

## ROZDZIAŁ 13. KODOWANIE PASMOWE

W technice pasmowego kodowania (ang. subband encoding - SBE) obraz jest poddawany filtracji w dziedzinie 2D przy pomocy filtrów dolno- i górnoprzepustowych. W jej efekcie powstaje szereg 2D zbiorów współczynników, z których każdy zawiera dane z określonego zakresu częstotliwości przestrzennych (pasma będące częścią pasma obrazu oryginalnego).

Ponieważ każdy z subpasmowych obrazów ma zredukowaną szerokość pasma w stosunku do obrazu oryginalnego, mogą być one próbkowane z mniejszą rozdzielczością (ang. downsampling). Ta pierwsza faza procesu kompresji składająca się z filtracji i próbkowania nazywana jest analizą. Potem następuje kodowanie każdego z obrazów pasmowych. Można do tego celu użyć tego samego lub różnych dekodery dla każdego pasma, przy czym powinny być one dostosowane do charakterystyki obrazu danego pasma. Stosuje się nie tylko różne stopnie kompresji dla poszczególnych pasm, lecz nawet w pewnych zastosowaniach zupełnie odmienne techniki kompresji.

Proces dekompresji polega na próbkowaniu zwiększającym rozdzielczość rekonstruowanych przy użyciu odpowiednich filtrów obrazów pasmowych, a następnie zsumowaniu tych obrazów w jeden całościowy obraz. Proces ten nazywa się synteza obrazu.

Tworzenie pasmowych obrazów samo w sobie nie wprowadza elementu kompresji. Ze względu jednak na znaczącą redukcję korelacji pomiędzy wartościami pikseli w stosunku do obrazu oryginalnego, obrazy te mogą być znacznie efektywniej kodowane niż obraz oryginalny.

Do kluczowych elementów techniki SBE należy zestaw filtrów stosowanych do analizy i syntezy obrazu, a także algorytmy kodowania obrazów pasmowych.

Przykładowo aby dokonać dekompozycji całego pasma sygnału 1-D na dwa podpasma, zestaw filtrów do analizy powinien zawierać filtr dolnoprzepustowy i górnoprzepustowy z 'nie nakładającymi' się charakterystykami częstotliwościowymi, ciągłymi, ze stałym wzmocnieniem w całym paśmie. Idealne filtry w praktyce są jednak nie realizowalne, stąd też by uniknąć eliminacji pewnych częstotliwości z pasma sygnału podczas analizy, stosuje się filtry o pokrywających się charakterystykach częstotliwościowych, co z kolei prowadzi do zniekształceń utożsamiania (ang. aliasing). By rozwiązać ten problem stosuje się najczęściej lustrzane filtry kwadratowe (ang. quadrature mirror filter - *QMF*), które pozwalają wyeliminować błąd utożsamiania i uniknąć błędów kodowania.

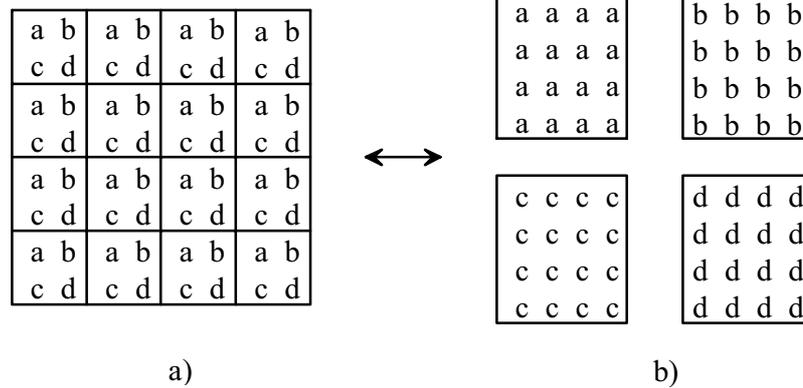
### 13.1. Schemat podstawowy

Dwuwymiarowej filtracji dokonuje się najczęściej stosując filtry 1-D najpierw w jednym, a potem w drugim wymiarze. Przykład analizy i syntezy obrazu przedstawia rys. 13.1.



Blokki po transformacji

Pasmowe obrazy



Rys.13.2. Porównanie transformatorowej i subpasmowej reprezentacji obrazu; a) transformatorowa, b) subpasmowa.

Rysunek 13.2.a) przedstawia współczynniki transformaty (n.p. DCT) od a (najniższa częstotliwość) do d (najwyższa częstotliwość) obliczonej w blokach pikseli  $2 \times 2$  obrazu oryginalnego. Natomiast rysunek 13.2.b) zawiera powstałe w wyniku analizy, przy użyciu filtrów  $2 \times 2$  i próbkowaniu z poziomem 2:1 zmniejszającym dwukrotnie liczbę pikseli w każdym kierunku, obrazy pasmowe z wartościami a, b, c, d (lokującymi się w poszczególnych pod-pasmach).

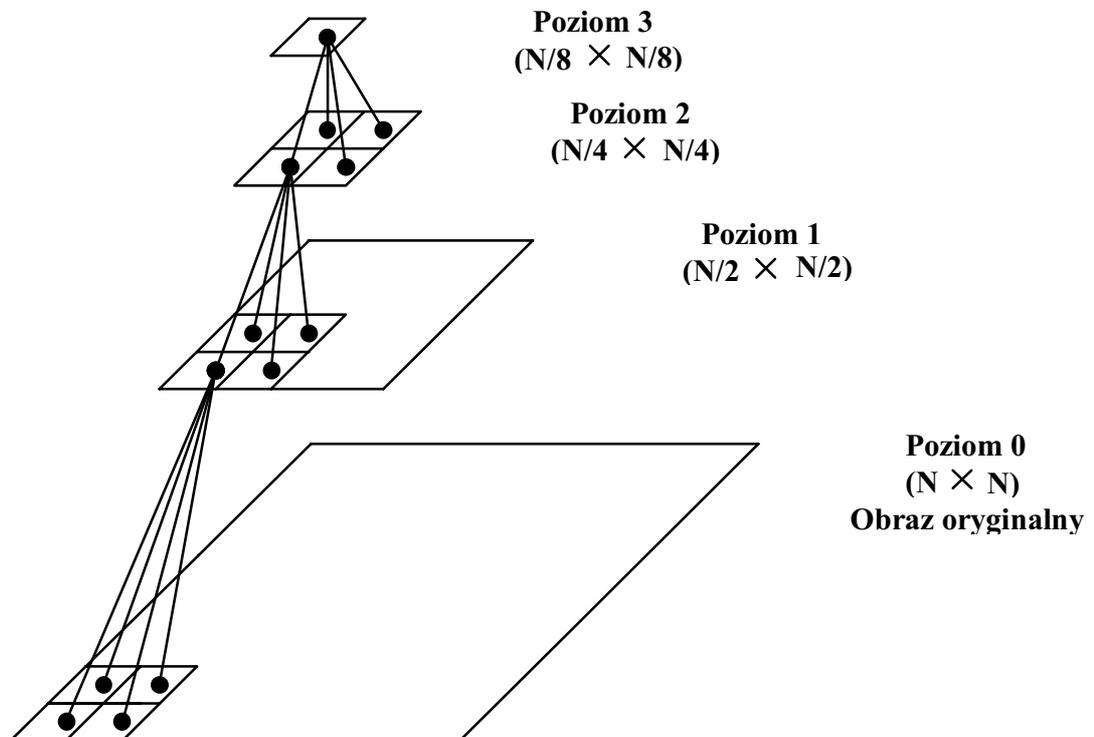
### 13.3. Kodowanie hierarchiczne

Technika hierarchicznego kodowania (ang. hierarchical encoding - HE) polega na takim kodowaniu obrazu, które umożliwia jego zrekonstruowanie na różnym poziomie jakości lub w wersjach o różnej rozdzielczości. Przykładem aplikacji techniki HE są obrazowe bazy danych z możliwością szybkiego przeszukiwania, gdzie użytkownik korzysta z wygodnej opcji przeglądania niskorozdzielczych wersji obrazu w celu odnalezienia poszukiwanego obrazu. Następnie można wysubtelnić ten obraz w sposób progresywny. Ten rodzaj algorytmów hierarchicznych nazywany jest progresywną transmisją. Może być on także stosowany do transmisji obrazów przez kanały o wąskim paśmie w stosunku do liczby danych, np. linie telefoniczne, szczególnie w przypadku, gdy szybkie rozpoznanie obrazu jest konieczne lub gdy czas transmisji jest ograniczony. Przykładowo technika ta znajduje zastosowanie w poligrafii, w wojsku, czy też w telemedycynie. W progresywnej transmisji informacja zawarta w obrazie oryginalnym jest dzielona na części, które są kodowane (transmitowane) według odpowiedniej sekwencji. Następnie następuje kolejne dekodowanie (odbieranie) informacji i rekonstruowanie coraz doskonalszej aproksymacji obrazu oryginalnego. Do podstawowych cech progresywnej transmisji należy zaliczyć:

- uzyskanie dużych stopni kompresji dla wczesnych aproksymacji oryginału;
- maksymalne wykorzystanie poprzednio transmitowanych danych i zminimalizowanie dodatkowej informacji przekazywanej w kolejnych etapach;

- możliwość uzyskania efektywnej kompresji całego obrazu oraz rekonstrukcji z dużą wiernością względem oryginału (jeśli transmitowana jest różnicowa mapa błędów wówczas możliwa jest transmisja bezstratna);
- algorytmy kodowania/dekodowania są względnie szybkie i podatne hardware'owym implementacjom.

Urządzenia wyjściowe zainstalowane w systemie z bazą danych obrazowych mogą wymagać wersji obrazu o różnej rozdzielczości lub jakości. Jest to przykład wieloużytkowego środowiska (ang. multiuse environment), które jest innym rodzajem zastosowań HE. Do tych urządzeń można najczęściej zaliczyć monitor HDTV, wysokorozdzielcza drukarka laserowa, niskorozdzielczy nadajnik do linii telefonicznych itd. Cechy charakterystyczne technik HE stosowanych w środowiskach wieloużytkowych są podobne jak w progresywnej aproksymacji. Można jedynie dodać szybki i skuteczny dostęp do wersji obrazu o danej rozdzielczości czy pożądanym poziomie jakości. Ponadto techniki te winny zapewnić efektywne algorytmy dekompresji ze względu na częste czytanie (dekodowanie) obrazu przy jednokrotnym jego wpisaniu do bazy.



Rys.13.3. Struktura piramidy charakterystyczna dla technik HE wykorzystujących hierarchię o zmiennej rozdzielczości.

Wspólnym aspektem obu zastosowań jest potrzeba konstrukcji hierarchii obrazów, tzn. sposobu organizacji danych obrazowych ze względu na ich ważność. Zasadniczo poszczególne poziomy w hierarchii odnoszą się do obrazu rekonstruowanego z pewną rozdzielczością lub poziomem jakości. Choć hierarchie danych znajdują liczne zastosowanie w przetwarzaniu obrazów poza zagadnieniem kompresji i stosowane są różne wskaźniki do wyznaczenia poziomów, jednak w technikach kompresji do najczęściej stosowanych należą hierarchie o stałej i zmiennej rozdzielczości. W przypadku hierarchii o stałej rozdzielczości

rekonstruowany obraz jest takiego samego rozmiaru jak obraz oryginalny i wartość każdego piksela jest korygowana przy przechodzeniu z jednego poziomu na drugi. Przyrost średniej liczby bitów kodu przypadających na piksel obrazu jest mniej więcej stały, jednak niekoniecznie odpowiada to stałej poprawie jakości rekonstruowanego obrazu. Jest ona szczególnie przydatna w algorytmach progresywnej transmisji. Natomiast techniki wykorzystujące hierarchie o zmiennej rozdzielczości (szczególnie przydatne w wieloużytkowym środowisku), w których obrazy odpowiadające poszczególnym poziomom różnią się przestrzenną rozdzielczością, charakteryzują się eksponentywną zmianą przyrostowej średniej bitowej przy przechodzeniu pomiędzy różnymi poziomami. W technikach tych powstaje charakterystyczna piramidalna struktura, gdzie bazą piramidy jest obraz o pełnej rozdzielczości obrazu oryginalnego. Przykład takiej piramidy przedstawia rysunek 13.3, gdzie zależności pomiędzy wartościami pikseli na kolejnych poziomach tworzą typową strukturę drzewa.

Do typowych technik z hierarchią o stałej rozdzielczości należy:

- kodowanie map bitowych, kiedy to binarna mapa każdego bitu wartości pikseli (o rozmiarach oryginału) może być progresywnie kodowana (transmitowana) w sekwencji, zaczynając od mapy bitu najbardziej znaczącego;
- wektorowa kwantyzacja o strukturze drzewa, gdzie poszczególne etapy progresywnej transmisji mogą zawierać kolejne bity indeksu najlepszego przybliżenia obrazu oryginalnego w książce kodowej. Zakodowanie w ostatnim etapie różnicowej mapy błędów może umożliwić bezstratną rekonstrukcję obrazu;
- hierarchiczne kodowanie z wykorzystaniem transformat ortogonalnych, w którym progresywnie można kodować w każdym etapie kolejne współczynniki transformaty (zaczynając od składowej stałej) lub też kodować jednocześnie wszystkie współczynniki, ale przy ograniczonej liczbie bitów wyrażających ich wartości. Poszczególne rozwiązania obok rodzaju zastosowanej transformaty różnią się także metodą hierarchicznego uporządkowania informacji zawartej w wartościach współczynników.

W technikach z hierarchią o zmiennej rozdzielczości używa się takie metody budowania struktury piramidy jak algorytmy kolejnego próbkowania zmniejszającego rozdzielczość danych reprezentujących obraz oryginalny na danym etapie (ang. subsampling pyramid), zastępowania bloków pikseli (typowo  $2 \times 2$ ) średnią ich wartości (ang. mean pyramid), formowania predykcji z ograniczonej liczby danych i jej odejmowania od obrazu oryginalnego, w celu uzyskania obrazu resztkowego (ang. prediction/residual pyramid) itp.

#### 13.4. Algorytmy waveletowej kompresji obrazów

Transformacja danych oryginalnych do nowej przestrzeni jako pierwszy etap wielu schematów kompresji pozwala silnie zredukować lub nawet usunąć różnego typu nadmiarowości występujące w oryginalnej reprezentacji przetwarzanego zbioru danych. Rodzaj zastosowanej transformaty warunkuje efektywność różnych rozwiązań algorytmów skalarnej lub wektorowej kwantyzacji w dziedzinie częstotliwościowej lub czasowo-częstotliwościowej. Statystyczna charakterystyka uzyskiwanych po kwantyzacji danych służy do wyboru optymalnych metod bezstratnego kodowania współczynników. Minimalna długość nowej reprezentacji danych oryginalnych warunkowana jest dopuszczalnym poziomem zniekształceń dla danego typu danych. Większość efektywnych metod kompresji wykorzystuje bardzo dobre własności silnej dekorrelacji danych przez transformatę kosinusową, zawierając, jako serce tych metod, procedury obliczania szybkiej dyskretnej

transformaty kosinusowej, niezależnie w blokach o stałych rozmiarach, najczęściej  $8 \times 8$  (standard JPEG). Największą wadą tych metod jest nieciągłość funkcji jasności na granicach bloków, znana jako efekt blokowy. Pośród technik redukcji tego zjawiska to najskuteczniejszych można zaliczyć metody obliczające lokalną transformatę kosinusową w zachodzących na siebie blokach, pokrywających z nadmiarem cały obraz. Można także zwiększyć adaptacyjność całego algorytmu poprzez dopasowanie rozmiarów bloków w poszczególnych fragmentach obrazów do ich lokalnej statystyki, stosując optymalne z punktu widzenia efektywności kompresji techniki segmentacji obrazów. Wszystkie te zabiegi pozwalają zwiększyć jedynie nieznacznie efektywność kompresji i to najczęściej tylko dla wybranych klas obrazów.

Nowy etap w rozwoju technik kompresji stanowi wykorzystanie różnego typu transformat wavelet, które otwierając znacznie szersze możliwości budowania modeli kwantyzacji i kodowania pod kątem optymalnej jakości psychowizualnego odbioru rekonstruowanych obrazów, pozwalają znacząco zwiększyć efektywność kompresji w szerokiej gamie aplikacji. Praktycznie nieograniczona ilość możliwych postaci funkcji bazowych transformaty znacznie rozszerza możliwości tworzenia filtrów wykorzystywanych w znanych już od wielu lat podstawowych technikach pasmowego kodowania (ang. subband coding). Transformata wavelet pozwala uzyskać zbliżony do transformaty kosinusowej poziom dekorrelacji danych oryginalnych. Sekwencyjna przestrzenna filtracja danych obrazowych w coraz mniejszej skali nie wprowadza artefaktów blokowych. Dla małych średnich bitowych (dużych stopni kompresji) pojawiają się natomiast zjawiska rozmycia i zniekształceń pierścieniowych. Zasadniczo rozmycie należy do zniekształceń najlepiej akceptowalnych przez obserwatora, natomiast znacznie mniej przyjemnymi zjawiskami są efekty blokowe i pierścieniowe. Lokalne funkcje bazowe przekształceń falkowych, definiowane jako niezerowe na skończonym nośniku (o niewielkich najczęściej rozmiarach w stosunku do długości zbioru), mogą być konstruowane w zależności od lokalnych własności kompresowanych zbiorów danych. Funkcjom tym odpowiada określona postać filtrów o skończonej długości (skończonej odpowiedzi impulsowej). Tworząc algorytmy kompresji oparte na przekształceniu wavelet można dokonywać selekcji optymalnych banków filtrów poprzez stosowanie metod wyboru najlepszej bazy w pakiecie waveletów (ang. best basis wavelet packets), budować złożone adaptacyjne techniki przestrzenno-częstotliwościowej kwantyzacji oraz kontekstowego binarnego kodowania. Cechy waveletowej transformaty pozwalają uzyskać lepszą reprezentację rzeczywistych niestacjonarnych sygnałów. Z ich pomocą można skuteczniej wykorzystać dostępną a priori i a posteriori wiedzę o przetwarzanym zbiorze danych do zachowania wszystkich istotnych z punktu widzenia konkretnych zastosowań cech obrazu, np. można lepiej rekonstruować diagnostycznie ważne cechy obrazów medycznych. Funkcje waveletowe o ograniczonym nośniku (ang. compact support) są efektywne w procesie kompresji, a długość odpowiadających im filtrom jest proporcjonalna do gładkości i regularności waveletów.

Algorytmy, które pozwalają uzyskać największą skuteczność kompresji obrazów wykorzystują wielorozdzielczą analizę obrazów, dokonywaną przy pomocy czasowo-częstotliwościowej transformaty falkowej (waveletowej). Czasowo-częstotliwościowa dekompozycja obrazu pozwala zachować informację przestrzenną o rozkładzie krawędzi w obrazie oryginalnym w kolejnych wersjach tegoż obrazu o malejącej skali, co daje nadmiarowość łatwą do wyeliminowania w algorytmie kodowania. Jednocześnie sama informacja o położeniu krawędzi jest przy doborze właściwej bazy falkowej silnie upakowana w miejscu występowania krawędzi. W wersjach obrazu o większej skali (pierwsze etapy dekompozycji) dokładniejszej lokalizacji w przestrzeni towarzyszy mniej dokładna informacja częstotliwościowa. Z kolei przy małej skali miejsce wystąpienia krawędzi jest

określone bardzo zgrubnie natomiast analiza częstotliwościowa sygnału oryginalnego jest znacznie dokładniejsza.

Funkcje bazowe przekształcenia wavelet, świetnie nadające się do reprezentacji sygnałów niestacjonarnych, winny dobrze oddawać charakter (być do nich podobne) kompresowanych danych, najczęściej gładkich krawędzi o zróżnicowanym, ale niezbyt wielkim gradiencie. Powinny więc być to funkcje gładkie o skończonym, najlepiej niewielkim w stosunku do rozmiarów obrazu nośniku. Stopień gładkości filtru odpowiadającego funkcji falkowej rośnie wraz z długością nośnika przy pogarszającej się jednocześnie przestrzennej lokalizacji sygnału w dziedzinie transformaty. Potęgują się też niekorzystne zjawiska na granicy określoności sygnału (na krawędzi obrazu), szczególnie niewygodne przy wielokrotnej filtracji w analizie wielorozdzielczej. Konstrukcja bowiem pary filtrów ortogonalnych o odpowiedzi impulsowej skorelowanej z kształtem falki i funkcji skalującej jest możliwa jedynie przy nieliniowej fazie dla filtrów SOI, dając dodatkowe rozkłady znaczących współczynników na granicach obszaru określoności nie zawierające żadnej istotnej informacji o obrazie, a znacznie ograniczające kompresję. Powoduje to na granicach filtrowanych zbiorów danych o skończonej długości. W celu zminimalizowania tego zjawiska wprowadza się na granicach cykliczne rozszerzenia zbiorów danych. Rezygnacja z warunku ortogonalności umożliwi budowę filtrów biortogonalnych czy nadmiarowych o liniowej fazie i braku zniekształceń na granicach pasm, ale odbywa się to oczywiście kosztem słabszej dekorelacji danych w przekształceniu falkowym. Pary dolno- i górnoprzepustowych filtrów realizujących przekształcenie wavelet stanowią najczęściej parę QMF (ang. quadrature mirror filters).

### 13.5. Wielorozdzielcza analiza obrazu w dziedzinie wavelet

Waveletami są funkcje generowane z jednej funkcji matki  $\psi$  poprzez elementarne operacje rozciągania (zmiana skali czasu) i przesuwania

$$\psi^{s,x}(t) = \frac{1}{\sqrt{|s|}} \psi\left(\frac{t-x}{s}\right), \quad s \neq 0. \quad (13.1)$$

Wavelet matka  $\psi$  musi spełniać równanie  $\int \psi(t) dt = 0$ , które wymusza co najmniej kilka oscylacji. Warunek na funkcję matkę może być też formułowany inaczej:

$$\int \frac{|\Psi(\omega)|^2}{\omega} d\omega < \infty, \quad (13.2)$$

gdzie  $\Psi$  jest transformatą Fouriera funkcji  $\psi$ .

Jeśli  $\psi(t)$  zanika szybciej niż  $|t|^{-1}$  dla  $t \rightarrow \infty$ , wówczas oba warunki są równoważne. Kształt kolejnych funkcji waveletowych zależy od parametru  $s$ . Jeśli  $s < 1$  są to funkcje coraz węższe, jeśli natomiast  $s > 1$  następuje stopniowe rozszerzanie funkcji matki.

Główną ideą transformaty wavelet jest przedstawienie dowolnej funkcji  $f$  jako superpozycji waveletów stanowiących jądro transformaty. Każde takie przekształcenie jest dekompozycją funkcji  $f$  na różne poziomy rozdzielczości przestrzennej (lub czasowej). Jednym ze sposobów otrzymania takiej reprezentacji funkcji  $f$  jest zapis w postaci całki po parametrach  $s$  i  $x$  rodziny funkcji wavelet  $\psi^{s,x}$  z odpowiednimi współczynnikami. Ze względów praktycznych wygodniej jest wyrazić funkcję  $f$  w postaci dyskretnej superpozycji zastępując całkę operatorem sumowania. Wprowadza się wtedy najczęściej dyskretyzację

postaci  $s = s_0^m, x = nx_0 s_0^m$ , gdzie  $m, n \in \mathbb{C}$ , oraz przyjęte wartości  $s_0 > 1, x_0 > 0$ . Wówczas waveletowa dekompozycja funkcji przedstawia się następująco:

$$f = \sum_{m,n \in \mathbb{C}} c_{m,n}(f) \psi_{m,n}, \quad (13.3)$$

gdzie  $\psi_{m,n}(t) = \psi_{s_0^m, nx_0 s_0^m}^-(t) = s_0^{-m/2} \psi(s_0^{-m} t - nx_0)$ .

Dla wartości  $s_0 = 2, x_0 = 1$  wybieramy ortonormalną bazę  $\psi_{m,n}$ . W tym przypadku

$$c_{m,n}(f) = \langle \psi_{m,n}, f \rangle = \int \psi_{m,n}(x) f(x) dx. \quad (13.4)$$

Różne bazy o takiej właśnie naturze skonstruowali Stromberg, Meyer, Lemarie, Battle, Daubechies. Wszystkie te zbiory funkcji nawiązują do matematycznego narzędzia wprowadzonego przez Mallata, zwanego analizą wielorozdzielczą. Narzędzie to może być dobrze przystosowane do analizy obrazów z wykorzystaniem baz waveletowych, pozwalając jednocześnie na konstrukcję szybkich algorytmów obliczeniowych.

W waveletowej analizie wielorozdzielczej występują dwie zasadnicze funkcje: wavelet matka  $\psi$  oraz funkcja skalująca  $\phi$ . Konstruuje się analogiczny do  $\psi_{m,n}$  zbiór funkcji uzyskany poprzez skalowanie i translację funkcji skalującej jako  $\phi_{m,n}(x) = 2^{-m/2} \phi(2^{-m} x - n)$ . Dla ustalonej wartości  $m$  funkcje  $\phi_{m,n}$  są ortonormalne. Oznaczmy przez  $V_m$  przestrzeń rozpiętą przez funkcje  $\phi_{m,n}$ . Takie przestrzenie  $V_m$  opisują przestrzenie sukcesywnej aproksymacji  $\dots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \dots$ , każda z rozdzielczością  $2^m$ . Dla każdej wartości  $m$  zbiór funkcji  $\psi_{m,n}$  rozpina przestrzeń  $W_m$ , która stanowi dokładnie ortogonalne dopełnienie  $V_m$  w przestrzeni  $V_{m-1}$ . Stąd współczynniki obliczone jako iloczyn skalarny  $\langle \psi_{m,n}, f \rangle$  opisują informację traconą podczas przechodzenia od aproksymacji funkcji  $f$  z rozdzielczością  $2^{m-1}$  do bardziej zgrubnej aproksymacji o rozdzielczości  $2^m$ .

Z powyższego wynika algorytm obliczania współczynników  $c_{m,n}(f) = \langle \psi_{m,n}, f \rangle$  opisany następującymi równaniami:

$$\begin{aligned} c_{m,n}(f) &= \sum_k g_{2n-k} a_{m-1,k}(f) \\ a_{m,n}(f) &= \sum_k h_{2n-k} a_{m-1,k}(f) \end{aligned} \quad (13.5a)$$

$$\text{gdzie} \quad g_l = (-1)^l h_{-l+1} \quad (13.5b)$$

$$\text{oraz} \quad h_n = 2^{1/2} \int \phi(x-n) \phi(2x) dx. \quad (13.5c)$$

W rzeczywistości wartości  $a_{m,n}(f)$  są współczynnikami charakteryzującymi projekcję funkcji  $f$  na przestrzeń  $V_m$ .

Jeśli funkcja  $f$  dana jest w postaci dyskretnej, można potraktować jej próbki jako współczynniki aproksymacji o najwyższej rozdzielczości  $a_{0,n}$ , a powyższe równania opisują algorytm waveletowej dekompozycji wartości tych próbek przy pomocy dolnoprzepustowego filtra postaci  $h$  i górnoprzepustowego filtra postaci  $g$ . Ponieważ postać tych filtrów jest bezpośrednio związana z ortonormalną bazą waveletową, przekształcenia te pozwalają uzyskać dokładną rekonstrukcję próbek funkcji  $f$  w sposób następujący:

$$a_{m-1,l}(f) = \sum_k [h_{2n-l} a_{m,n}(f) + g_{2n-l} c_{m,n}(f)]. \quad (13.6)$$

Większość ortonormalnych baz waveletowych oparta jest na funkcjach  $\psi$  o nieskończonym nośniku, którym odpowiadają filtry  $h$  i  $g$  o nieskończonej ilości współczynników. Istnieją jednak metody konstrukcji funkcji  $\psi$  o skończonym nośniku wraz z odpowiadającymi im filtrami SOI (o skończonej odpowiedzi impulsowej), będące podstawą baz ortonormalnych wykorzystywanych w metodach analizy wielorozdzielczej. Uzyskuje się wówczas dokładną rekonstrukcję obrazów oryginalnych poprzez zastosowanie tych samych filtrów SOI zarówno do analizy (proces dekompozycji), jak i do syntezy (proces rekonstrukcji). Zasady konstrukcji takich filtrów opublikowali między innymi Smith i Barnwell, a także Vetterli.

Dodatkowym efektem ortonormalnej waveletowej dekompozycji obrazu jest tworzenie nowej reprezentacji danych jako superpozycji stosunkowo gładkich bloków elementarnych pokrywających przestrzeń obrazu. Filtry muszą spełniać dodatkowy warunek:

$\prod_{k=1}^{\infty} H(2^{-k} \xi)$  zanika szybciej niż  $C(1+|\xi|)^{-\epsilon-0.5}$  przy  $|\xi| \rightarrow \infty$  dla dowolnego  $\epsilon > 0$ , gdzie  $H(\xi) = 2^{-l/2} \sum_n h_n e^{-jn\xi}$ . Te wymaganie regularności jest niespełnione przez większość filtrów dokładnej rekonstrukcji dostępnych w literaturze.

#### *Biortogonalne bazy waveletowe*

Ponieważ obrazy są zazwyczaj w przeważającej części gładkie (wyłączając pojawiające się gdzieś krawędzie), wydaje się iż w schematach wielorozdzielczej dekompozycji, pozwalających dokładnie rekonstruować obraz oryginalny na podstawie waveletowej reprezentacji, do analizy powinna być wykorzystywana baza ortonormalna zbudowana na podstawie stosunkowo gładkiej waveletowej funkcji matki. Aby uzyskać szybki algorytm dekompozycji obrazu filtry powinny mieć krótki nośnik. Nie mogą być jednak zbyt krótkie, bo tracimy wówczas gładkość bazowych funkcji przekształcania. Ponadto bardzo korzystną cechą filtrów wykorzystywanych w hierarchicznej (piramidalnej) dekompozycji danych jest ich liniowa faza. Nie ma wówczas potrzeby kompensacji fazy na poszczególnych stopniach dekompozycji w występujących w praktyce warunkach ograniczonej długości przekształcanego zbioru danych. Niestety nie ma nietrywialnych postaci ortonormalnych filtrów SOI o liniowej fazie pozwalających na dokładną rekonstrukcję sygnału, nie zważając na jakiegokolwiek warunki regularności filtrów. Jedyne symetryczne filtry pozwalające na dokładną rekonstrukcję wynikają z bardzo prostej waveletowej bazy Haara, która jest wysoce niekorzystna z punktu widzenia wymagań gładkości (regularności).

Aby zachować liniowość fazy filtrów, co odpowiada warunkowi symetryczności funkcji waveletowych, rezygnuje się z wymagania ortonormalności bazy, tworząc bazę biortogonalną (funkcje bazowe są liniowo niezależne, ale nie są ortonormalne). Można wówczas również konstruować bazy o stosunkowo wysokiej regularności. W schemacie analizy i syntezy z bazą biortogonalną dekompozycja jest identyczna jak w przypadku bazy ortonormalnej (13.5), natomiast rekonstrukcja przebiega następująco:

$$a_{m-1,l}(f) = \sum_k [\tilde{h}_{2n-l} a_{m,n}(f) + \tilde{g}_{2n-l} c_{m,n}(f)] \quad (13.7)$$

gdzie filtry  $\tilde{h}$ ,  $\tilde{g}$  mogą być różne od  $h$ ,  $g$ . Aby uzyskać dokładną rekonstrukcję muszą być spełnione warunki:

$$\begin{aligned} \tilde{g}_n &= (-1)^n h_{-n+1} & \sum_n h_n \tilde{h}_{n+2k} &= \delta_{k,0} \\ g_n &= (-1)^n \tilde{h}_{-n+1} \end{aligned} \quad (13.8)$$

Tak więc jedyną różnicą w schemacie waveletowej dekompozycji (w stosunku do bazy ortonormalnej) jest fakt, iż filtry stosowane do analizy i syntezy obrazu są różne. Jeśli natomiast filtry spełniają dodatkowy warunek:

$$\prod_{k=1}^{\infty} \tilde{H}(2^{-k} \xi) \quad \text{i} \quad \prod_{k=1}^{\infty} H(2^{-k} \xi). \quad (13.9a)$$

zanika szybciej niż  $C(1+|\xi|)^{-\varepsilon-0.5}$  przy  $|\xi| \rightarrow \infty$  dla dowolnego  $\varepsilon > 0$ , gdzie

$$\tilde{H}(\xi) = 2^{-1/2} \sum_n \tilde{h}_n e^{-jn\xi} \quad H(\xi) = 2^{-1/2} \sum_n h_n e^{-jn\xi}, \quad (13.9b)$$

wówczas można sformułować następującą interpretację równań (13.5) i (13.7). Zdefiniujemy funkcje  $\phi$  i  $\tilde{\phi}$  jako  $\phi(x) = \sum_n h_n \phi(2x - n)$  i  $\tilde{\phi}(x) = \sum_n \tilde{h}_n \tilde{\phi}(2x - n)$ . Ich transformaty

Fouriera są dokładnie nieskończonymi iloczynami z równania (5a) i stąd są one dobrze zdefiniowanymi funkcjami całkowalnymi z kwadratem, o skończonym nośniku jeśli filtry  $h$ ,  $\tilde{h}$  są filtrami SOI. Zdefiniujemy ponadto  $\psi(x) = \sum_n g_n \phi(2x - n)$  i  $\tilde{\psi}(x) = \sum_n \tilde{g}_n \tilde{\phi}(2x - n)$ .

Wówczas równanie (13.6) może być napisane w postaci:

$$\begin{aligned} a_{m,n}(f) &= \langle \phi_{m,n}, f \rangle = 2^{-m/2} \int \phi_{m,n}(x) f(x) dx \\ c_{m,n}(f) &= \langle \psi_{m,n}, f \rangle = 2^{-m/2} \int \psi_{m,n}(x) f(x) dx \end{aligned} \quad (13.10)$$

a zapis rekonstrukcji jest wówczas bardzo prosty:

$$f = \sum_{m,n} \langle \psi_{m,n}, f \rangle \tilde{\psi}_{m,n}. \quad (13.11)$$

Jeśli nieskończone iloczyny z równania (13.9a) znikają nawet szybciej niż to było wspomniane wyżej, wtedy  $\phi$  i  $\tilde{\phi}$  oraz odpowiednio  $\psi$  i  $\tilde{\psi}$  będą stosunkowo gładkie. W szczególności zbieżność nieskończonych iloczynów (13.9a) jest możliwa tylko wtedy, jeśli  $\sum_n h_n = 2^{1/2}$  i  $\sum_n \tilde{h}_n = 2^{1/2}$ . Ponadto równanie (13.11) jest spełnione jedynie, gdy  $\sum_n (-1)^n h_n = 0$  i  $\sum_n (-1)^n \tilde{h}_n = 0$ . Większość schematów waveletowej dekompozycji wiernie rekonstruujących obraz nie spełnia tych warunków.

Prowadzono także szereg rozważań nad regularnością funkcji baz biortogonalnych. Wykazano (Cohen, Daubechies, Feauveau), że można osiągnąć stosunkowo dużą regularność funkcji  $\psi$  i  $\tilde{\psi}$ , budując jednocześnie odpowiednio długie skojarzone z nimi filtry. Konkretnie, jeśli funkcje  $\psi$  i  $\tilde{\psi}$  są odpowiednio  $(k-1)$  i  $(\tilde{k}-1)$  razy różniczkowalne w sposób ciągły (odpowiednie pochodne są ciągłe), to trygonometryczne wielomiany  $H(\xi)$  i

$\tilde{H}(\xi)$  muszą być podzielne odpowiednio przez  $(1+e^{-j\xi})^k$  i  $(1+e^{-j\xi})^{\tilde{k}}$ , tak że długość odpowiadających im filtrów  $h, \tilde{h}$  musi przekraczać wartości  $k, \tilde{k}$ .

Z równań (13.8) wynika, że podzielność  $\tilde{H}(\xi)$  przez  $(1+e^{-j\xi})^{\tilde{k}}$  oznacza iż funkcja  $\psi$  będzie miała  $\tilde{k}$  kolejnych zerowych momentów:

$$\int x^l \psi(x) dx = 0, \quad \text{dla } l = 0, 1, \dots, \tilde{k} - 1. \quad (13.12)$$

Wiadomo (można to wykazać poprzez zastosowanie rozwinięcia z szereg Taylora), że jeśli  $\psi$  ma  $\tilde{k}$  momentów zerowych, wtedy współczynniki  $\langle \psi_{m,n}, f \rangle$  będą reprezentacją funkcji  $f$ , które są  $\tilde{k}$  razy różniczkowalne, wysoce korzystną z punktu widzenia skutecznej filtracji, tj. wiele współczynników będzie miało nieznaczące bardzo małe wartości. Nieliniowa filtracja polega na odrzucaniu tych współczynników i zachowaniu wartości niewielkiej liczby współczynników zawierających niemal całą użyteczną informację o sygnale (obrazie).

### 13.6. Algorytmy kompresji

Trzy zasadnicze elementy przedstawianego algorytmu to: wybór najlepszej postaci filtrów do wielorozdzielczej dekompozycji obrazów, przestrzenno-częstotliwościowa kwantyzacja oraz entropijne kodowanie strumienia danych nadające ostateczny kształt nowej reprezentacji kompresowanego zbioru danych.

Ważnym problemem w kodowaniu pasmowym jest projektowanie efektywnych filtrów do analizy i syntezy obrazu. Właściwy wybór odpowiedniej do konstrukcji banku filtrów rodziny waveletowej z takimi kryteriami jak: kształt odpowiedzi impulsowej i skokowej, regularność, liniowość, ortogonalność lub biortogonalność, symetria, pozwala znacząco poprawić efektywność kompresji całego algorytmu

Ze względu na: niestacjonarność procesu modelującego kompresowany zbiór danych, dużą różnorodność cech obrazu istotnych z punktu widzenia zastosowań, różny poziom jakości obrazów oryginalnych (stosunek sygnału do szumów, przestrzenna rozdzielczość, częstotliwościowe widmo sygnału itp.) z różnych systemów obrazowania, bardzo trudnym zadaniem okazuje się opracowanie metody konstrukcji filtrów optymalnych z punktu widzenia efektywności kompresji. Na obecnym poziomie wiedzy wiele zagadnień odnoszących się do wyboru najbardziej skutecznych w kompresji filtrów jest nadal nierozwiązanych.

Interesujące są odpowiedzi na dwa pytania: jaki rodzaj dekompozycji winien być zastosowany (diadyczny, równomierny, z wyborem najlepszej bazy i pakietem waveletów) oraz jaka rodzina waveletów (o jakich własnościach, jaki bank filtrów)?

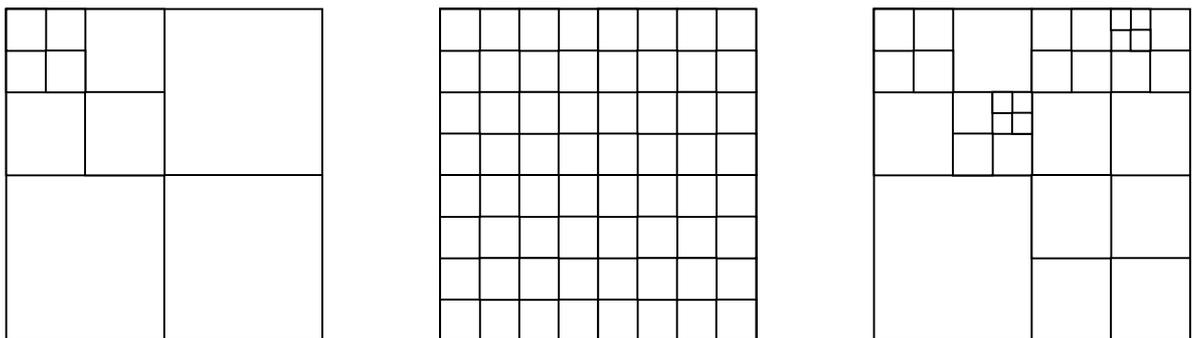
Z przeprowadzonej analizy, badań własnych oraz literaturowych i internetowych, wynikają następujące wnioski:

- diadyczny schemat dekompozycji obrazu o dość zróżnicowanej liczbie poziomów, zależnej od obrazu i techniki kwantyzacji-kodowania, okazuje się w większości przypadków najefektywniejszy (świadczą o tym chociażby wyniki porównań efektywności najlepszych waveletowych koderów na stronie UCLA:

[http://www.icsl.ucla.edu/~ipl/psnr\\_results.html](http://www.icsl.ucla.edu/~ipl/psnr_results.html));

- schematy równomiernej dekompozycji w pewnych przypadkach dają zadawalające rezultaty [11], jednak są to bardzo rzadkie przypadki i raczej nie najskuteczniejsze;
- adaptacyjny wybór najlepszej bazy wymaga zapisu dodatkowej informacji w kodzie wyjściowym o kształcie drzewa dekompozycji obrazu, ogranicza także w pewnym stopniu możliwości usuwania nadmiarowości na etapie kodowania; w efekcie daje niewspółmierne wyniki w stosunku do kosztów kosztowo-obliczeniowych (np. poprzez zastosowanie pakietu waveletów efektywność metody SFQ [12] w kompresji obrazu testowego Lenna zwiększyła się zaledwie o 0.02 dB dla 0.25bpp i 0.05 dB dla 0.5bpp w stosunku do metody z dekompozycją diadyczną); nie wprowadza więc to rozwiązanie żadnej nowej jakości;
- rodziny waveletów niezmienniczych względem przesunięcia lub też dzielących pasmo sygnału na więcej niż dwa kanały, zróżnicowane na każdym poziomie dekompozycji obrazu itd. jakkolwiek pozwalają uzyskać w pewnych przypadkach konkurencyjną skuteczność kompresji, jednak w szerszej gamie zastosowań nie są konkurencyjne w stosunku do najefektywniejszych baz waveletowych generujących banki filtrów biortogonalnych;
- bazy biorthogonalne, dwukanałowe, są najskuteczniejsze do zastosowań kompresji obrazów;
- wpływ szeregu parametrów funkcji waveletowych, takich jak regularność, liczba momentów znikających, resztkowy aliasing, długość filtrów, na skuteczność kompresji jest niejednoznaczny; wydaje się natomiast że dwa parametry są najsilniej skorelowane z efektywnością kompresji: wzmocnienie kodowe oraz geometryczna parametryzacja kształtu funkcji [5]; **stąd też powstał nasz pomysł przybliżania kształtu waveletu do średniego kształtu falki wyznaczonej dla obrazu czy grupy obrazów; koncepcja ta nie została jednak jeszcze zrealizowana - poszukiwane są narzędzia wyznaczające taką średnią falkę dla obrazu, jak też skorelowaną z nimi metodę generacji filtrów;**

Przykładowe schematy dekompozycji pokazano na rys. 13.4.

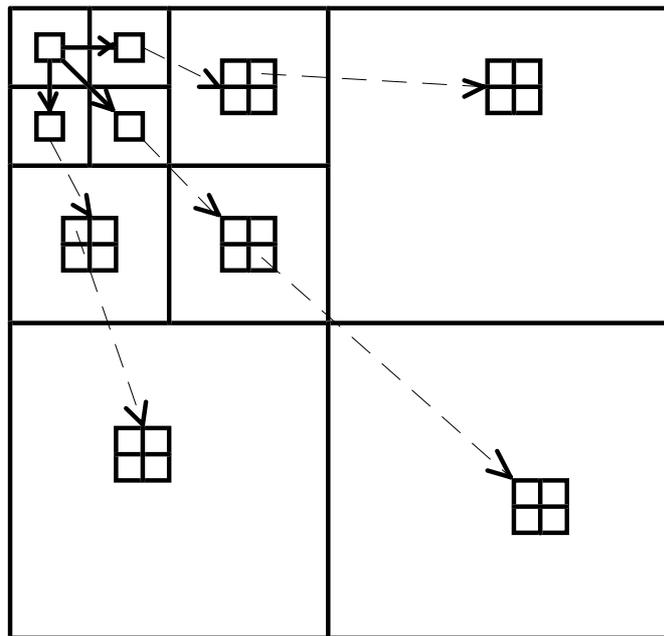


Rys. 13.4. a) dekompozycja diadyczna, b) dekompozycja równomierna, c) adaptacyjny wybór najlepszej bazy.

Typowy schemat kwantyzacji skalarnej lub wektorowej może być ulepszony poprzez wykorzystanie przestrzenno-częstotliwościowej charakterystyki waveletowej reprezentacji. Ponieważ waveletowa reprezentacja danych zawiera obok częstotliwościowej także składową przestrzenną (czasową), przestrzenny model grupowania i kwantyzacji danych jest możliwy.

Kolejność kodowania skwantowanych wartości współczynników transformaty wavelet wynika z ilości informacji o obrazie oryginalnym zawartych w poszczególnych pasmach.

Selekcja widma pozwalająca na progresywną transmisję kompresowanych obrazów jest wykonywana w tradycyjny sposób. Jako pierwsze kodowane jest pasmo najniższych częstotliwości LL zawierające najwięcej informacji o obrazie (licząc na pojedynczy współczynnik należący do pasma), potem kolejne bloki o najmniejszej skali: LH, HL i na końcu HH, gdzie L- dolnoprzepustowa filtracja, H- górnoprzepustowa filtracja. Są to bloki sąsiednie w przestrzeni obrazu, odpowiednio: lewy-górny, prawy-górny, lewy-dolny, prawy-dolny określone przez horyzontalną relację drzewa dekompozycji - zobacz rys. 13.5. Następnie kodowane są trzy bloki kolejnego poziomu drzewa będące w relacji rodzic-dzieci w stosunku do bloków przetwarzanych poprzednio: prawy-górny, lewy-dolny, prawy-dolny itd.



Rys.13.5. Podstawowy schemat waveletowej dekompozycji obrazu. — - relacje poziome (horyzontalne) drzewa, - - - relacje rodzic - dzieci.

Typowa postać drzewa hierarchicznej dekompozycji obrazu w podstawowym waveletowym algorytmie została przedstawiona na rys. 13.5. Cztery pasma składowych o najniższych częstotliwościach stanowią szczyt tego drzewa. Dane należące do tego poziomu nie mają rodzica i są rodzicami pierwszej generacji dla wszystkich skojarzonych przestrzennie współczynników. Każdy współczynnik na poziomie różnym od podstawy drzewa rozrasta się w grupę czterech współczynników kolejnego poziomu o większej skali, będąc z nimi w bezpośredniej relacji rodzic-dzieci. Struktura drzewa zer odpowiada dokładnie strukturze drzewa dekompozycji obrazu oryginalnego, przy czym w węzłach drzewa znajdują się początkowo jedynie symbole dwuelementowego alfabetu: znaczący lub nieznaczący. Po kwantyzacja tworzone jest więc drzewo binarne, gdzie w węzłach z niezerowymi współczynnikami występuje 1 (węzeł znaczący), a z zerowymi 0 (węzeł nieznaczący). Jedynie dla wartości współczynników z pasma LL symbole drzewa zer przypisuje się nieco inaczej. Gałęzmi tego drzewa są relacje horyzontalne i rodzic-dzieci z rys. 13.5.

Procedura okrajania drzewa zer rozpoczyna się na najniższym poziomie drzewa (podstawa) o największej skali. Dla kolejnych grup czterech węzłów sprawdzane są dwa warunki: czy ich rodzic jest nieznaczący i czy wszystkie dzieci są nieznaczące. W przypadku

pozytywnej odpowiedzi na oba pytania drzewo jest okrajane: usuwane są cztery nieznaczące węzły podstawy drzewa wraz z gałęziami łączącymi je z rodzicem, a węzłowi rodzica przypisywany jest nowy symbol PBN (ang. pruned branch node) oznaczający węzeł obciętych gałęzi. Tak więc w wyniku okrajania drzewa alfabet symboli opisujących węzły staje się trzelementowy. Z przeprowadzonych badań wynika, iż dalsze rozszerzanie alfabetu drzewa jest niekorzystne z punktu widzenia efektywności kompresji. Na poziomach środkowych drzewa zer okrajanie przebiega analogicznie, z małą jednak różnicą, że warunkiem okrajania jest wystąpienie nieznaczącego rodzica i czterech dzieci-węzłów opisanych symbolem PBN. Ostatni etap okrajania drzewa zer wygląda nieco inaczej. Redukcja drzewa jest dokonywana na tym samym, szczytowym poziomie, a rozważane są korelacje wartości czterech węzłów według relacji horyzontalnych. Jeśli wartość współczynnika z pasma LL jest znacząca, a pozostałe trzy węzły określone jako PBN, wówczas likwidowane są trzy gałęzie wraz z węzłami, a węzeł z pasma LL oznaczany jako PBN. Przestrzenna korelacja danych na niższych poziomach drzewa jest zbyt mała, aby wprowadzić analogiczny mechanizm okrajania drzewa według relacji horyzontalnych także na tych poziomach. Z przeprowadzonych badań wynika, iż uzyskuje się wówczas zmniejszenie efektywności kompresji.

### 13.7. Konstrukcja filtrów waveletowych

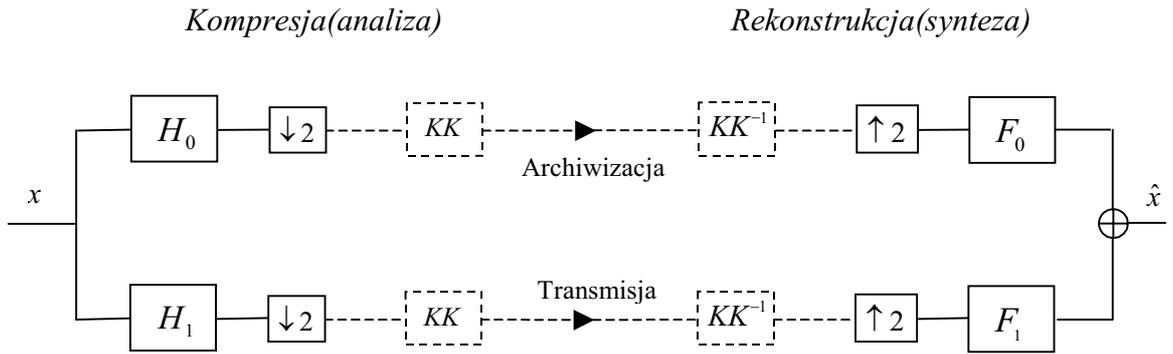
W rozdziale tym przedstawione są podstawowe zagadnienia dotyczące przede wszystkim podstaw techniki tworzenia banków filtrów, które mogą być zastosowane w schematach kompresji danych. Przytoczono kilka różnych metod konstrukcji filtrów, reprezentujących różne sposoby podejścia do zagadnienia optymalizacji podziału częstotliwościowego pasma sygnału na pod-pasma z zachowaniem informacji o skali (częstotliwości).

#### *Wprowadzenie*

Ponieważ interesuje nas praktyczna realizacja algorytmów kompresji danych cyfrowych, nasze rozważania będą prowadzone przy założeniu cyfrowej postaci sygnałów czy danych wejściowych, jak również cyfrowej postaci danych rekonstruowanych w procesie stratnej/bezstratnej kompresji. Konkretnie interesujące nas zagadnienie konstrukcji banków filtrów do kompresji dotyczy oczywiście filtrów cyfrowych.

Idea dekompozycji sygnału wejściowego w celu redukcji różnego typu nadmiarowości nie może być niestety zrealizowana przy pomocy jednego filtru do analizy i jednego filtru do syntezy (schemat jednokanałowy). Nie sposób uzyskać wówczas wiernej rekonstrukcji sygnału wejściowego (niemożliwy jest do spełnienia warunek doskonałej rekonstrukcji). Konieczna jest więc budowa co najmniej dwukanałowego modelu dekompozycji/rekonstrukcji i taka forma znajduje najczęstsze zastosowanie w algorytmach kompresji. Nie wyklucza to oczywiście stosowania schematów wielokanałowych, aczkolwiek ze względu na dużo bardziej złożony proces konstrukcji takich banków filtrów i brak znaczącej poprawy efektywności w stosunku do banków dwukanałowych, rozwiązania takie należą do rzadkości.

Elementarny schemat wykorzystania banku filtrów w metodzie pasmowego kodowania wygląda zazwyczaj tak jak na rys. 13.6.



Rys.13.6. Schemat dekompozycji (analiza) oraz rekonstrukcji (syntezy) sygnału przy pomocy dwukanałowego banku filtrów. Symbol  $\downarrow 2$  oznacza decymator, a  $\uparrow 2$  - ekspander. Z kolei  $KK$  to proces kwantyzacji/kodowania, odwracany przy rekonstrukcji sygnału oryginalnego ( $KK^{-1}$ ). W naszych rozważaniach pomijamy proces  $KK$  i archiwizacji domagając się spełnienia warunku doskonałej rekonstrukcji przez algorytm filtracji wraz z decymacją - ekspandowaniem.

Wejściowy sygnał  $x$  poddawany jest filtracji poprzez zastosowanie pary filtrów: dolnoprzepustowego  $H$  oraz górnoprzepustowego  $G$ . Proces filtracji może być prosto zrealizowany jako splot w dziedzinie oryginalnej, bądź też jako iloczyn w dziedzinie częstotliwościowej po odpowiednich przekształceniach dyskretnej transformaty Fouriera. Można także opisać filtrację w dziedzinie transformaty  $Z$ , która jest bardzo użyteczna w rozważaniach dotyczących filtrów cyfrowych.

Tak więc dolnoprzepustową filtrację sygnału wejściowego  $x$  można opisać jako splot w dziedzinie oryginalnej w sposób następujący:

$$y_0(n) = \sum_{k=0}^L h(k)x(n-k), \quad (13.13)$$

gdzie  $y_0(n)$  oznacza wartości sygnału wyjściowego po filtracji dolnoprzepustowej,  $L+1$  - długość filtru,  $h(k)$ - współczynniki filtru dolnoprzepustowego; rozważamy oczywiście praktycznie realizowalne filtry SOI (o skończonej odpowiedzi impulsowej), gdzie  $L < \infty$ . Splot ten można zapisać prościej jako:  $y_0 = h \otimes x = H \cdot x$

Iloczyn w dziedzinie częstotliwościowej przedstawia równanie

$$Y_0(\omega) = H(\omega) \cdot X(\omega), \quad (13.14)$$

gdzie  $X(\omega) = \sum_n x(n) \cdot e^{-j\omega n}$  - dyskretna transformata Fouriera sygnału wejściowego,

$H(\omega) = \sum_{n=0}^L h(n) \cdot e^{-j\omega n}$  - transformata filtru.

Analogicznie filtracja jako iloczyn w dziedzinie transformaty  $Z$  wygląda jak niżej

$$Y_0(z) = H(z) \cdot X(z), \quad (13.15)$$

gdzie  $X(z) = \sum_n x(n) \cdot z^{-n}$  - transformata  $Z$  sygnału wejściowego,  $H(z) = \sum_{n=0}^L h(n) \cdot z^{-n}$  - transformata  $Z$  filtru przyczynowego o skończonej odpowiedzi impulsowej.

W analogiczny sposób uzyskujemy wartości  $y_1 = G \cdot x$  po filtracji górnoprzepustowej, przy czym liczba próbek zarówno  $y_0(n)$  jak i  $y_1(n)$  jest równa ilości próbek sygnału wejściowego. Nastąpiło więc dwukrotne zwiększenie ilości próbek, co jest niewątpliwie niekorzystne z punktu widzenia kompresji. Istnieje poza tym zagrożenie zniekształceniem typu *aliasing* wzmacniania pewnych fragmentów sygnału poprzez nakładanie się pasm z obu kanałów (dolno- i górnoprzepustowego), powodowane niemożnością realizacji filtrów o idealnie pasmowych charakterystykach.

Rozwiązaniem tych obu problemów jest decymacja  $\downarrow 2$  ciągu próbek  $y_0(n)$  i  $y_1(n)$ , czyli wybieranie co drugiej próbki. Otrzymujemy wówczas ciągi  $v_0(n) = (\downarrow 2)y_0(n) = y_0(2n)$  oraz  $v_1(n) = (\downarrow 2)y_1(n) = y_1(2n)$ , czy czym długość każdego z nich jest równa połowie długości wejściowego wektora danych.

Otrzymany zbiór danych  $v_0$  i  $v_1$  może być następnie zakodowany tworząc skompresowaną reprezentację wejściowego wektora danych  $x$ . Odwrotny proces rekonstrukcji, będący złożeniem działań ekspandera  $\uparrow 2$  (dopisanie zera po każdej próbce zdekodowanych ciągów  $v$ ) oraz filtrów  $\tilde{H}$  i  $\tilde{G}$ , przedstawia następujące wyrażenie:

$$\hat{x} = \tilde{H}(\uparrow 2)v_0 + \tilde{G}(\uparrow 2)v_1. \quad (13.16)$$

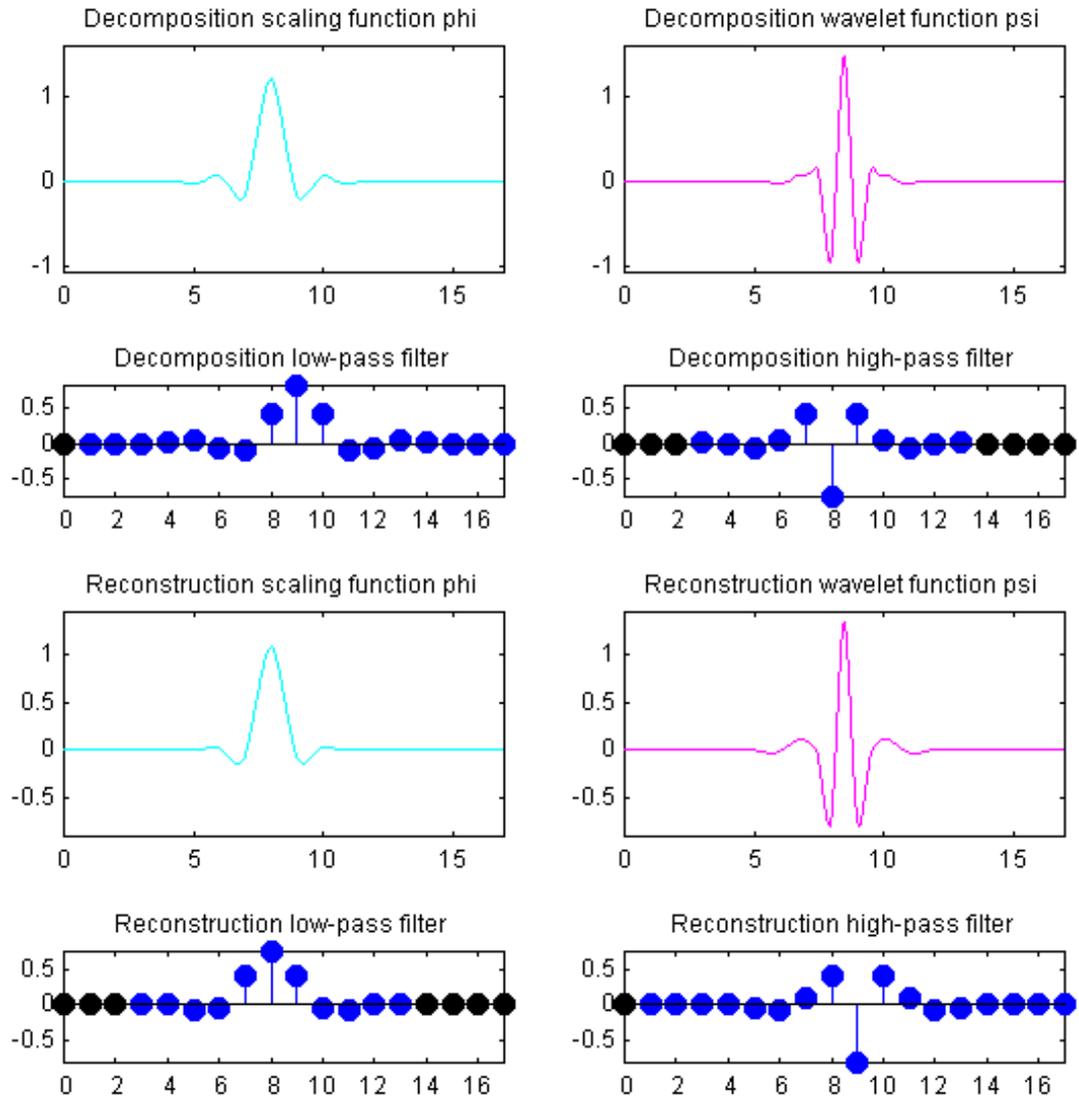
Celem jest takie zaprojektowanie schematu analizy i syntezy, aby spełniony był warunek  $x = \hat{x}$ . Bank filtrów powinien więc spełniać warunek doskonałej rekonstrukcji, czyli wiernego odtworzenia sygnału oryginalnego.

### Przykłady filtrów

Charakterystyki jednego z efektywnych biortogonalnych banków filtrów zaprezentowano na rys. 13.7. Współczynniki kilku innych popularnych zestawów filtrów zawiera tabela 13.1. Obok trzech zestawów filtrów ortogonalnych, definiowanych jednoznacznie przez filtr dolnoprzepustowy  $h$  oraz skojarzony z nim filtr górnoprzepustowy  $g$  według równania (13.5b) umieszczono tam także banki filtrów biortogonalnych. W tym przypadku relacje pomiędzy odpowiednimi filtrami dolno- i górnoprzepustowymi określają wyrażenia (13.8), wystarczy więc podać postać filtrów dolnoprzepustowych do analizy i syntezy:  $h$  i  $\tilde{h}$ .

Tabela 13.1. Współczynniki wybranych banków filtrów, zarówno ortogonalnych (Daubechies i Adelson) jak i biortogonalnych Antonini i Villasenor).

| Daubechies<br>10 | Daubechies<br>14 | Adelson [] | Antonini [] |             | Villasenor [] |             |
|------------------|------------------|------------|-------------|-------------|---------------|-------------|
| $h$              | $h$              | $h$        | $h$         | $\tilde{h}$ | $h$           | $\tilde{h}$ |
| 0.160102         | 0.077852         | 0.028220   | 0.037828    | -0.064539   | 0.288526      | 0.000954    |
| 0.138428         | 0.396539         | -0.060394  | -0.023849   | -0.040689   | 0.00008       | -0.000003   |
| -0.242294        | 0.729132         | -0.073882  | -0.110624   | 0.418092    | -0.157526     | -0.009452   |
| -0.032245        | 0.469782         | 0.413948   | 0.377403    | 0.788486    | 0.076790      | -0.002528   |
| 0.077571         | -0.143906        | 0.798430   | 0.852699    | 0.418092    | 0.758908      | 0.030834    |
| 0.603829         | -0.224036        | 0.413948   | 0.377403    | -0.040689   | 0.758908      | -0.013765   |
| 0.724309         | 0.071309         | -0.073882  | -0.110624   | -0.064539   | 0.076790      | -0.085661   |
| -0.006241        | 0.080613         | -0.060394  | -0.023849   |             | -0.157526     | 0.163369    |
| -0.012581        | -0.038030        | 0.028220   | 0.037828    |             | 0.00008       | 0.623360    |
| 0.003336         | -0.016575        |            |             |             | 0.288526      | 0.623360    |
|                  | 0.012551         |            |             |             |               | 0.163369    |
|                  | 0.000430         |            |             |             |               | -0.085661   |
|                  | -0.001802        |            |             |             |               | -0.013765   |
|                  | 0.000354         |            |             |             |               | 0.030834    |
|                  |                  |            |             |             |               | -0.002528   |
|                  |                  |            |             |             |               | -0.009452   |
|                  |                  |            |             |             |               | -0.000003   |
|                  |                  |            |             |             |               | 0.000954    |



Rys. 13.5. Charakterystyka - funkcje waveletowe i skalujące oraz współczynniki biortogonalnego banku filtrów o rozmiarach 17/11. Wartości współczynników tego filtru są następujące:

$$h = \{0.001909, -0.001914, -0.016991, 0.011935, 0.049733, -0.077263, -0.094059, 0.420796, 0.825923, 0.420796, -0.094059, -0.077263, 0.049733, 0.011935, -0.016991, -0.001914, 0.001909\}$$

oraz

$$\tilde{h} = \{0.014427, 0.014467, -0.078722, -0.040368, 0.417849, 0.758908, 0.417849, -0.040368, -0.078722, 0.014467, 0.014426\}$$

### 13.7. Charakterystyka najefektywniejszych algorytmów kompresji

Do grupy najefektywniejszych koderów falkowych zaliczyć należy przede wszystkim: SPIHT, PACC, EQ, SFQ, C/B oraz PC-AUTQ. Niektóre z tych technik zostały scharakteryzowane w rozdziale 15. Techniki te różnią się przede wszystkim sposobem

kwantyzacji oraz kodowania współczynników falkowych. W każdym przypadku zredukowana jest nadmiarowość korelacji w przestrzeniach położenia oraz skali poprzez bądź odpowiednio dobrany algorytm czasowo-częstotliwościowej kwantyzacji (EQ, SFQ, C/B oraz PC-AUTQ), bądź też praktycznie w każdym przypadku odpowiednio dobrany sposób kodowania. Oparty on jest w każdym przypadku na strukturze drzewa zer bardzo dobrze opisującej zależność w przestrzeni falkowej dekompozycji obrazu. Obok bardzo szybkich algorytmów o pełnej kontroli długości i kodu oraz progresji w transmisji informacji w algorytmie (SPIHT) znajdują się tutaj bardzo złożone i czasochłonne algorytmy (EQ, SFQ), których jedyną zaletą jest bardzo duża skuteczność kompresji, nieosiągalna żadną inną metodą.

Falkowa metoda kompresji, budowana na podstawie tych właśnie algorytmów stanowi opracowywany, nowy standard kompresji pojedynczych obrazów JPEG 2000. Jest to niewątpliwym dowodem pierwszeństwa tych metod w konstrukcji współczesnych koderów o szerokiej gamie zastosowań. Wyniki porównania efektywności różnych technik waveletowych przedstawia tabela 13.2.

Tabela 13.2. Porównanie efektywności kompresji różnych koderów falkowych.

| Compression technique | Lenna       |        | Barbara     |        | Goldhill    |        | MR          |        |
|-----------------------|-------------|--------|-------------|--------|-------------|--------|-------------|--------|
|                       | 0.25bp<br>p | 0.5bpp | 0.25bp<br>p | 0.5bpp | 0.25bp<br>p | 0.5bpp | 0.25bp<br>p | 0.5bpp |
| SPIHT                 | 34.13       | 37.24  | 27.79       | 31.72  | 30.63       | 33.19  | 34.85       | 39.10  |
| SFQ                   | 34.33       | 37.36  | 28.29       | 32.15  | 30.71       | 33.37  | -           | -      |
| C/B                   | 34.45       | 37.59  | 28.38       | 32.22  | 30.77       | 33.43  | 34.98       | 39.30  |
| PACC                  | 34.50       | 37.50  | 28.62       | 32.52  | 30.81       | 33.49  | 35.36       | 39.58  |
| PC-AUTQ               | 34.46       | 37.56  | -           | -      | 30.78       | 33.46  | -           | -      |
| EQ                    | 34.57       | 37.68  | -           | -      | 30.76       | 33.42  | -           | -      |