównomierna), ci sami Huang i Schultheiss wprowadzili zależność na alokację bitów postaci

$$r_i = R + \frac{1}{2 \ln 2} \left\{ \ln \sigma_i^2 - \frac{1}{N} \sum_{i=1}^N \ln \sigma_i^2 \right\} \quad (11.66)$$

Z kolei Wintz i Kurtenbach [17] zaproponowali bardzo podobną postać schematu przydziału bitów dla Gaussowskich zmiennych losowych przy założeniu optymalnej równomiernej kwantyzacji według liczby przydzielonych bitów

$$r_i = R + \frac{2}{\ln 10} \left\{ \ln \sigma_i^2 - \frac{1}{N} \sum_{i=1}^N \ln \sigma_i^2 \right\} \quad (11.67)$$

Podobieństwo wynika z podobieństwa optymalnych kwantyzatorów równomiernych i nierównomiernych dla źródeł Gaussowskich.

Poniżej przedstawiono dwa przykłady techniki alokacji bitów dla transformacyjnego kodowania z blokową oraz pełnokadrową transformatą.

PRZYKŁAD 11.2. Najprostszy, często wykorzystywany schemat przydziału bitów wygląda następująco: obraz dzielony jest na bloki, obliczana jest transformatą, a każdemu współczynnikowi o współrzędnych (u, v) w dziedzinie transformaty danego bloku przypisywana jest liczba bitów według zależności:

$$r(u, v) = \begin{cases} \log_2 \sigma_{F(u, v)}^2 + p, & \text{dla } \sigma_{F(u, v)}^2 \geq p \\ p & , \text{ dla } \sigma_{F(u, v)}^2 < p \text{ i } p \geq 0 \\ 0 & , \text{ wpp} \end{cases} \quad (11.68)$$

gdzie p jest wartością rzeczywistą, dobieraną w zależności od założonego stopnia kompresji (w tym rozwiązaniu nie zawsze jest kontrolowana średnia bitowa reprezentacji wyjściowej ze względu na dodatkowy etap kodowania strumienia danych po kwantyzacji, pozwalający zwiększyć efektywność kompresji). Taka tablica przydziału bitów z wartościami $r(u, v)$, ustalona w koderze, musi być przesłana do dekodera wraz z maksymalnymi wartościami poszczególnych współczynników $F_{\max}(u, v)$. Można także poszczególne bloki obrazu przypisać do pewnego zbioru klas o różnych własnościach, a następnie wyznaczyć tablice przydziału bitów oraz maksymalne wartości współczynników dla każdej z tych klas, co daje rozwiązanie bardziej adaptacyjne kosztem większej ilości dodatkowej informacji przesyłanej do dekodera.

Na podstawie przydziału według (11.68) odbywa się proces równomiernej kwantyzacji:

$$q_F(u, v) = \text{round}\{(2^{r(u, v)} - 1) \cdot F(u, v) / F_{\max}(u, v)\}, \quad (11.69)$$

gdzie $F_{\max}(u, v)$ - maksymalna wartość współczynnika (u, v) we wszystkich blokach obrazu (lub danej klasy). Odtwarzanie kwantowanych wartości w dekodrze wygląda następująco:

$$\tilde{F}(u, v) = F_{\max}(u, v) \cdot q_F(u, v) / (2^{r(u, v)} - 1). \quad (11.70)$$

PRZYKŁAD 11.3. Kwantyzację za pomocą techniki przydziału bitów można także zrealizować dla pełnokadrowej transformacji obrazu, aczkolwiek będzie to wersja uproszczona. Ponieważ rozważamy alokację bitów dla całego obrazu, mamy do dyspozycji tylko jeden duży blok współczynników, a więc nie sposób oszacować momentów statystycznych domniemanych zmiennych losowych. Wariancję muszą więc zastąpić wartości poszczególnych współczynników transformaty. Każdemu współczynnikowi o wartości $F(u, v)$ w dziedzinie transformaty $N \times N$ całego obrazu przypisywana jest liczba bitów według równania:

$$r(u, v) = \begin{cases} \lceil \log_2 [F(u, v)] \rceil + p, & \text{dla } |F(u, v)| \geq 1 \\ p, & \text{wpp} \end{cases}, \quad (11.71)$$

gdzie p jest liczbą rzeczywistą, dobieraną w zależności od żądanego stopnia kompresji. Tablica przydziału bitów zawiera wartości $r(u, v)$, przy czym każdej wartości współczynnika $F(u, v)$ przypisana jest jedna wartość. Jeśli tablica ta, ustalona w koderze, będzie przesłana do dekodera wraz z wartością p , trudno będzie mówić o kompresji, gdyż liczba elementów tej tablicy jest równa liczbie pikseli obrazu. Dodatkowo trzeba przecież przesłać jeszcze skwantowane wartości współczynników. Podział według wyznaczonych wartości $r(u, v)$ musi wobec tego zostać przybliżony schematem, który wykorzystuje jedynie maksymalną liczbę bitów r_{\max} przydzielonych największej wartości współczynnika F_{\max} . Proces kwantyzacji realizowany jest według następującego równania:

$$q_F(u, v) = \text{round}\{(2^{r_{\max}-1} - 1) \cdot F(u, v) / F_{\max}\}, \quad (11.72)$$

gdzie r_{\max} - liczba bitów przydzielona wartości maksymalnej. Wykorzystanie w równaniu (11.72) wyrażenia $r_{\max} - 1$ w wykładniku dwójki pozwala zarezerwować jeden bit na znak współczynnika. Można także na przydzielonej liczbie bitów zapisywać jedynie moduł wartości (jak w schemacie według (11.69)), znak kodując osobno, co oczywiście należy uwzględnić przy doborze wartości p .

Odtwarzanie kwantowanych wartości w dekodерze wygląda następująco:

$$\tilde{F}(u, v) = F_{\max} \cdot q_F(u, v) / (2^{r_{\max}-1} - 1). \quad (11.73)$$

Przy takim rozwiązaniu dekodерowi oprócz skwantowanych wartości współczynników przesyłane są dodatkowo tylko dwie wartości: F_{\max} i r_{\max} (lub F_{\max} i p).

Selekcja próbek

Techniki przydziału bitów są często zastępowane metodami selekcji próbek, w których kontrola długości reprezentacji wyjściowej jest mniejsza, ale efektywność kompresji zazwyczaj rośnie ze względu na dobre sprzężenie strumienia wyjściowego kwantyzatora ze skutecznymi algorytmami odwracalnego kodowania. Występują także techniki kwantyzacji łączące selekcję próbek z metodami przydziału bitów.

Istnieją dwa zasadnicze schematy skalarnej kwantyzacji wartości współczynników transformaty z bloków z selekcją próbek. Pierwszy z nich wykorzystuje strefową selekcję próbek (ang. zonal sample selection), zachowując wartości współczynników tylko w pewnej strefie transformowanego bloku danych, poza którą wszystkie pozostałe współczynniki są zerowane bez względu na wartość. Ponieważ większość obrazów ma dolnopasmowe spektrum częstotliwościowe, taki model kwantyzacji pozwala zachować większość informacji zawartej w obrazie, usuwa jednak bezwzględnie współczynniki wyższych częstotliwości. Można powiedzieć, że techniki przydziału bitów są pewnym przypadkiem realizacji idei strefowej selekcji próbek. Drugim schematem kwantyzacji jest progowa selekcja próbek (ang. threshold sample selection), która pozwala na zachowanie każdego współczynnika w przypadku, gdy jego wartość bezwzględna przekracza pewną wartość progową.

Na rys. 11.9. widać ograniczenia obu metod selekcji. Transformata DCT bloku obrazu pokazuje, że występują w nim wyraźniej zarysowane krawędzie poziome i współczynniki w kierunku pionowym mają dużo większe wartości od współczynników wzdłuż osi poziomej. Przy kwantyzacji z selekcją strefową przy standartowym kształcie strefy zachowującej

symetrycznie dziesięć współczynników niskoczęstotliwościowych przenoszących najczęściej zasadniczą informację o sygnale zgubione zostają współczynniki o dużych wartościach. Do lokalnej charakterystyki bloku dostosowana jest kwantyzacja z selekcją progową, zachowująca wszystkie znaczące współczynniki. Problemem jednak w tym przypadku jest nieregularny kształt obszaru znaczącego, co powoduje dodatkowe trudności w efektywnym zakodowaniu wartości tych współczynników. W obu przypadkach jako kwantyzacje wyselekcjonowanych współczynników zastosowano jedynie przybliżenie do najbliższej liczby całkowitej.

a)

109	101	91	79	80	85	84	89
87	87	89	99	98	91	91	93
100	102	109	127	112	99	87	84
84	89	87	88	83	90	100	94
100	101	108	112	110	114	113	109
103	103	95	89	89	89	82	86
84	97	100	100	108	103	106	112
111	115	111	113	107	100	109	107

b)

784.2	6.630	-6.45	-5.46	2.25	0.000	-3.82	-2.80
-34.9	10.89	-0.46	5.318	4.209	3.641	3.270	-2.62
0.726	10.11	12.68	-0.72	2.956	-5.22	-0.91	3.229
-19.8	2.843	21.02	7.299	-5.43	-7.31	7.198	2.446
9.000	-2.50	20.10	1.802	-8.00	5.061	-0.48	1.331
-21.7	-3.61	8.453	-1.76	-5.13	-3.66	0.128	0.899
2.405	37.31	-0.16	3.307	2.564	0.964	2.072	-3.61
40.72	-3.57	-18.2	-1.70	3.892	0.977	2.199	-0.53

c)

784	7	-6	-5	2	0	0	0
-35	11	0	5	0	0	0	0
1	10	13	0	0	0	0	0
-20	3	0	0	0	0	0	0
9	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

d)

784	0	0	0	0	0	0	0
-35	11	0	0	0	0	0	0
0	10	13	0	0	0	0	0
-20	0	21	0	0	0	0	0
0	0	20	0	0	0	0	0
-22	0	0	0	0	0	0	0
0	37	0	0	0	0	0	0
41	0	-18	0	0	0	0	0

e)

99	96	93	90	88	86	85	85
97	96	94	93	92	90	89	89
96	97	98	100	99	97	95	94
95	97	101	104	104	101	99	98
93	96	100	103	102	99	97	97
93	96	99	100	98	96	96	97
100	102	103	103	100	99	101	103
109	109	109	108	106	105	108	112

f)

106	99	89	81	77	80	85	89
87	89	91	93	95	97	97	98
100	106	114	118	114	103	90	82
85	85	85	86	88	91	94	96
100	102	104	107	110	112	113	113
102	99	94	90	88	88	89	91
90	93	99	104	108	110	109	109
114	112	108	106	104	105	106	107

Rys. 11.9. Fragment obrazu - blok 8×8 - (a) poddany transformacji DCT (b) oraz kwantyzacji w dwu schematach: ze strefową (c) oraz progową (d) selekcją próbek. Przyjęty kształt strefy zaznaczono kolorem szarym, a wartość progę wynosi 10. Niosące znaczącą informację wartości sześciu współczynników: 21, 20, -22, 37, 41 i -18, pominięte w strefowej selekcji próbek, zostały zachowane w przypadku progowania. Rekonstrukcja bloku obrazu na podstawie skwantowanych wartości po selekcji strefowej i progowej została przedstawiona odpowiednio jako e) i f).

Przy strefowej selekcji próbek stosowane są często różne techniki definiowania strefy, w tym także metody przydziału bitów. W zależności od wiedzy dostępnej a priori na temat użytecznego widma kompresowanych obrazów lub też utartego schematu, np. zachowujemy wartości dziesięciu harmonicznych o najniższych częstotliwościach czy też dajemy dużo bitów dla reprezentacji współczynników niskich częstotliwości i bardzo mało dla współczynników częstotliwości wyższych, określany jest statyczny obszar współczynników istotnych, taki sam dla każdego bloku obrazu czy też danej klasy bloków. Strefa może być określona także metodą estymacji wariancji poszczególnych współczynników transformaty w bloku, którym przydzielana jest konkretna ilość bitów, przy czym zwykle w obszarze wyższych częstotliwości przydział jest zerowy. Oznacza to, że globalna miara statystyczna, jaką jest wariancja, czy też przyjęty a priori schemat decyduje o eliminacji pewnej grupy współczynników. Może wówczas wystąpić sytuacja, gdy w małej liczbie bloków z istotnych obszarów obrazu o ostrych, wyraźnych krawędziach pojawią się znaczące współczynniki wysokoczęstotliwościowe, niosące ważną informację o kształcie krawędzi. Zostaną one jednak zgubione (jak w przykładzie z rys. 11.9), gdyż znajdą się poza przyjętą strefą współczynników istotnych. Bardziej adaptacyjne metody, np. wprowadzające ruchomą strefę o nieregularnych kształtach, śledzącą lokalne własności obrazu wymagają jednak dużego nakładu informacji, którą trzeba dodatkowo zakodować.

Prosty sposób wyznaczenia strefy polega na zakreśleniu obszaru znaczącego w dziedzinie transformaty bloku. Wielkość tej strefy jest zależna od planowanego stopnia kompresji. Kształt strefy jest natomiast zwykle regularny i współczynniki ze strefy są kodowane kolejno według ustalonego porządku. Nie potrzeba w tym przypadku żadnej dodatkowej informacji o położeniu. Kwantyzacja jest zwykle równomierna, a liczbę bitów do zapisu wartości kolejnych współczynników ze strefy (i związaną z nią liczbę poziomów kwantyzacji) można przyjąć na stałe według dostępnej wiedzy a priori. Wykorzystuje się także adaptacyjny przydział różnej ilości bitów współczynnikom w zachowywanej strefie (przy stałej całkowitej liczbie bitów), minimalizując błąd kwantyzacji. Jest to rodzaj adaptacji wprzód, który wymaga dwukrotnego przeglądania danych oraz przesłania dodatkowej informacji do kodera. Najpierw wyznaczana jest średnia wariancja dla każdej pozycji współczynnika w strefie, a potem przydzielana jest liczba bitów do zapisu kwantowanych wartości proporcjonalnie do logarytmu wariancji każdego współczynnika. Macierz przydziału bitów musi być dostępna w procesie dekompresji.

W niektórych modelach kwantyzacji ze strefową selekcją próbek stosuje się nierównomierny, optymalny w sensie minimalnego średniokwadratowego błędu kwantyzacji kwantyzator Lloyda-Maxa, projektowany dla każdego współczynnika, przy czym zerowy współczynnik (składowa stała) jest modelowany najczęściej rozkładem Rayleigh'a, a pozostałe współczynniki rozkładem Gaussa lub Laplace'a. Liczba poziomów kwantyzacji jest ustalana proporcjonalnie do logarytmu (o podstawie 2) wariancji tych rozkładów.

Wszystkie te metody, wykorzystujące nawet najbardziej złożone modele statystyczne rozkładu wartości współczynników w blokach, mają podstawową wadę, którą jest mała adaptacyjność do istotnych szczegółów. Wszelkie próby przydziału bitów, w zależności od globalnej czy lokalnej statystyki, wprowadzają bowiem mechanizm strefowej selekcji próbek, eliminując rzadko występujące, wysokoczęstotliwościowe składowe obrazu.

Drugim schematem kwantyzacji jest progowa selekcja próbek (ang. threshold sample selection). Pozwala ona na wyeliminowanie największej wady strefowego schematu selekcji, przejawiającej się w możliwości pominięcia wartości współczynnika zawierającego znaczącą energię, który znalazł się poza zachowywaną strefą, co może powodować duży lokalny błąd rekonstrukcji. W technice tej dobiera się odpowiedni poziom progę, sterując uzyskiwanym

stopniem kompresji, powyżej którego współczynniki podlegają kwantyzacji i kodowaniu. Ponieważ współczynniki te mogą wystąpić w dowolnym miejscu w bloku, pojawia się konieczność pamiętania pozycji współczynników, które przekraczają wartość progu w każdym bloku. Do zakodowania pozycji używa się najczęściej technik kodowania długości sekwencji. W schemacie z progową selekcją stosuje się zarówno równomierne jak i nierównomierne modele kwantyzacji, przy czym kwantyzacja równomierna uzupełniona jest technikami entropijnego kodowania.

Struktura kwantyzatora z selekcją progową może być skonstruowana adaptacyjnie, gdzie zmienna wartość progu jest funkcją położenia poszczególnych współczynników w bloku. Tablica, określająca wartość progu dla każdego współczynnika w bloku może być stosowana dla wszystkich bloków w obrazie, przy czym może być całościowo przeskalowywana w celu osiągnięcia różnych stopni kompresji, z jednoczesną zmianą błędu kwantyzacji i poziomu zniekształceń w rekonstruowanym obrazie. Wartości progów występujące w takiej tablicy kwantyzacji winny być związane z funkcją czułości kontrastu HVS, minimalizując zmiany w percepcji obrazu, bądź też inne (np. zmiany wiarygodności diagnostycznej w przypadku obrazów medycznych), które są wprowadzane przez błąd kwantyzacji poszczególnych współczynników.

Jedno z najbardziej skutecznych rozwiązań realizujących kwantyzację w metodzie transformacyjnego kodowania w oparciu o technikę selekcji próbek przyjęto w standardzie JPEG. Stosowana jest tam koncepcja równomiernej kwantyzacji połączona z progową selekcją próbek na bazie tablicy kwantyzacji. Tablica ta dla każdego współczynnika bloku $c_{DCT}(u, v) = c(u, v)$, gdzie (u, v) , $u, v = 0, \dots, 7$ to współrzędne w bloku, określa wartość progu, jak również przedział kwantyzacji. Sama kwantyzacja opisana jest przez następujące wyrażenie:

$$\hat{c}(u, v) = \text{round} \left\{ \frac{c(u, v)}{z(u, v)} \right\}, \quad (11.74)$$

gdzie $z(u, v)$ oznacza element tablicy kwantyzacji (liczba całkowita dodatnia), a 'round' jest operatorem przybliżającym do najbliższej liczby całkowitej.

Oznacza to, że dla przykładowej wartości z tablicy $z(u, v)$ równej Δ (jest to szerokość przedziału kwantyzacji dla współczynników o współrzędnych (u, v) - zobacz rysunek 11.10.a) skwantowana wartość współczynnika $\hat{c}(u, v)$ wynosi κ wtedy i tylko wtedy, jeśli

$$\kappa\Delta - \frac{\Delta}{2} \leq c(u, v) < \kappa\Delta + \frac{\Delta}{2}. \quad (11.75)$$

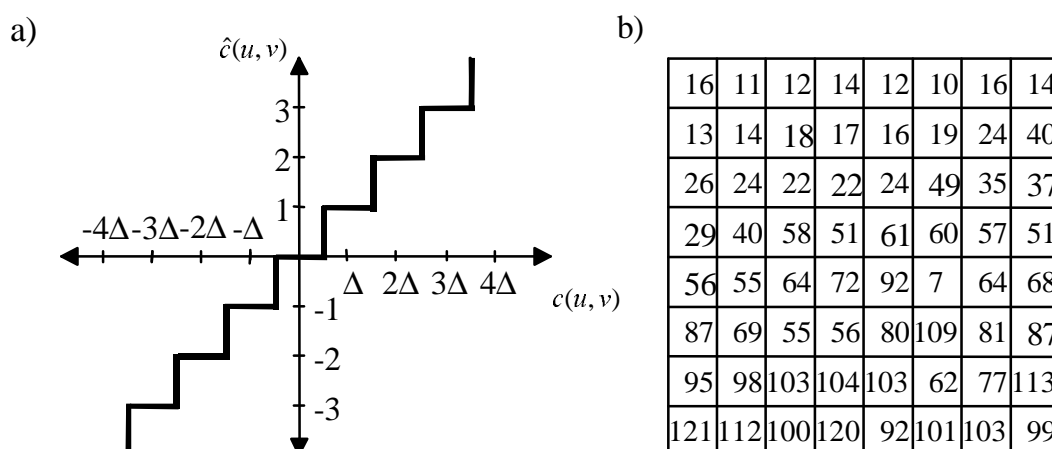
Jeśli $z(u, v) > 2/c(u, v)$, wówczas $\hat{c}(u, v) = 0$, czyli następuje eliminacja tego współczynnika ze zbioru współczynników znaczących (niosących informację w przestrzeni DCT). Wobec tego wartość progu w selekcji próbek wynosi $p = \frac{z(u, v)}{2}$. Przykładowa postać tablicy kwantyzacji została przedstawiona na rys. 11.10 b).

W celu osiągnięcia różnych stopni kompresji tablica ta jest modyfikowana w następujący sposób:

$$z'(u, v) = \text{round}\{z(u, v) \cdot k\}, \quad u, v = 0 \dots 7 \quad (11.76)$$

gdzie $z'(u, v)$ jest nową wartością w tablicy zastępującą $z(u, v)$, a dodatnia liczba rzeczywista k jest stałą dobieraną przez użytkownika.

Zastosowanie równomiernej kwantyzacji współczynników DCT, których rozkład można w przybliżeniu uważać za Gaussowski, daje, według Wintza i Kurtenbacha [17], podobny poziom zniekształceń jak optymalne metody kwantyzacji nierównomiernej.



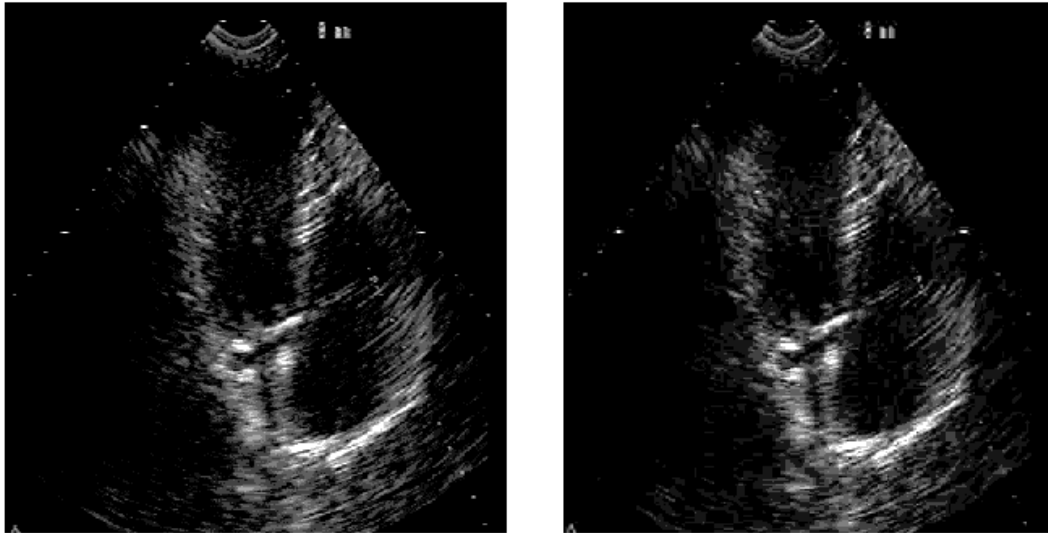
Rys. 11.10. a) krzywa równomiernej kwantyzacji z selekcją progową; b) tablica kwantyzacji optymalizująca jakość rekonstruowanych obrazów pod kątem oceny psychowizualnej (zamieszczona w normie standardu JPEG, jednak nie obligatoryjna).

Powyższy schemat kwantyzacji pozwala więc przy niewielkim błędzie kwantyzacji zachować współczynniki o wyższych częstotliwościach, a ponadto, poprzez dobranie odpowiednich wartości tablicy kwantyzacji, zgodnie z perceptualną wagą poszczególnych współczynników, wysoką psychowizualną jakość obrazu. Dodatkową zaletą w stosunku do optymalnej kwantyzacji Lloyda-Maxa jest zmniejszenie efektów blokowych poprzez zastosowanie tego samego przedziału kwantyzacji dla składowej stałej we wszystkich blokach obrazu.

Redukcja efektów blokowych

Do podstawowych wad metod blokowych należą efekty blokowe występujące przy wyższych stopniach kompresji. Są to artefakty związane z nieciągłością funkcji jasności na granicach pomiędzy poszczególnymi blokami, powstające wskutek silnej kwantyzacji przeprowadzanej niezależnie w każdym bloku. Przykład artefaktów blokowych przedstawia rys. 11.11. Efekty te występują szczególnie intensywnie przy silnej kwantyzacji składowej stałej. Inną wadą to ewentualna nadmiarowość (blokowa) związana z korelacją pomiędzy wartościami odpowiednich współczynników transformaty w sąsiednich blokach. Przy pełnokadrowej transformacji korelacja pomiędzy wartościami współczynników całego obrazu jest minimalna. Poniżej przedstawiono kilka przykładowych rozwiązań redukcji efektów blokowych.

Nieco zmodyfikowany algorytm kompresji wykorzystującej pełnokadrową DCT podał Chan i in. [13]. Kompresując obrazy CT ($512 \times 512 \times 12b$) i MR ($256 \times 256 \times 8b$), wprowadził on adaptacyjną dwuwymiarową tablicę przydziału bitów, powodując wierniejsze odtwarzanie ostrych krawędzi w obrazie. Proponowano też inne metody przy kompresji obrazów CT głowy [9], w których występuje duża liczba ostrych krawędzi pomiędzy czaszką, tłem i tkanką mózgu, polegające na umieszczeniu czaszki i tkanki mózgowej w dwu oddzielnych zbiorach danych i kompresji obrazu czaszki przy pomocy RLE, a tkanki mózgowej przy pomocy DCT. Pozwala to uzyskać diagnostycznie lepszą jakość zrekonstruowanych obrazów.



Rys. 11.11. Oryginalny obraz USG (z lewej) oraz obraz zrekonstruowany po silnej stratnej kompresji (CR=32) z widocznymi efektami blokowymi.

Ishigaki [18] zaproponował zmniejszenie efektów blokowych poprzez zastosowanie adaptacyjnej kwantyzacji współczynników DCT, usuwającej w poszczególnych blokach współczynniki o wartości poniżej pewnego progu, zakładając że rozkład szumów jest równomierny w dziedzinie częstotliwości. Otrzymane przez nich wyniki przy kompresji obrazów CT o parametrach $2000 \times 2510 \times 10b$ potwierdziły te przypuszczenia - przy obserwacyjnej ocenie jakości rekonstruowanych obrazów (ROC, subiektywna klasyfikacja) uzyskano wyższe dopuszczalne stopnie kompresji niż w konwencjonalnej metodzie blokowej, jednak przy 2 dB spadku wartości SNR.

Porównanie metody blokowej (8×8) i pełnokadrowej przez Lee [19] wykazało średnio o 13 dB większą wartość SNR przy zastosowaniu techniki blokowej (kompresowano obrazy CT głowy). Modelowano rozkład współczynników DCT w blokach za pomocą znormalizowanych funkcji rozkładu (Gausa, Laplace'a) i optymalizowano kwantyzację współczynników, przy czym kodowano następnie współczynniki DC przy pomocy DPCM, a współczynniki AC metodą Huffmana. Natomiast Ahn i in. [20] porównywali algorytm Lo i Huanga [8] ze standardem JPEG (jego wersją progresywną). Wyniki także potwierdziły wyższość metod blokowych, zarówno pod względem osiąganych wartości SNR, jak też w ocenie wizualnej, w zastosowaniu do kompresji obrazów MR 8-mio i 12-to bitowych.

MacMahon [21] zaproponował technikę kompresji wykorzystującą blokową DCT o nieco większym rozmiarze bloków - 16×16 . Dla każdego bloku ustalany jest rozmiar przedziału równomiernej kwantyzacji oraz częstotliwość odcięcia (powyżej której wszystkie współczynniki są usuwane) na podstawie histogramu oraz przestrzennego rozkładu współczynników. Pozwala to zredukować artefakty (efekty blokowe) oraz zmniejszyć degradację obrazu w istotnych jego fragmentach przy wyższych stopniach kompresji. Przy kompresji cyfrowych radiogramów płuc $2048 \times 2048 \times 10b$ uzyskano dopuszczalny stopień kompresji 25:1 według oceny ROC, do której wykorzystano opinie 12 radiologów.

Reeve i Lim [22] zaproponowali dwie metody ograniczenia artefaktów. Pierwsza polega na zastosowaniu zachodzących na siebie bloków, tak że piksele należące do regionów pokrywanych przez dwa lub więcej bloki są kodowane, a następnie dekodowane w każdym z tych bloków, po czym zrekonstruowane ich wartości są średnią obliczaną z dekodowanych wartości. Pozwala to w znacznym stopniu zmniejszyć efekty blokowe, jednak kosztem

osiąganych stopni kompresji (przy blokach 16×16 długość kodu wydłuża się o 13%, a dla mniejszych bloków jeszcze bardziej). Drugie rozwiązanie zakłada normalną strukturę nie zachodzących na siebie bloków i niezmienną formę kompresji/dekompresji, po której następuje dolnoprzepustowa filtracja wygładzająca obraz wzdłuż granic bloków. Powoduje to jednak często niepożądaną redukcję wysokoczęstotliwościowej użytecznej informacji zawartej w obrazie, a poza tym powoduje jedynie rozmycie ostrej struktury artefaktów, nie likwidując jej.

Zamiast pokrywających się bloków, niektóre rozwiązania proponują nakładanie bazowych funkcji transformaty, t.j. technika wykorzystująca FFT krótkiej przestrzeni (ang. short space) [23], czy też zachodzące transformaty ortogonalne (ang. lapped orthogonal transform) [24]. Wprowadzają one jednak silne efekty pierścieniowe.

Nakagawa i Miyahara [25] zaproponowali uogólnioną koncepcję KLT jako miarę zniekształceń w zależności od położenia piksela w bloku. Umożliwia to zaprojektowanie transformaty, która pozwala uzyskać mały błąd na granicy bloków kosztem pikseli leżących w ich centrum. Symulacja tej metody wykazuje większy średni poziom błędu rekonstrukcji w stosunku do tradycyjnej KLT, jednak psychowizualny odbiór obrazów ulega poprawie (występuje większa ciągłość funkcji jasności na granicach bloków).

Meiri i Yudilevich, oraz Jain i Farrelle przedstawili szereg prac, w których zasadniczy schemat redukcji efektów blokowych polega na zamodelowaniu obrazu przy pomocy stochastycznych cząstkowych równań różniczkowych lub różnicowych (SPDE) w formie

$$Lf = \varepsilon, \quad (11.77)$$

gdzie L jest liniowym operatorem różniczkowym lub różnicowym, f jest ograniczonym segmentem obrazu, a ε jest błędem stochastycznym, który decyduje o modelu. Wartości pikseli każdego bloku przedstawiane są jako suma dwu źródeł. Pierwsze jest predykcją lub inaczej graniczną odpowiedzią (ang. boundary response), $f_b(n)$, która jest rozwiązaniem równania $Lf = 0$ dla wartości granicznych bloku. Drugie jest błędem predykcji lub resztą $e(n)$ taką, że

$$e(n) = f(n) - f_b(n), \quad (11.78)$$

i jest to składnik wartości piksela, która wynika ze stochastycznej natury modelu oraz z elementów deterministycznych jak np. krawędzie, które nie są ujęte w modelu. Następnie graniczne odpowiedzi są kodowane bezstratnie przy użyciu np. techniki PCM, a reszta - za pomocą transformacji sinusowej, która w tym zastosowaniu daje lepsze wyniki niż kosinusowa.

Dekompozycja wartości pikseli w bloku na dwa źródła (f_b, e) redukuje efekty blokowe z kilku zasadniczych powodów. Jeśli dzielimy wartości graniczne, najczęściej o szerokości jednego piksela, pomiędzy dwa sąsiednie bloki, wtedy graniczna odpowiedź, która interpoluje te wartości, będzie ciągła przy przejściu z jednego bloku do drugiego, bez względu na poziom zniekształcenia na granicach. Ponadto o predykcja zaprojektowana jest tak, że reszty z sąsiednich bloków są nieskorelowane, co oznacza że nie oczekujemy ciągłości tego składnika na granicy. Bloki są więc także fizycznie separowane przez graniczne piksele, co eliminuje blokowe artefakty przy kompresji reszt nawet ze znacznym zniekształceniem.

Poszczególne techniki różnią się przede wszystkim sposobem rozwiązywania równań PDE. Meiri i Yudilevich [26] w swojej metodzie PSTC (ang. pinned sine transform coder) stosują ciągłe PDE, a następnie aproksymują ciągłe rozwiązanie granicznej odpowiedzi, podczas gdy Farrelle w RBC (ang. recursive block coding) proponuje model dyskretny [4]. Jakkolwiek rezultaty osiągnięte przy pomocy techniki RBC wykazują skuteczną eliminację

efektów blokowych, związane jest to jednak często z silniejszym zamazywaniem ostrych krawędzi obrazu.

Zhang [27] oraz Bassiouni [12] przedstawili inną metodę modyfikacji blokowej DCT w celu minimalizacji efektów blokowych, zwaną łączonym kodowaniem transformacyjnym (ang. combined-transform coding-CTC). Polega ona na rozdzieleniu wartości pikseli na dwa zbiory: bardziej i mniej znaczących bitów, a następnie oddzielnej kompresji każdego z tych zbiorów. Bardziej znaczące bity o znacznie większej korelacji są kompresowane za pomocą technik odwracalnych (RLE + kodowanie Huffmana lub LZW) bez dzielenia pikseli na bloki.

Zbiór mniej znaczących bitów jest kompresowany z wykorzystaniem blokowej DCT, przy czym korelacja pomiędzy wartościami współczynników w sąsiednich blokach jest zmniejszona. Wyniki porównań z klasyczną metodą blokowej DCT przedstawiono w przykładzie 11.4.

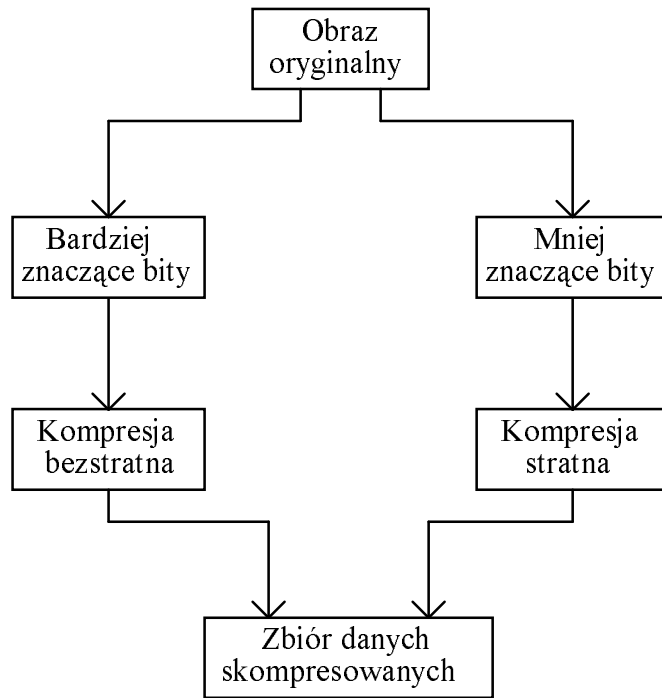
Łączone kodowanie transformacyjne

Algorytm ten jest pewnym uogólnieniem metody sukcesywnej aproksymacji występującej w rozszerzonej, progresywnej wersji standardu JPEG. Rozdzielenie wartości danych na bardziej i mniej znaczące bity następuje o dwa etapy wcześniej, a więc nie po transformacji i kwantyzacji (jak w JPEG), ale już na samym początku procesu kompresji.

Poprzez efektywne bezstratne zakodowanie silniej skorelowanego zbioru bardziej znaczących bitów wartości pikseli, bez podziału na bloki, można uzyskać duży stopień kompresji i silną redukcję przestrzennej nadmiarowości występującą w tym zbiorze. Zastosowanie blokowej DCT do kompresji słabo skorelowanego zbioru mniej znaczących bitów pozwala z kolei na efektywną, nieodwracalną redukcję części informacji z obrazu. Technika ta może być łatwo implementowana w aplikacjach medycznych, ze względu na prostą możliwość dostosowania algorytmu do danych 8-mio, 12-to czy 16-to bitowych.

Ogólny schemat algorytmu CTC wykorzystujący efektywną metodę kompresji bezstratnej oraz blokową DCT przedstawiony jest na rys. 11.12. Proces dekompresji przebiega dokładnie odwrotnie.

Minimalizację efektów blokowych uzyskuje się poprzez wierne odtworzenie bardziej znaczących bitów wartości pikseli całego obrazu, a więc także na granicach bloków, podczas gdy zniekształcenia związane są jedynie ze stratną kompresją młodszych bitów. Ponieważ starsze bity w większości systemów prezentacji decydują o psychowizualnej jakości obrazów, artefakty stają się mniej widoczne.



Rys. 11.12. Ogólny schemat algorytmu kompresji łączonej CTC.

Proste wyjaśnienie mechanizmu redukcji blokowych artefaktów w technice CTC w stosunku do konwencjonalnych metod transformacyjnych przedstawia się następująco. Ponieważ artefakty te związane są z blokową transformacją i niezależną kwantyzacją współczynników transformaty, są one proporcjonalne do błędu kwantyzacji. Niech f oznacza 2-D macierz obrazu oryginalnego, natomiast przetransformowaną macierz obrazu oznaczmy przez F , a przez φ - macierz transformaty unitarnej. Wobec tego mamy:

$$F = \varphi f \varphi^T, \quad (11.79)$$

a średniokwadratowy błąd kwantyzacji jest równy:

$$\sigma_F^2 = E\{(F - Q[F])^2\}. \quad (11.80)$$

W technice CTC oryginalny obraz f jest dzielony na dwie macierze: mniej znaczących bitów, oznaczoną przez f' , i bardziej znaczących bitów f'' . Macierz f' jest przekształcana przy pomocy blokowej transformaty, podczas gdy f'' jest kodowana bezstratnie. Stąd też błąd kwantyzacji związany jest jedynie z przekształceniem pierwszej macierzy. Tak więc, jeśli przez σ_1^2 i σ_2^2 oznaczmy błędy kwantyzacji metod: konwencjonalnej i CTC, wówczas mamy:

$$\sigma_2^2 = \sigma_{\varphi f' \varphi^T}^2 \leq \sigma_{\varphi f' \varphi^T + \varphi f'' \varphi^T}^2 = \sigma_{\varphi f \varphi^T}^2 = \sigma_1^2, \quad (11.81)$$

co należało wykazać.

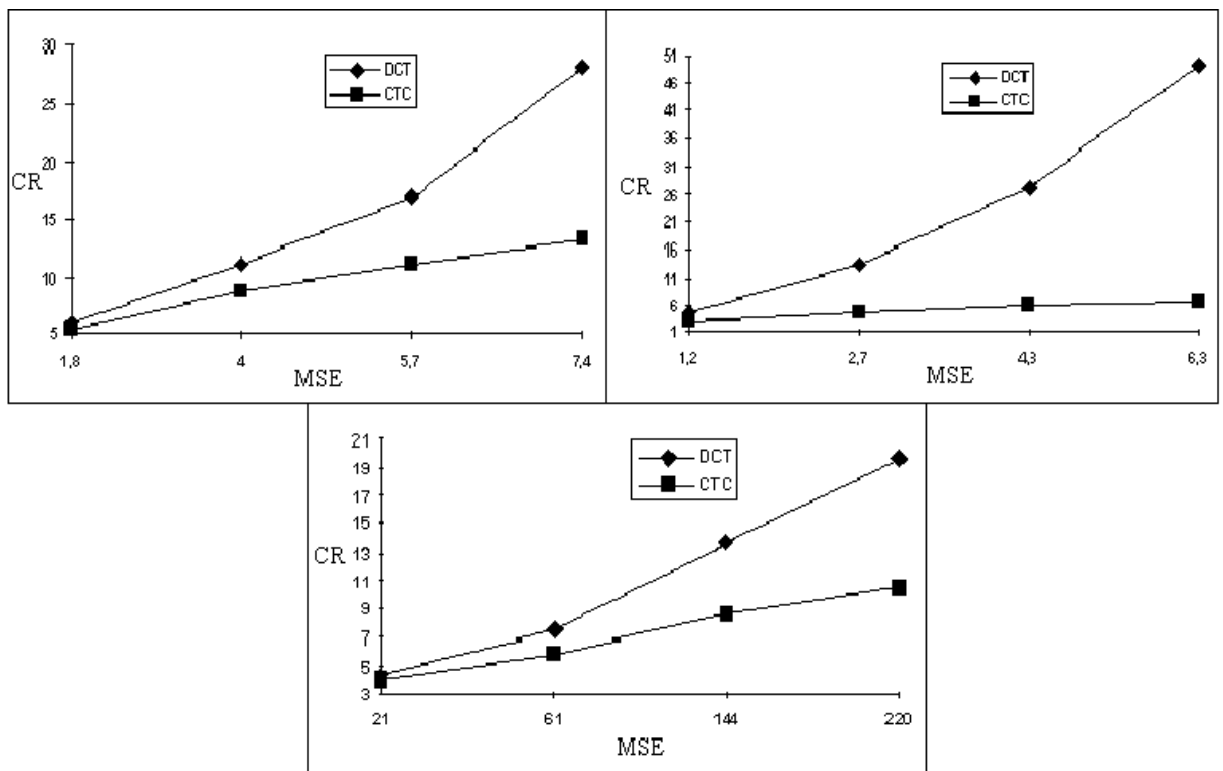
Zmniejszenie efektów blokowych odbywa się kosztem mniejszej skuteczności kompresji (wskutek bezstratnej kompresji starszych bitów). Stąd też, aby uzyskać taki sam stopień kompresji jak w metodzie konwencjonalnej, przeprowadza się silniejszą redukcję mniej znaczących bitów, co często powoduje wspomniane efekty zniekształceń szczegółów obrazu oraz podkreślenia krawędzi.

Technika CTC może być szczególnie efektywna w kompresji obrazów, w których występują stosunkowo niewielkie obszary istotne diagnostycznie o dużych wartościach

pikseli wobec dużych obszarów mniej użytecznych o małych poziomach wartości pikseli (najlepiej nie przekraczających progu bardziej znaczących bitów).

PRZYKŁAD 11.4. Porównywano skuteczność kompresji techniki CTC z klasyczną techniką kompresji wszystkich bitów wartości pikseli z pomocą blokowej DCT na grupie obrazów medycznych. Zbiór bardziej znaczących bitów kompresowano przy użyciu algorytmu kodowania arytmetycznego.

Wstępne wyniki wykazały większą efektywność podziału wartości poszczególnych pikseli 2-6 (dwa bardziej znaczące bity kompresowane są bezstratnie, a sześć mniej znaczących bitów stratnie z wykorzystaniem DCT) w stosunku do podziału 4-4, stąd też taki model podziału w technice CTC został zastosowany (dla obrazów 8-mio bitowych). Dla 11-to bitowych danych obrazów MR zastosowano podział 3-8. Prezentowane wykresy są średnią wyników uzyskanych z kompresji grupy 4 obrazów każdego rodzaju. Wyniki oceny średniokwadratowej przedstawiono na rys. 11.13. W ocenie subiektywnej dla różnych stopni kompresji specjaliści (zarówno lekarze jak i inżynierowie) byli zgodni co do wyższej jakości obrazów kompresowanych techniką DCT.



Rys. 11.13. Badanie efektywności techniki CTC - obrazy scyntygraficzne (górn-lewo), obrazy rentgenowskie (górn-prawo), obrazy MR (dół).

Kodowanie współczynników transformacji

Odwracalne metody kompresji kwantowanych współczynników transformacji mogą być różnorodne, a o wielu z nich wspomniano w rozdziałach dotyczących kompresji bezstratnej oraz w poprzednich punktach tego rozdziału. Podsumowując te rozważania należy zwrócić uwagę na kilka elementów.

Po pierwsze algorytm kodowania powinien być dopasowany do struktury strumienia wyjściowego danych z kwantyzatora. Dobrym przykładem jest tutaj znormalizowany koder Huffmana ze standardu JPEG, opisany w punkcie 3.4. Stosując koder arytmetyczny należy odpowiednio zamodelować alfabet źródła, skonstruować kontekst modelu statystycznego (rozmiar i kształt) i mechanizmy adaptacji tego modelu do lokalnej statystyki, potrzebna jest więc możliwie pełna informacja o własnościach kodowanego strumienia danych.

Ponadto, stopień złożoności kodera winien korelować ze złożonością całego algorytmu kompresji, w tym przede wszystkim ze sposobem realizacji procesu kwantyzacji współczynników transformaty. W praktyce najczęściej występuje zamienność pomiędzy rozbudowaniem schematu kwantyzacji oraz kodowania. Przy złożonej adaptacyjnej metodzie przydziału bitów, selekcji próbek czy optymalizowanym zbiorze kwantyzatorów równomiernych stosowanie nawet bardzo rozbudowanych koderów entropijnych nie daje spodziewanych efektów, czyniąc algorytm kompresji bardzo czasochłonnym. Mogą one natomiast przynieść dużą poprawę skuteczności kompresji dla prostych kwantyzatorów skalarnych – równomiernych. Rozwiązaniem najbardziej chyba efektywnym jest zastosowanie zarówno kwantyzatora jak i kodera o niewielkim stopniu złożoności, zapewniając dobre ich sprzężenie (jak w standardzie JPEG).

Coraz większą rolę zaczyna odgrywać elastyczność strumienia wyjściowego algorytmu kompresji, tj. łatwa możliwość zmiany charakteru progresji przekazywanej informacji np. w transmisjach sieciowych, odporność na zakłócenia np. w transmisji bezprzewodowej kosztem pewnej nadmiarowości reprezentacji kodowej, zdolność osadzania (zagnieżdżania) informacji w strumieniu, która pozwoli odbierać czy odtwarzać jedynie wystarczającą w danym zastosowaniu porcję informacji bez potrzeby oczekiwania na przesłanie całego zbioru skompresowanych danych, czy też odczytywania z bazy danych całego zarchiwizowanego zbioru.

W pełnej realizacji powyższych własności strumienia danych potrzebna jest odpowiednia technika formatowania tego strumienia w paczki, rozdzielne i np. niezależne strumienie (jak w technikach wielokrotnego opisu – ang. multiple description), hierarchiczne struktury informacji, itd., która umożliwi dopisanie w efektywny sposób informacji o prawach autorskich związanych z przesyłaną informacją, umieszczenie dodatkowej informacji użytkownika w określonej, efektywnej strukturze (np. XML), itp. Zagadnienia te zostały częściowo rozwiązane w czasie realizacji standardów JPEG, MPEG-1, MPEG-2 i MPEG-4, są także przedmiotem optymalizacji nowego standardu JPEG2000, który zachowuje kompatybilność z JPEG, opartym na kodowaniu transformacyjnym.

Bibliografia:

1. N. Ahmed, T. Natarajan, K.R. Rao, *Discrete Cosine Transform*, IEEE Trans. Computers, str. 90-93, January 1974.
2. D.F. Elliott, *Handbook of Digital Signal Processing*, Academic Press, Inc., 1987.
3. C. Loeffler, A. Ligtenberg, G. Moschytz, *Practical Fast 1-D DCT Algorithms with 11 Multiplications*, Proc. Int'l Conf. on Acoustics, Speech, and Signal Processing (ICASSP'89), 988-991, 1989.
4. Farrelle Paul Michael, *Recursive Block Coding for Image Data Compression*, Springer-Verlag New York Inc., 1990.

5. J.D. Villasenor, *Alternatives to the Discrete Cosine Transform for Irreversible Tomographic Image Compression*, IEEE Trans. Medical Imaging, 12(4):803-811, Dec. 1993.
6. P. Yip, K.R. Rao, *Energy packing efficiency for generalized discrete transforms*, IEEE Trans. Commun. COM-26:1257-1262, 1978.
7. H. Kitajima, *Energy Packing Efficiency of Hadamard Transforms*, IEEE Trans. Commun. COM-24:1256-1258, 1976.
8. S.-C. Lo, H.K. Huang, *Radiological Image Compression: Full-Frame Bit-Allocation Technique*, Radiology, 155(3):811-817, June 1985.
9. J.M. Bramble, H.K. Huang, M.D. Murphy, *Image Data Compression*, Investigative Radiology, 23:707-712, Oct. 1988.
10. A.W. Templeton., M.D. Murphey, L.T. Cook, M.A. McFadden, *Clinical Examples of Image Compression Using a Full Frame 2-Dimensional Discrete Cosine Transform*, Symposium for Computer Assisted Radiology S/CAR'92, USA, 1992.
11. H. Wang, D. Rosenfeld, M. Braun, H. Yan, *Compression and Reconstruction of MRI Images using 2D DCT*, Magnetic Resonance Imaging, 10(3):427-432, 1992.
12. M.A. Bassiouni, N.S. Tzannes, M.C. Tzannes, *High-fidelity integrated lossless/ lossy compression and reconstruction of images*, Optical Engineering, 32(8):1848-1853, Aug. 1993.
13. K.K. Chan, S.-L. Lou, H. K. Huang, *Full-Frame Transform Compression of CT and MR Images*, Radiology, 171(3):847-851, June 1989.
14. A. Segall, *Bit allocation and encoding for vector sources*, IEEE Trans. Inform. Theory, IT-22:162-169, March 1976.
15. A. Gersho, R.M. Gray, *Vector quantization and Signal Compression*, Kluwer Academic Publishers, 1992.
16. J.J.Y. Huang, P.M. Schultheiss, *Block Quantization of Correlated Gaussian Random Variables*, IEEE Trans. Commun. Syst., CS-11:289-296, Sept. 1963.
17. P.A. Wintz, A.J. Kurtenbach, *Waveform error control in PCM telemetry*, IEEE Trans. Inform. Theory, IT-14:650-661, Sept. 1968.
18. T. Ishigaki, S. Sakuma, M. Ikeda, Y. Itoh, M. Suzuki, S. Iwai, *Clinical Evaluation of Irreversible Image Compression: Analysis of Chest Imaging with Computed Radiography*, Radiology, 175(3):739-743, June 1990.
19. H. Lee, Y. Kim, A.H. Rowberg, E.A. Riskin, *Statistical Distributions of DCT Coefficients and Their Application to an Interframe Compression Algorithm for 3-D Medical Images*, IEEE Trans. on Medical Imaging, 12(3):478-485, Sep. 1993.
20. C.B. Ahn, I.Y. Kim, S.W. Han, *Medical Image Compression using JPEG Progressive Coding*, Medical Imaging IEEE Conference, 1994.

21. H. MacMahon, K. Doi, S. Sanada, S. Montner, M.L. Giger, Ch.E. Metz, N. Nakamori, F. - F. Yin, X.-W. Xu, H. Yonekawa, H. Takeuchi, *Data Compression: Effect on Diagnostic Accuracy in Digital Chest Radiography*, *Radiology*, 178(1):175 -179, Jan. 1991.
22. H.C. Reeve, J.S. Lim, *Reduction of Blocking Effect in Image Coding*, Proc. ICASSP'83, 1212-1215, 1983.
23. B.L. Hinman, J.G. Bernstein, D.H. Staelin, *Short-space Fourier transform image processing*, Proc. ICASSP'84, San Diego, CA, March 1984.
24. P. M. Cassereau, D.H. Staelin, G.D. Jager, *Encoding of images based on lapped orthogonal transform*", *IEEE Trans. Commun.*, Feb. 1989.
25. M. Nakagawa, M. Miyahara, *Generalized Karhunen-Loeve transformation I (theoretical consideration)*, *IEEE Trans. Commun.*, COM-35:215-223, Feb. 1987.
26. A.Z. Meiri, E.Yudilevich, *A Pinned Sine Transform Image Coder*, *IEEE Trans. Commun.*, COM-29:1728-1735, Dec. 1981.
27. Y.-Q. Zhang, M.H. Loew, R.L. Pichholtz, *A Combined-Transform Coding (CTC) Scheme for Medical Images*, *IEEE Trans. Medical Imaging*, 11(2):196-202, June 1992.

W. D. Ray, and R. M. Driver, *Further decomposition of the Karhunen-Loeve series representation of a stationary random process*, *IEEE Trans. Inf. Theory*, IT-16(6):663-668, 1970.