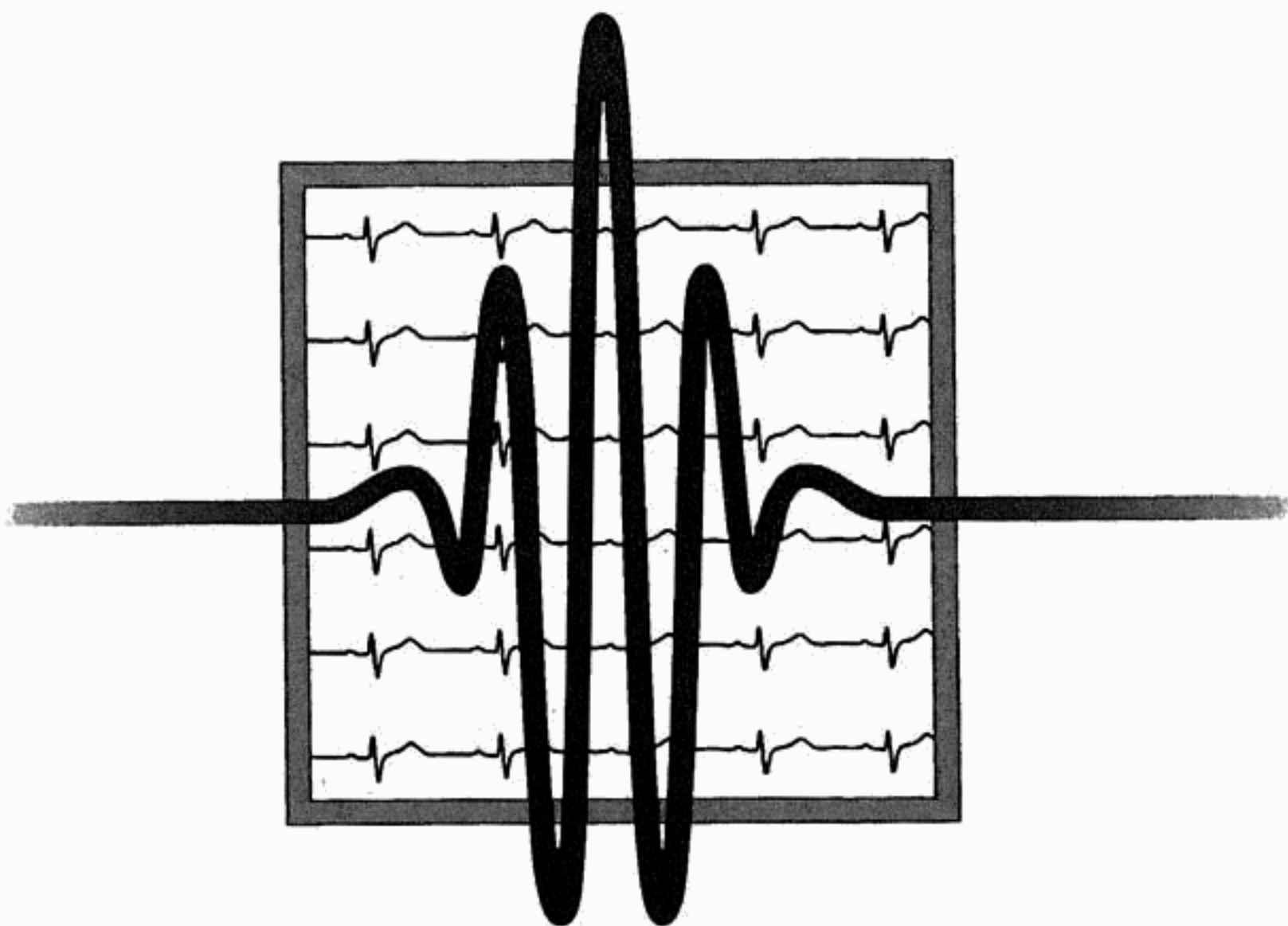


**PIOTR AUGUSTYNIAK**



**TRANSFORMACJE FALKOWE  
W ZASTOSOWANIACH  
ELEKTRODIAGNOSTYCZNYCH**



UCZELNIANE WYDAWNICTWA NAUKOWO-DYDAKTYCZNE

KRAKÓW 2003

KU 0083 pozycja wydawnictw naukowych  
Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie

© Wydawnictwa AGH, Kraków 2003  
ISBN 83-89388-10-3

Redaktor Naczelny Uczelnianych Wydawnictw  
Naukowo-Dydaktycznych: *dr inż. Jan Sas*

Z-ca Redaktora Naczelnego: *mgr Beata Barszczewska-Wojda*

Komitety Naukowe UWND AGH: *prof. dr hab. inż. Janusz Kowal (przewodniczący), prof. dr hab. inż. Tadeusz Banaszkowski, prof. dr hab. Bogdan Choczewski, dr hab. Zdzisław Cięciwa, prof. AGH, prof. dr hab. inż. Edward Fraś, prof. dr hab. inż. Ryszard Uberman*

Recenzent: *prof. dr hab. inż. Ryszard Tadeusiewicz*

Praca dofinansowana przez Zakłady Elektroniki Medycznej ASPEL S.A.

Projekt okładki i strony tytułowej: *Beata Barszczewska-Wojda i Piotr Augustyniak*

Opracowanie edytorskie: *Joanna Ciągala*

Skład komputerowy: „Andre”, tel. 423-10-10

Redakcja Uczelnianych Wydawnictw Naukowo-Dydaktycznych AGH  
al. Mickiewicza 30, 30-059 Kraków  
tel. 617-32-28, tel./fax 636-40-38  
e-mail: [wydagh@uci.agh.edu.pl](mailto:wydagh@uci.agh.edu.pl), [redakcja@wydawnictwoagh.pl](mailto:redakcja@wydawnictwoagh.pl)  
<http://www.WydawnictwoAGH.pl>

---

# Spis treści

<b>Spis wybranych skrótów .....</b>	<b>9</b>
<b>Wstęp .....</b>	<b>11</b>
<b>ROZDZIAŁ 1</b>	
<b>Pojęcia podstawowe .....</b>	<b>15</b>
1.1. Dyskretna reprezentacja sygnałów .....	15
1.2. Analiza sygnałów w dziedzinie częstotliwości .....	17
1.2.1. Analiza częstotliwościowa sygnałów stacjonarnych .....	18
1.2.2. Analiza częstotliwościowa sygnałów niestacjonarnych .....	21
1.3. Analiza sygnałów w dziedzinie czasowo-częstotliwościowej .....	27
1.3.1. Transformacja falkowa .....	27
1.3.2. Analiza wielorozdzielcza .....	34
1.3.3. Filtracja pasmowa sygnału .....	38
1.3.4. Rekonstrukcja .....	43
1.3.5. Zasada zachowania energii .....	45
1.4. Falki .....	45
1.4.1. Funkcja Haara .....	45
1.4.2. Falka Daubechies .....	46
1.4.3. Falki Coiflets .....	46
1.4.4. Falki biortogonalne .....	47
1.4.5. Porównanie parametrów falek .....	48
<b>ROZDZIAŁ 2</b>	
<b>Rozwinięcia metod dekompozycji sygnału .....</b>	<b>49</b>
2.1. Transformacja <i>Wavelet Packets</i> .....	49
2.1.1. Dekompozycja sygnału z użyciem pakietów falkowych .....	49
2.1.2. Poszukiwanie najlepszego drzewa dekompozycji .....	51
2.1.3. Pojęcie entropii .....	52
2.2. Techniki poszukiwania dopasowań .....	54
2.2.1. Wstęp .....	54
2.2.2. Algorytm poszukiwania dopasowań .....	55
2.2.3. Dobór elementów słownika .....	57

## **ROZDZIAŁ 3**

<b>Analiza falkowa – od teorii do implementacji .....</b>	<b>59</b>
3.1. Falki o nośniku zwartym .....	59
3.2. Projektowanie własnych falek .....	62
3.3. Predykcyjna transformacja falkowa .....	64
3.4. Transformacja falkowa w reprezentacji całkowitoliczbowej .....	67

## **ROZDZIAŁ 4**

<b>Ekstrakcja informacji diagnostycznych z reprezentacji czasowo-częstotliwościowej .....</b>	<b>71</b>
4.1. Detekcja późnych potencjałów komorowych w dziedzinie czasowo-częstotliwościowej .....	71
4.1.1. Definicja i znaczenie diagnostyczne mikropotencjałów .....	71
4.1.2. Detekcja mikropotencjałów w dziedzinie czasu .....	73
4.1.3. Detekcja mikropotencjałów w dziedzinie czasowo-częstotliwościowej .....	75
4.2. Detekcja zjawisk padaczkowych .....	77
4.2.1. Zastosowanie elektroencefalografii w detekcji padaczki .....	77
4.2.2. Detekcja iglic w dziedzinie czasowo-częstotliwościowej .....	80
4.3. Detekcja wrzecion snu w zapisie polisomnograficznym .....	83
4.4. Ocena zmienności rytmu serca za pomocą sieci falkowo-neuronowej .....	86

## **ROZDZIAŁ 5**

<b>Filtracja sygnału w dziedzinie czasowo-częstotliwościowej o parametrach dobieranych adaptacyjnie .....</b>	<b>89</b>
5.1. Eliminacja zakłóceń sygnału okoruchowego .....	89
5.1.1. Opis sygnału okoruchowego i jego zastosowań diagnostycznych .....	89
5.1.2. Pozyskiwanie sygnału okoruchowego .....	93
5.1.3. Zakłócenia sygnału okoruchowego i metody ich ograniczania .....	96
5.1.4. Filtracja sygnału okoruchowego w dziedzinie czasowo-częstotliwościowej .....	99
5.2. Ortogonalny podział pasma .....	104
5.2.1. Wyodrębnianie podpasm częstotliwościowych .....	104
5.2.2. Zastosowanie podziału pasma w sekwencyjnej kompresji wektorkardiogramów .....	107

## **ROZDZIAŁ 6**

<b>Kompresja sygnałów biomedycznych z wykorzystaniem reprezentacji czasowo-częstotliwościowych .....</b>	<b>111</b>
6.1. Stratne i bezstratne kodowanie reprezentacji czasowo-częstotliwościowej .....	113



6.1.1. Eliminacja progowa współczynników czasowo-częstotliwościowych .....	114
6.1.2. Kodowanie współczynników czasowo-częstotliwościowych .....	115
6.2. Kompresja optymalnego drzewa dekompozycji <i>Wavelet Packets</i> .....	118
6.2.1. Zastosowanie pakietów falkowych do bezstratnej kompresji elektrokardiogramu .....	120
6.2.2. Stratny algorytm kompresji elektrokardiogramu z użyciem pakietów falkowych .....	121
6.3. Kompresja z użyciem dodatkowych transformacji .....	124
6.3.1. Reprezentacja czasowo-częstotliwościowa sygnałów zdekorelowanych .....	125
6.3.2. Dekorelacja czasowo-częstotliwościowych reprezentacji sygnałów .....	127
6.4. Kompresja z modulacją dystrybucji zniekształceń .....	129
6.4.1. Aspekt lokalnej istotności elektrokardiogramu .....	129
6.4.2. Realizacja algorytmu kompresji z użyciem transformacji falkowej .....	130
6.4.3. Realizacja algorytmu kompresji z użyciem pakietów falkowych .....	136

## ROZDZIAŁ 7

<b>Odszumianie i modelowanie tła</b> .....	140
7.1. Redukcja reprezentacji czasowo-częstotliwościowej w poszczególnych pasmach .....	140
7.1.1. Jednostajne progowanie reprezentacji czasowo-częstotliwościowej .....	140
7.1.2. Adaptacyjne progowanie reprezentacji czasowo-częstotliwościowej .....	143
7.2. Modelowanie tła EKG z ekstrapolacją szumu izolinii .....	145
7.2.1. Wprowadzenie .....	145
7.2.2. Szczegóły implementacji .....	146
7.2.3. Własności algorytmu odszumiania z ekstrapolacją szumu izolinii .....	150

<b>Zakończenie</b> .....	152
--------------------------	-----

<b>Dodatek</b> .....	154
----------------------	-----

D1. Próbkowanie, ograniczanie i warianty transformacji Fouriera .....	154
---	-----

D2. Transformacja Karhunen–Loevego .....	155
--	-----

D3. Dyskretna transformacja kosinusowa .....	158
--	-----

<b>Bibliografia</b> .....	159
---------------------------	-----



*Mojej Mamie*



## Spis wybranych skrótów

- AZTEC – ang. *amplitude zone time epoch coding*  
CORTES – ang. *coordinate reduction time encoding system*  
CP – ang. *canal parensis* (parametr określający symetrię pobudzenia błęd-  
ników)  
CSE – Common Standard for Quantitative Electrocardiography  
CT – ang. *computerized tomography* (tomografia komputerowa)  
DCT – ang. *discrete cosine transform* (dyskretna transformacja kosinusowa)  
DP – ang. *directional preponderance* (parametr określający przewagę kierun-  
kową oczopląsu)  
DWT – ang. *discrete wavelet transform* (dyskretna transformacja falkowa)  
EEG – elektroencefalogram (elektroencefalografia)  
EKG – elektrokardiogram (elektrokardiografia)  
ENG – elektronystagmogram (elektronystagmografia)  
EZW – ang. *embedded zerotree wavelet* (algorytm wykorzystujący zagnieżdżo-  
ne drzewa zerowe dekompozycji)  
FIR – ang. *finite impulse response* (filtr o skończonej odpowiedzi impulsowej)  
HR-ECG – ang. *high resolution electrocardiography* (elektrokardiografia wysokiej  
rozdzielczości)  
HRV – ang. *heart rate variability* (zmiennosc rytmu serca)  
IIR – ang. *infinite impulse response* (filtr o nieskończonej odpowiedzi impul-  
sowej)  
KLT – ang. *Karhunen–Loeve transform* (transformacja Karhunen–Loevego)  
LSB – ang. *least significant bit* (najmniej znaczący bit)  
MIT-BIH – Massachusetts Institute of Technology & Beth Israel Hospital w Bostonie  
MPEG – Motion Picture Expert Group (komitet standaryzacyjny informacji au-  
diowizualnej)  
NMR – ang. *nuclear magnetic resonance* (tomografia z wykorzystaniem magne-  
tycznego rezonansu jądrowego)  
PRD – ang. *percent root-mean-square difference* (współczynnik odchylenia  
średniokwadratowego)  
QMF – ang. *quadrature mirror filter* (kwadraturowy filtr lustrzany)

- QRSd – ang. *QRS duration* (czas trwania zespołu QRS)
- REM – ang. *rapid eye movements* (fazy snu aktywnego)
- RMS40 – ang. *root-mean-square* (pierwiastek sumy kwadratów sygnału na odcinku 40 ms)
- RR – interwał międzyuderzeniowy serca
- TF – ang. *time-frequency* (czasowo-częstotliwościowy)
- WKG – wektokardiogram (wektokardiografia)
- WNN – ang. *wavelet neural network* (sieć falkowo-neuronowa)
- WV – ang. *Wigner–Ville transform* (transformacja Wignera–Ville’a)



# Wstęp

Jednym z pierwotnych aspektów życia człowieka jest tworzenie. Twórczość w jakiegokolwiek dziedzinie jest wyrazem myśli, koncepcji, a często w jej przetwarzaniu bierze udział rysunek (partytura, projekt) albo słowo (poezja). Z wyjątkiem improwizacji (muzycznych, malarskich) twórczość człowieka charakteryzuje regularność, która jest wyrazem pierwotności zamysłu. Już na etapie koncepcji człowiek stara się zapewnić swemu dziełu określoną trwałość. Przeciwnymi zasadami rządzi się tworzenie w przyrodzie. Tu, na skutek zjawisk podległych prawom fizyki, ale także towarzyszących im anomalii podległych rachunkowi prawdopodobieństwa, powstają formy nieregularne, których trwałość jest weryfikowana w praktyce. Niekiedy jest ona tak znikoma, że aż trudna do zmierzenia, kiedy indziej zaś przekracza znacznie trwałość wytworów człowieka.

Innym, równie istotnym aspektem życia człowieka jest poznanie. Jest ono podstawą lokalizacji w otaczającym świecie, ale dalsze przetworzenie obserwacji umożliwia rozumienie zasad, którym ten świat podlega. Na tym etapie także używane są formy utrwalenia informacji – rysunek, słowo, a dziś, dzięki postępowi technologii, także cyfrowy zapis informacji na nośniku magnetycznym albo optycznym.

Analityczny umysł człowieka starający się wyodrębnić z pola obserwacji tylko najistotniejsze szczegóły, a także ograniczony zasób środków dostępnych do utrwalenia informacji są przyczynami, dla których do opisu nieregularnych form spotkanych w przyrodzie stosowane są metody charakteryzujące się regularnością. Stąd wywodzi się opis różnych zjawisk przyrody realizowany za pomocą matematyki. Opis, na który zgadzamy się przede wszystkim dlatego, że jest podstawą rozumienia praw przyrody, zawiera jednak zwykle tylko wycinek obserwacji, często nie bierze pod uwagę warunków początkowych i jest w rzeczywistości uproszczoną reprezentacją zjawiska. Wiemy wszyscy z doświadczenia, że nawet jabłko z drzewa nie spada nigdy dwa razy dokładnie tak samo.

O ile jednak do opisu działania regularnych wytworów człowieka te zakładające regularność metody okazują się w większości przypadków wystarczające, to jednak opis wytworów przyrody bywa niekiedy uproszczony ponad miarę. W większości przypadków użycie łatwego opisu matematycznego zakładającego stabilność i powtarzal-

ność prowadzi do zgubienia w procesie redukcji informacji szczegółów istotnych dla opisywanego zjawiska, a w konsekwencji – znacznych rozbieżności oczekiwań czy wyobrażeń obserwatora zjawiska od jego rzeczywistego przebiegu. Uwaga ta dotyczy także narządów, których funkcjonowanie może być obserwowane za pośrednictwem sygnałów elektrodiagnostycznych.

Dlatego wraz z postępem technologii pomiarów i metodologii opisu coraz bardziej atrakcyjna staje się rezygnacja z wątpliwego założenia o stacjonarności analizowanego sygnału. Użycie trudniejszych w percepcji form opisu pozwala często odkryć nowe aspekty obserwowanych zjawisk, a niejednokrotnie niestacjonarność staje się kluczem do zrozumienia ich istoty. Przyroda może być obecnie opisywana z użyciem metod charakterystycznych dla jej nieregularności, zamiast tych, których jedyną zaletą jest łatwość interpretacji. Warto w tym miejscu zauważyć przyspieszenie, jakie miało miejsce w rozwoju metod opisu sygnałów – niespełna 40 lat temu wynaleziono szybką transformację Fouriera, która dokonuje dekompozycji na bazie nieskończonych funkcji sinusoidalnych, 20 lat temu użyto opisu sygnału z wykorzystaniem funkcji o pojedynczym zafalowaniu, a już dziś stosowane są techniki pogoni za dopasowaniem, w których dekompozycja oparta jest na poszukiwanych fragmentach sygnału.

Sygnały elektrodiagnostyczne są częścią opisu przyrody, bardziej jednak istotną od opisu otaczającego świata, bowiem reprezentują zjawiska zachodzące w funkcjonujących narządach człowieka. Tak więc ciekawość, z jaką człowiek traktuje sygnały elektrodiagnostyczne, jest częścią zainteresowania samym sobą. Sygnały te, istniejące w sposób nie absorbujący uwagi przez całe życie człowieka, mogą w pewnych sytuacjach odegrać zasadniczo decydującą rolę. Poprawna interpretacja zawartych w nich informacji o pracy poszczególnych narządów jest bowiem warunkiem skutecznej terapii.

Warto jeszcze zdać sobie sprawę z roli, jaką współcześnie odgrywa elektrodiagnostyka w medycynie. Zważywszy, że przyczyną wszelkich akcji komórek jest zmiana własności elektrycznych ich błony (depolaryzacja), łatwo dojść do wniosku, że zjawiska elektryczne reprezentują przesyłanie informacji w organizmie. Mimo że w istocie nie jest to jedyny „kanał informacyjny” (istnieją jeszcze m.in. hormony), wiele informacji przekazywanych sobie przez narządy udaje się przechwycić metodami elektrodiagnostycznymi. Stąd już tylko krok, chociaż nie zawsze łatwy, do identyfikacji zaburzeń. Pomimo gwałtownego rozwoju innych metod diagnostyki medycznej, z diagnostyką obrazową na czele, oraz występujących niekiedy trudności pomiarowych i interpretacyjnych informacje pozyskane drogą elektryczną są wciąż cenione przez lekarzy. Powodem są przede wszystkim dwie cechy elektrodiagnostyki:

- nieinwazyjność, która czyni diagnozę powtarzalną wiele razy;
- niski koszt, który czyni diagnozę powszechnie dostępną.

Głównym celem autora jest przedstawienie wybranych problemów elektrodiagnostyki medycznej i sposobów, w jaki udało się zautomatyzować ich rozwiązywanie za pomocą metod czasowo-częstotliwościowych.

Przedstawiono kilkanaście algorytmów automatycznie przetwarzających różne sygnały elektrodiagnostyczne o szerokim zakresie zastosowań, do którego m.in. należy:

- automatyczna analiza i poszukiwanie cech,
- filtracja czasowo-częstotliwościowa sygnału,
- kompresja sygnału,
- dyskryminacja szumu.

Intencją autora jest zainspirowanie obecnych i przyszłych projektantów i użytkowników metod czasowo-częstotliwościowych. Dlatego każdy opis zawiera sformułowanie problemu diagnostycznego oraz przykład jego rozwiązania poparty omówieniem podstawowych argumentów projektowych. Oczywiście, nie leży w zamyśle autora próba wyczerpania możliwych rozwiązań ani tym bardziej możliwych zastosowań. Tych ostatnich zresztą przybywa z dnia na dzień, i oby wśród autorów przyszłych rozwiązań znalazł się ktoś z grona czytelników tego skromnego przeglądu.

Jak wykazują liczne doniesienia z innych dziedzin nauki, metody czasowo-częstotliwościowe, będące jeszcze na początku lat 90. XX wieku swoistą ciekawostką matematyczną, obecnie uzyskały prawo bytu na równi z takimi kanonami analizy sygnałów, jak przekształcenie Fouriera czy funkcja korelacji wzajemnej. Szeroki zakres ich stosowania obejmuje rozległe dziedziny wiedzy: od sejsmografii i radiolokacji (echolokacji) poprzez akustyczną diagnostykę maszyn w ruchu, rozpoznawanie mowy, przetwarzanie obrazów (także tych medycznych) aż po modelowanie turbulencji aero- i hydrodynamicznych w nowych konstrukcjach inżynierii cywilnej. Warto także dodać, że metody czasowo-częstotliwościowe nie są zarezerwowane wyłącznie dla bardzo wyszukanych dziedzin nauki i techniki. Już dziś istnieją karty przechwytywania sygnału wideo przeznaczone do komputerów osobistych, w których kompresja obrazu oparta jest na transformacji falkowej, a jutro być może kompresja taka zawita pod przysłowiowe strzechy w postaci milionów krążków nagranych w formacie MPEG-4. A jeśli kogoś nie przekonują te przykłady reprezentujące świat wysokich technologii, niech spojrzy w nuty. Czyż nie jest to zapis czasowo-częstotliwościowy?

Nie jest to kolejna książka o falkach, bogata literatura tego przedmiotu, także w języku polskim, jest już bowiem powszechnie dostępna. Zawartość trzech pierwszych rozdziałów ma raczej charakter informacyjny, służy także ujednoliceniu pojęć i notacji. Niemniej czytelnik pragnący pobieżnie przeglądnąć własności przekształceń czasowo-częstotliwościowych, może te informacje potraktować jako dalsze wskazówki do studiowania, eksperymentów i poszukiwań literaturowych.

Niektóre z zaprezentowanych zastosowań metod czasowo-częstotliwościowych w elektrodiagnostyce medycznej są oryginalnymi rozwiązaniami autora. Opisy pozostałych są oparte na powszechnie dostępnej literaturze naukowej, dołożono wszelkich

starań, aby niniejszy opis odpowiadał intencjom ich autorów. Spośród zapożyczonych z literatury przykładów wybrano tylko takie, które prezentują przełomowe rozwiązania istotnie usprawniające automatyczne przetwarzanie sygnałów oraz takie, których zastosowanie pokazuje nowy zakres możliwości analizy sygnałów elektrofizjologicznych w dziedzinie czasowo-częstotliwościowej.



## ROZDZIAŁ 1

# Pojęcia podstawowe

### 1.1. Dyskretna reprezentacja sygnałów

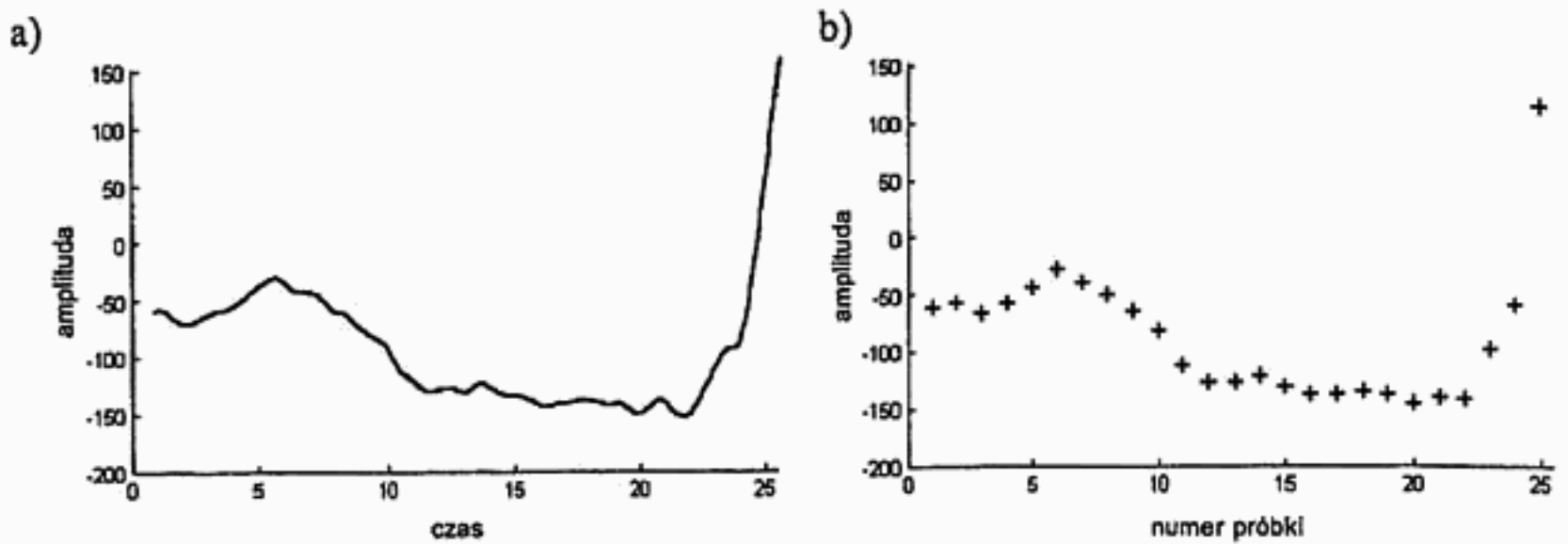
Próbkowanie sygnału analogowego dokonywane przez przetwornik analogowo-cyfrowy jest związane z przekształceniem niektórych własności sygnału analogowego w odmienne własności sygnału dyskretnego. Odmienne własności sygnału analogowego oraz jego cyfrowej reprezentacji mogą być jednym z decydujących czynników konstrukcyjnych – zwłaszcza dotyczy to wyboru sprzętowej lub programowej realizacji niektórych etapów przetwarzania sygnałów elektrodiagnostycznych.

Do często przyjmowanych domyślnie (ale tutaj wymagających wyeksponowania) należą następujące założenia o dyskretyzacji sygnału ciągłego:

- próbkowanie sygnału analogowego jest dokonywane ze stałą precyzją amplitudową (choć niekiedy spotyka się systemy o zmiennej precyzji próbkowania);
- próbkowanie sygnału analogowego jest dokonywane ze stałym interwałem czasowym (choć niekiedy spotyka się systemy o zmiennej częstotliwości próbkowania – jest to wielce obiecująca, nowa gałąź nauki).

Warte uwagi jest przechowanie oprócz wynikowego ciągu liczb naturalnych także wartości interwału (lub częstotliwości) próbkowania i precyzji dyskretyzacji amplitudy [ $\mu\text{V}/\text{LSB}$ ] (ang. *least significant bit*). Są one jedynymi łącznikami pomiędzy ciągiem liczb niemianowanych a światem jednostek fizycznych (ms, mV itp.). Znane są przypadki, gdy zagubienie lub pomylenie jednej z tych liczb spowodowało, że setki megabajtów zmudnie zarejestrowanych sygnałów stały się bezużyteczne.

Proces dyskretyzacji sygnału analogowego (rys. 1.1) potocznie kojarzony jest przede wszystkim z kwantyzacją wartości amplitudy, tj. z przedstawieniem ciągłej wartości wejściowej za pomocą jednej z wartości ze zbioru skończonego (o liczbie elementów zdefiniowanej przez rozdzielczość przetwornika). Wielkość wejściowa może przyjmując dowolną wartość pomiędzy minimum i maksimum, liczba możliwych wartości jest więc nieskończona. Wyjściowa wielkość dyskretna to najmniej odległa od wielkości wejściowej cyfrowa jej reprezentacja. Niedoskonałość tego odwzorowania jest nazywana szumem kwantyzacji.

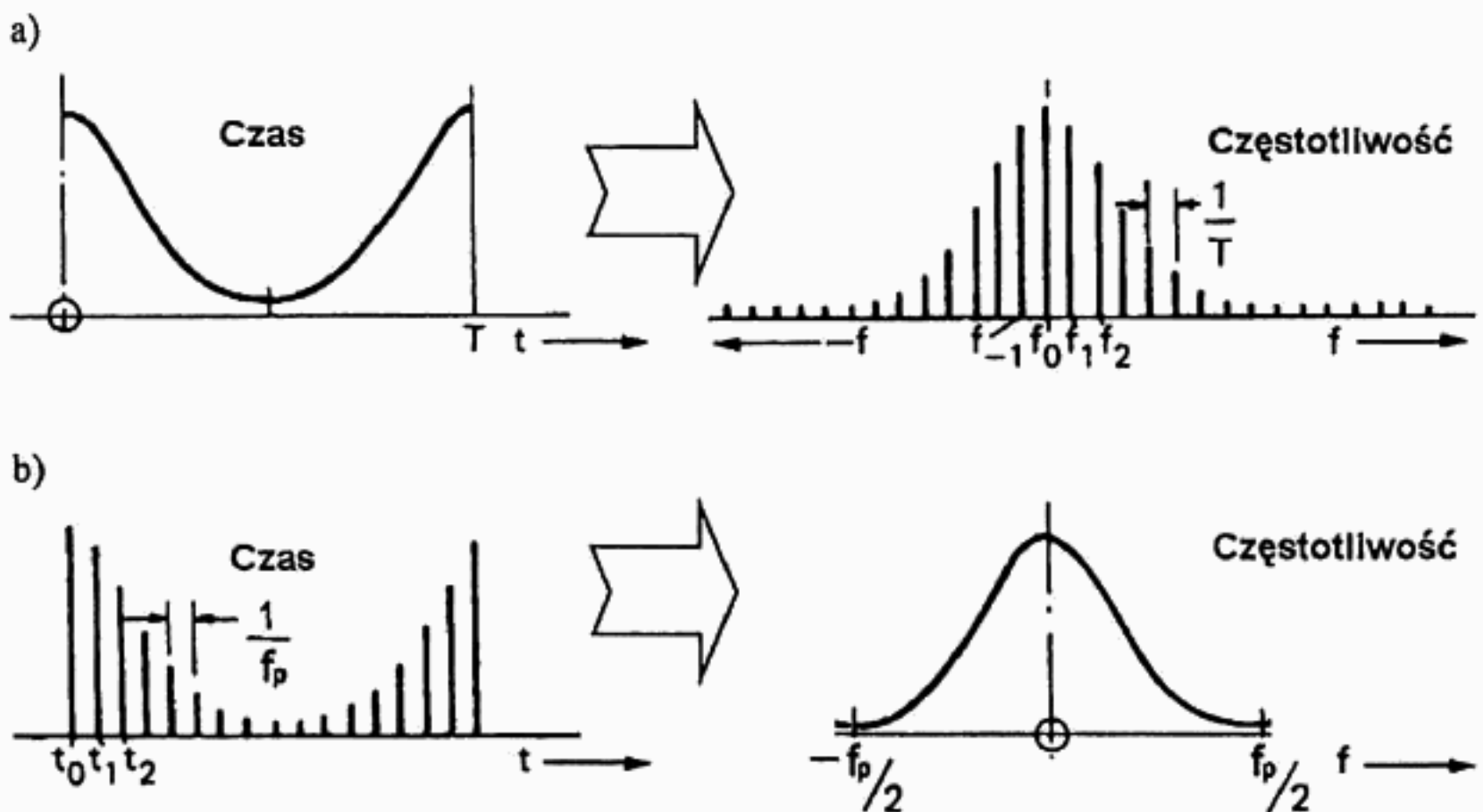


Rys.1.1. Sygnał analogowy (a) i jego dyskretna reprezentacja (b) w dziedzinie czasu

Dyskretyzacja sygnału analogowego to także dwa inne istotne zjawiska w dziedzinie czasu (rys. 1.2, dodatek D1):

- próbkowanie (dyskretyzacja w dziedzinie czasu) to zmiana ciągłej funkcji sygnału analogowego w funkcję istniejącą tylko w chwilach próbkowania, pomiędzy tymi chwilami przebieg reprezentacji sygnału jest nieznany;
- ograniczanie to zmiana nieograniczonej w dziedzinie czasu funkcji sygnału analogowego w funkcję ograniczoną długością pomiaru.

Z punktu widzenia teorii i analizy sygnałów, zarówno próbkowanie, jak i ograniczanie jest utratą informacji występujących być może w reprezentowanym sygnale.



Rys. 1.2. Ilustracja graficzna zjawisk ograniczania i próbkowania sygnału. Zwraca uwagę wzajemna odwrotność opisanych zjawisk: rezultatem ograniczania sygnału jest dyskretyzacja widma (a), natomiast rezultatem próbkowania (dyskretyzacji) sygnału jest ograniczenie widma (b)



Próbkowanie jest spowodowane niewykonalnością techniczną odwzorowania sygnału w sposób ciągły. Najszybsze aktualnie dostępne przetworniki oferują zawrotną częstotliwość próbkowania 200 Gs/s (interwał próbkowania równy 5 ps), nie jest to jednak próbkowanie ciągłe – cyfrowa reprezentacja sygnału nadal jest dyskretna w dziedzinie czasu, czyli nieokreślona dla przedziałów czasu pomiędzy momentami próbkowania. W rzeczywistości istnieją także sygnały z natury dyskretne, wielu przykładów dostarczają zwłaszcza nauki ekonomiczne (np. indeksy giełdowe).

Ograniczanie jest spowodowane niewykonalnością techniczną odwzorowania sygnału w sposób dowolnie długi. Z punktu widzenia procesu pomiarowego, sygnału źródła istnieje od zawsze (od  $-\infty$ ) i w nieskończoność (do  $+\infty$ ). Rzeczywiście, 10-sekundowy zapis elektrokardiogramu istniejącego u źródła przez ok. 75 lat dobrze przybliża matematyczny ideał nieskończoności. Zjawisko ograniczania jest spowodowane skończoną pojemnością nośników informacji cyfrowej (pamięci, taśm, dysków itp.).

Przetwarzanie cyfrowych reprezentacji sygnałów jest już obecnie dziedziną bardzo rozległą. Zazwyczaj, w klasycznym ujęciu tego terminu, obejmuje ono:

- realizacje filtrów cyfrowych różnych typów (dolnoprzepustowych, górnoprzepustowych, pasmowoprzepustowych i pasmowozaporowych), według różnych algorytmów (FIR, IIR) oraz o różnych charakterystykach (Butterwortha, Czebyszewa, ...);
- uśrednianie próbek (należących do tego samego sygnału) i uśrednianie sygnałów, często stosowane są współczynniki wagowe różnicujące wpływ poszczególnych składników na rezultat uśredniania;
- całkowanie i różniczkowanie sygnałów;
- aproksymacja z wykorzystaniem różnych funkcji (wielomianów, funkcji trygonometrycznych, funkcji sklepanych itd.);
- przekształcenia widmowe (przekształcenie Fouriera), także z wykorzystaniem widm wyższych rzędów;
- badanie współzmienności sygnałów metodami regresji, korelacji wzajemnej, funkcji koherencji itd...

Przedstawione powyżej techniki doczekały się wielu opracowań zarówno na poziomie teoretycznym, jak i implementacji we wszystkich chyba językach programowania. Czytelnik zainteresowany szczegółami z łatwością znajdzie odpowiadającą mu pozycję.

## 1.2. Analiza sygnałów w dziedzinie częstotliwości

Spośród metod opisu sygnałów bardzo użytecznym narzędziem okazały się metody transformacyjne pozwalające przedstawić sygnał w dziedzinie częstotliwości i zbadać jego właściwości widmowe. Analizy częstotliwościowe polegają na przedstawie-

niu badanego sygnału za pomocą kombinacji liniowej pewnych bazowych sygnałów elementarnych, które zazwyczaj charakteryzują się ortogonalnością względem siebie.

$$\begin{aligned}x(t) &= \sum_k a_k g_k(t) \\ a_k &= \int x(t) g_k^*(t) dt\end{aligned}\tag{1.1}$$

gdzie:

- $x(t)$  – analizowany sygnał;
- $g_k(t)$  –  $k$ -ty sygnał bazowy;
- $a_k(t)$  – współczynnik określający intensywność występowania w badanym sygnale bazy  $g_k(t)$ , tzn. wyznaczający stopień podobieństwa między tymi sygnałami;
- $g_k^*(t)$  – wartość sprzężona do  $g_k(t)$ .

Ideą metod częstotliwościowych jest minimalizacja niezerowych współczynników  $a_k$ , czyli minimalizacja informacji potrzebnej do opisu sygnału. Zatem bardzo ważnym czynnikiem jest wybór odpowiedniego zestawu funkcji bazowych. Im kształt funkcji bazowych będzie bardziej dopasowany do analizowanego sygnału, tym mniej tych funkcji będzie potrzeba do jego aproksymacji. W efekcie pociąga to za sobą większą liczbę zerowych współczynników  $a_k$  oraz stosunkowo małą liczbę „istotnych” współczynników reprezentujących badany przebieg. Opierając się na powyższym można stwierdzić, że sygnały stacjonarne powinny być analizowane funkcjami bazowymi o charakterze stacjonarnym. Takie są np. funkcje sinus i cosinus doskonale zlokalizowane w dziedzinie częstotliwości (nieskończony nośnik) – nadają się one do opisu przebiegów okresowych. Natomiast dla sygnałów niestacjonarnych należy wybrać funkcje o dobrej lokalizacji przestrzennej (czasowej), czyli o skończonym (selektywnym) nośniku.

### 1.2.1. Analiza częstotliwościowa sygnałów stacjonarnych

Najpopularniejszą, choć nie jedyną, metodą analizy sygnałów o własnościach niezmiennych w czasie jest przekształcenie Fouriera oraz jego wersje dla sygnałów dyskretnych. Stosowana baza dekompozycji jest zbiorem spiral w dziedzinie zespolonej, charakteryzowanych przez skok w dziedzinie czasu odpowiadający okresowi analizowanej częstotliwości. Baza taka spełnia warunek ortonormalności, dzięki czemu istnieje transformacja odwrotna, a złożenie prostej i odwrotnej transformacji Fouriera spełnia warunki perfekcyjnej rekonstrukcji sygnału. Jednocześnie sinusoidalny charakter funkcji bazowych najlepiej przybliża intuicyjne pojęcie częstotliwości, gdyż reprezentuje drgania występujące w układach najczęściej spotykanych w warunkach eksperymentalnych.

Przekształcenie Fouriera dla sygnałów ciągłych jest zdefiniowane parą równań:

$$\begin{aligned} G(f) &= \int_{-\infty}^{\infty} g(t) \cdot e^{-j \cdot 2\pi f t} dt \\ g(t) &= \int_{-\infty}^{\infty} G(f) \cdot e^{j \cdot 2\pi f t} df \end{aligned} \quad (1.2)$$

gdzie  $g(t)$  jest ciągłym nieograniczonym sygnałem w dziedzinie czasu, a  $G(f)$  jego reprezentacją w dziedzinie częstotliwości (widmem zespolonym), natomiast  $j = \sqrt{-1}$  jest wersorem osi urojonej.

Sygnały dyskretne, a tylko takie ma do dyspozycji użytkownik, który dokonał rejestracji za pomocą przetwornika A/C, są reprezentowane przez ciąg próbek. Są to liczby reprezentujące wielkość fizyczną, o której informację zawiera sygnał (np. napięcie). Domyślnie zakłada się, że próbki pobierane są co stały okres zwany interwałem próbkowania. Zamiast uzupełniać każdą wartość informacją o czasie, w którym została pobrana, wystarczy podać wartość częstotliwości próbkowania. Zgodnie z twierdzeniem Nyquista częstotliwość próbkowania powinna być dwa razy wyższa od maksymalnej częstotliwości występującej w sygnale. Niespełnienie tego warunku prowadzi do zjawiska aliasingu, w wyniku którego reprezentacja dyskretna zawiera komponenty nie występujące w sygnale oryginalnym, a więc jest zniekształcona. Ponieważ w praktyce w rejestrowanym sygnale mogą pojawić się dowolne składowe, konieczne jest stosowanie filtrów dolnoprzepustowych nawet wtedy, gdy częstotliwość próbkowania przewyższa dwukrotność maksymalnej częstotliwości spodziewanego sygnału użytecznego.

Dyskretna reprezentacja sygnału nie jest już funkcją ciągłą, toteż przekształcenie Fouriera dla sygnałów dyskretnych zdefiniowane jest wzorem, w którym całkę zastąpiono sumą wszystkich elementów reprezentacji:

$$\begin{aligned} G(f) &= \sum_{n=-\infty}^{\infty} g(t_n) \cdot e^{-j \cdot 2\pi f t_n} \\ g(t_n) &= \frac{1}{t_s} \int_{-\frac{t_s}{2}}^{\frac{t_s}{2}} G(f) \cdot e^{j \cdot 2\pi f t_n} df \end{aligned} \quad (1.3)$$

gdzie  $g(t_n)$  jest dyskretnym nieograniczonym sygnałem w dziedzinie czasu, a  $G(f)$  jego ciągłą ograniczoną reprezentacją w dziedzinie częstotliwości.

Sygnały zarejestrowane przy użyciu urządzeń pomiarowych są wycinkami sygnałów rzeczywistych, co wynika zwykle z ograniczonej pojemności nośnika informacji.

Okrojone sygnały nie są więc nieskończonymi funkcjami czasu, a zastosowanie przekształcenia Fouriera jest możliwe dzięki powszechnie stosowanemu fortelowi. Polega on na postrzeganiu posiadanej porcji sygnału jako okresu podstawowego pewnej nieskończonej funkcji okresowej. Oczywiście, przejście pomiędzy „okresami” tak powstałej funkcji musi być ciągłe, a więc dla wszystkich komponentów sygnału faza na początku „okresu” musi być identyczna jak na końcu. W tej sytuacji w sygnale mogą występować tylko komponenty o częstotliwości będącej wielokrotnością pewnej częstotliwości podstawowej, będącej odwrotnością długości zarejestrowanego sygnału. Reprezentacja okrojonego sygnału w dziedzinie częstotliwości składa się z izolowanych próbek, jest więc dyskretna. Nazywana jest ona szeregiem Fouriera, a wzór stanowiący definicję przekształcenia odwrotnego zastępuje całkowanie operacją sumy:

$$G(f_k) = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} g(t) \cdot e^{-j \cdot 2\pi f_k t} dt \quad (1.4)$$

$$g(t) = \sum_{k=-\infty}^{\infty} G(f_k) \cdot e^{-j \cdot 2\pi f_k t}$$

gdzie  $g(t)$  jest ciągłym ograniczonym sygnałem w dziedzinie czasu, a  $G(f_k)$  jego dyskretną nieograniczoną reprezentacją w dziedzinie częstotliwości.

W praktycznych realizacjach cyfrowych systemów pomiarowych reprezentacja sygnału jest zarówno dyskretna, jak i ograniczona. Dlatego najczęściej wykorzystywanym w praktyce przekształceniem jest tzw. dyskretna transformacja Fouriera, w której definicji obie operacje całkowania są zastąpione sumowaniem. Wynikiem transformacji jest szereg Fouriera dla sygnałów dyskretnych. Zarówno w dziedzinie czasu, jak i w dziedzinie częstotliwości reprezentacja sygnału jest dyskretna i ograniczona.

$$G(f_k) = \frac{1}{N} \sum_{n=0}^{N-1} g(t_n) \cdot e^{-j \cdot \frac{2\pi n k}{N}} \quad (1.5)$$

$$g(t_n) = \sum_{k=0}^{N-1} G(f_k) \cdot e^{j \cdot \frac{2\pi n k}{N}}$$

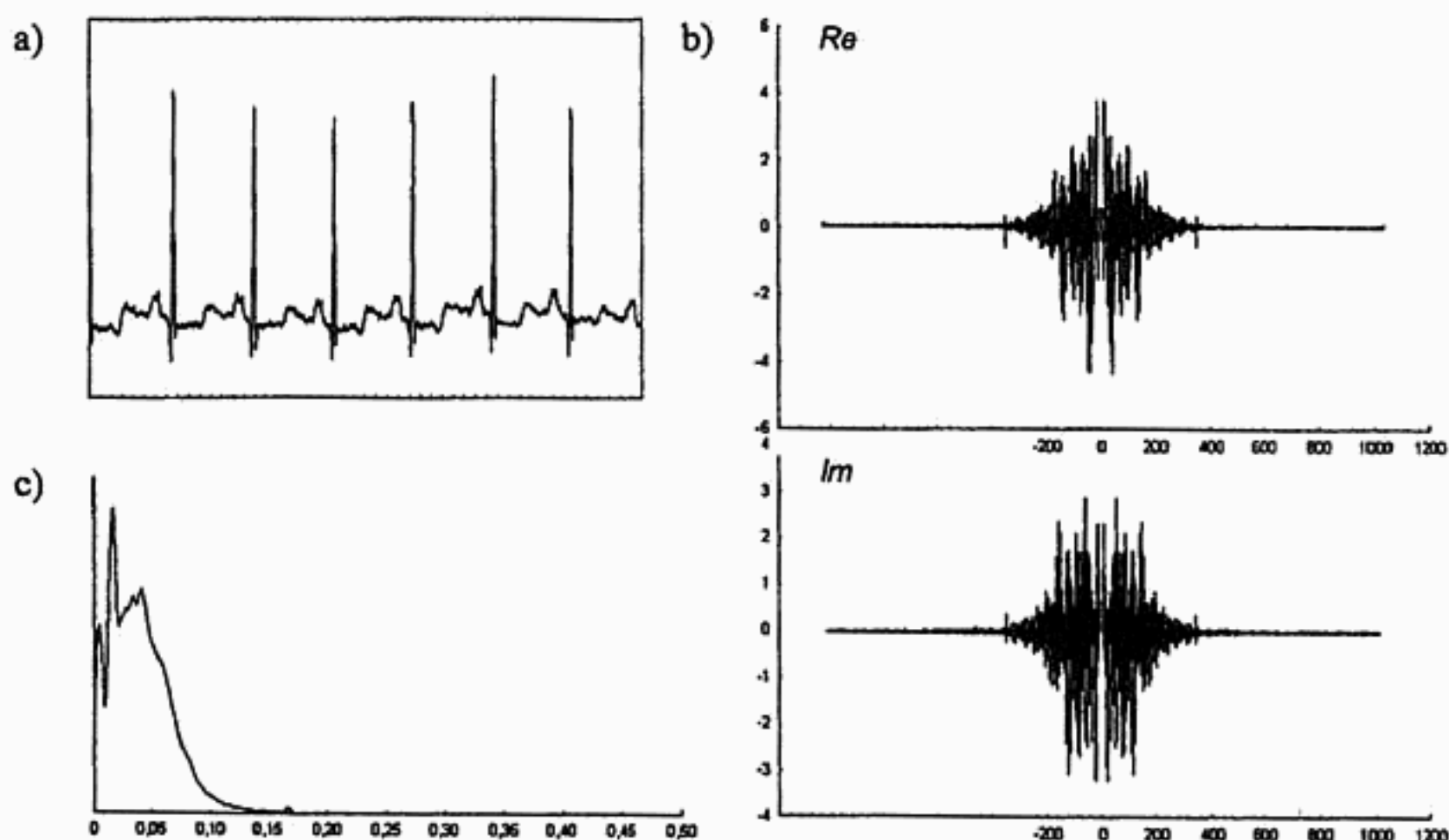
gdzie  $g(t_n)$  jest dyskretnym ograniczonym sygnałem w dziedzinie czasu, a  $G(f_k)$  jego dyskretną ograniczoną reprezentacją w dziedzinie częstotliwości.

Z powyższych wzorów wynika bardzo ciekawa własność: sygnał dyskretny w dziedzinie czasu oraz jego reprezentacja dyskretna w dziedzinie częstotliwości skła-



dają się z takiej samej liczby próbek – jest to jeden z objawów własności perfekcyjnej rekonstrukcji sygnału. Podczas gdy wartości sygnału są zwykle liczbami czysto rzeczywistymi (a nawet całkowitymi, pochodzą bowiem z przetwornika A/C), wartości widma są liczbami zespolonymi. Konsekwencją czysto rzeczywistego charakteru sygnału w dziedzinie czasu jest symetria jego widma, co zwykle pozwala analizować jego zakres w przedziale  $\langle 0 \div f_s/2 \rangle$ .

W przypadku gdy analiza widmowa prowadzona jest w celu identyfikacji komponentów sygnału lub określenia ich wzajemnego udziału w całkowitej energii sygnału, zamiast widma zespolonego chętniej wykorzystywane jest tzw. widmo mocy sygnału będące modulem widma zespolonego. Wygładzone widmo mocy jest estymatorem częstotliwościowego rozkładu energii sygnału, dlatego nosi nazwę funkcji gęstości mocy (rys. 1.3). Wygładzenie widma mocy powoduje już jednak zmianę jego zawartości informacyjnej, a więc odtworzenie sygnału na podstawie funkcji gęstości mocy nie jest możliwe.

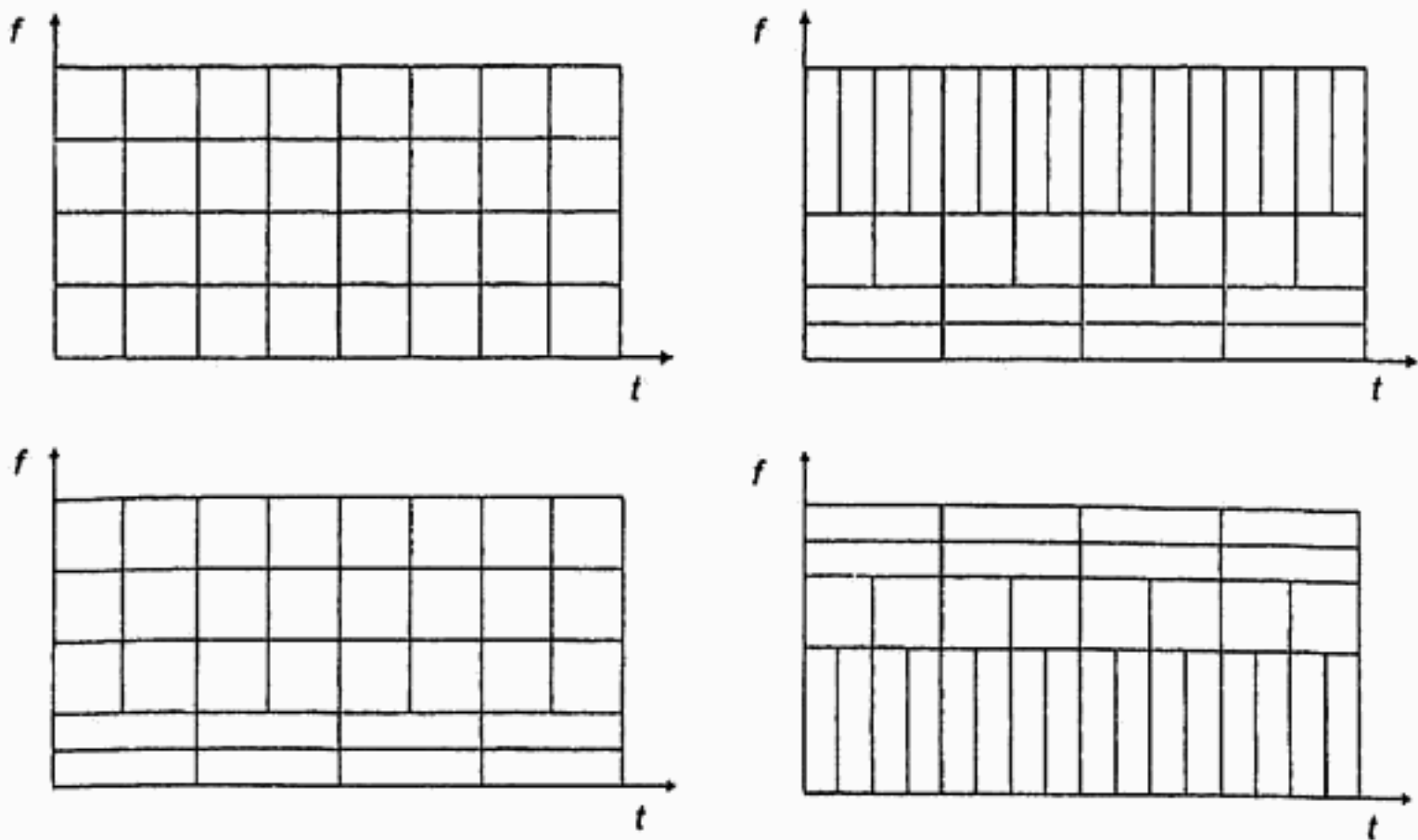


Rys. 1.3. Analiza częstotliwościowa elektrokardiogramu: a) szereg czasowy; b) transformata Fouriera; c) funkcja gęstości mocy

### 1.2.2. Analiza częstotliwościowa sygnałów niestacjonarnych

Analiza częstotliwościowa sygnałów niestacjonarnych powinna być prowadzona przy użyciu bazy dekompozycji złożonej z funkcji selektywnych w dziedzinie czasu i częstotliwości. Identyfikacja składowych (dekompozycja) sygnału jest dokonywana za pomocą przesuwania funkcji analizującej w dziedzinie czasu oraz przestrajania w dziedzinie częstotliwości. W praktyce stosowane są przebiegi impulsowe, których

część w dziedzinie czasu o istotnych wartościach wyznacza czasowy zakres analizy, a część widma o istotnych wartościach – częstotliwościowy zakres analizy. Najmniejszy możliwy do zidentyfikowania komponent sygnału, zwany także atomem z racji dalszej niepodzielności, odpowiada swoimi rozmiarami w dwuwymiarowej dziedzinie czasowo-częstotliwościowej prostokątowi, którego boki stanowią czasowy i częstotliwościowy zakres analizy funkcji bazowej (rys. 1.4).



Rys. 1.4. Przykłady konfiguracji atomów sygnału w dziedzinie czasowo-częstotliwościowej

Dążenie do jak najlepszej selektywności w obu dziedzinach, rozumiane intuicyjnie jako minimalna powierzchnia atomu sygnału, jest ograniczone zasadą nieoznaczoności Heisenberga

$$\sigma^2 = A^2 + B^2 \geq 2 \cdot A \cdot B \geq \frac{1}{2\Pi} \quad (1.6)$$

gdzie:

$$A^2 = \frac{1}{E} \int_{-\infty}^{\infty} t^2 |g(t)|^2 dt, \quad A \text{ oznacza średniokwadratową szerokość czasową,}$$

$$B^2 = \frac{1}{E} \int_{-\infty}^{\infty} f^2 |G(f)|^2 df, \quad B \text{ oznacza średniokwadratową szerokość częstotliwościową,}$$

$$E = \int_{-\infty}^{\infty} |g(t)|^2 dt = \int_{-\infty}^{\infty} |G(f)|^2 df.$$



Jako ciekawostkę należy przytoczyć, że równość we wzorze (1.6) jest spełniona tylko dla funkcji Gaussa

$$g(t) = (2 \cdot \alpha)^{\frac{1}{4}} \exp(-\alpha \Pi t^2) \quad (1.7)$$

Funkcja ta zapewnia minimalną powierzchnię atomu sygnału w dziedzinie czaso-częstotliwościowej jest więc optymalna z punktu widzenia precyzji lokalizacji komponentów.

Zależnie od celu analizy precyzja lokalizacji komponentów sygnału w dziedzinie czasu może być istotniejsza od precyzji w dziedzinie częstotliwości lub na odwrót. Wyboru można dokonać poprzez przeskalowanie funkcji z bazy dekompozycji (np. zmianę parametru  $\alpha$  funkcji Gaussa) wzdłuż osi czasu. Podczas gdy inne własności pozostają niezmiennicze, wydłużenie nośnika w dziedzinie czasu spowoduje wzrost precyzji lokalizacji komponentów w dziedzinie częstotliwości, podczas gdy skrócenie nośnika w dziedzinie czasu, związane ze wzrostem czasowej precyzji lokalizacji, spowoduje proporcjonalne wydłużenie widma funkcji analizujących. W wielu zastosowaniach kryterium doboru długości okna funkcji analizującej jest uzyskanie możliwie najlepszego dopasowania do własności analizowanego sygnału. Reprezentacja takiego sygnału będzie zawierała minimalną liczbę współczynników o istotnych wartościach.

Transformacja Gabora przedstawia ciągły sygnał  $x(t)$  jako kombinację liniową funkcji bazy dekompozycji  $g_{m,n}(t)$  i odpowiadających im współczynników rozkładu  $c_{m,n}$  (rys. 1.5)

$$x(t) = \sum_{m,n=-\infty}^{\infty} c_{m,n} \cdot g_{m,n}(t) \quad (1.8)$$

gdzie  $g_{m,n}(t)$  reprezentuje funkcję z bazy dekompozycji przesuniętą w czasie o  $mT$  oraz przestrojoną w częstotliwości o  $nF$

$$g_{m,n}(t) = g(t - mT) \cdot e^{j \cdot 2 \Pi n F t} \quad (1.9)$$

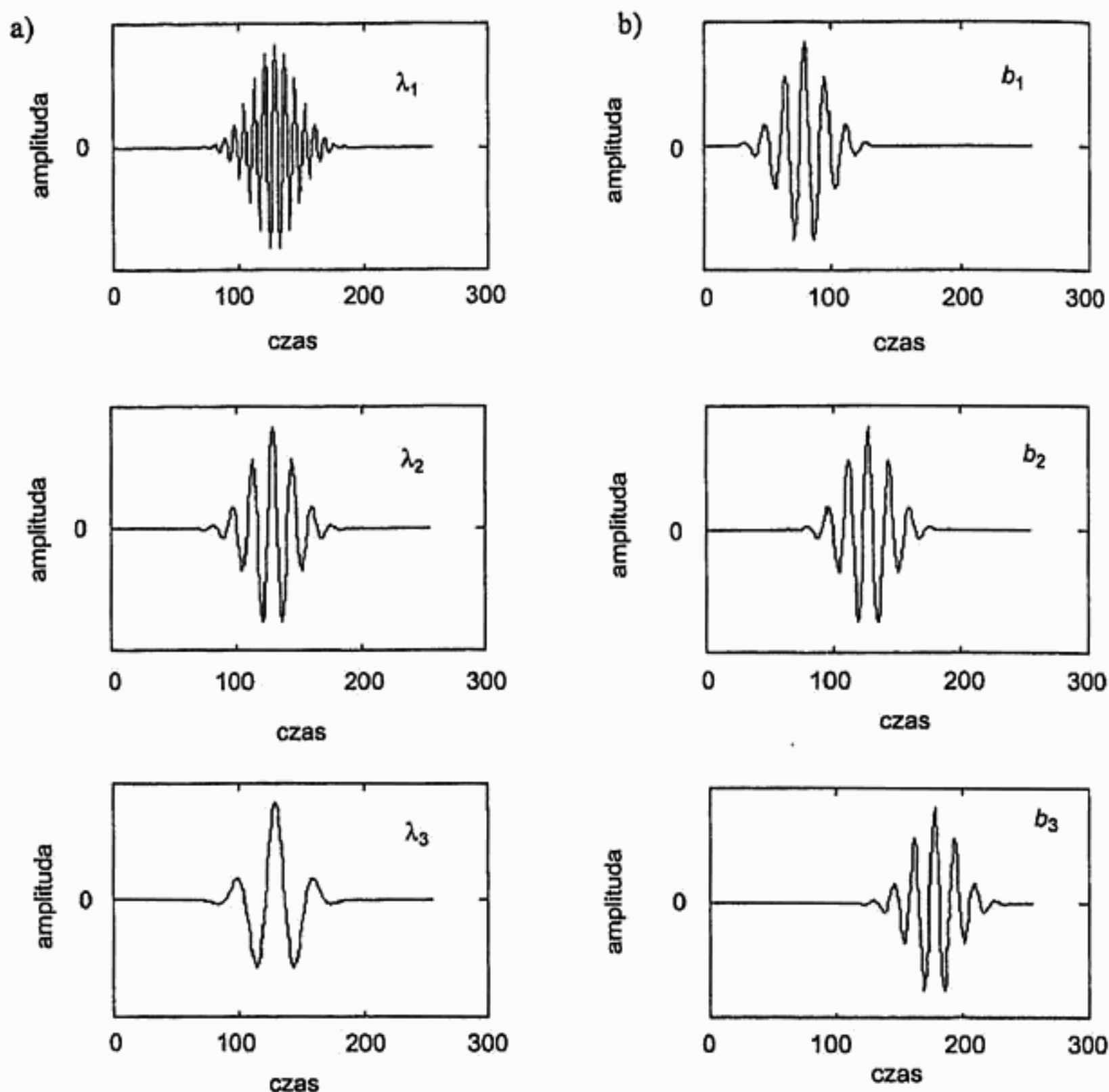
W roli funkcji bazowych  $g(t)$  mogą być użyte różne funkcje impulsowe, jednakże dla swych optymalnych własności funkcja Gaussa, zaproponowana oryginalnie przez Gabora, jest powszechnie używana w typowych zastosowaniach. Współczynniki dekompozycji Gabora, wyznaczające „udział” poszczególnych funkcji bazowych w sygnale  $x(t)$  można określić ze wzoru

$$c_{m,n}(t) = \int_{-\infty}^{\infty} x(t) \cdot \gamma_{m,n}^*(t) dt \quad (1.10)$$

gdzie  $\gamma_{m,n}(t)$  są funkcjami analizującymi danymi wzorem

$$\gamma_{m,n}(t) = \gamma(t - mT) \cdot e^{j \cdot 2 \Pi n F t} \quad (1.11)$$

natomiast funkcja  $\gamma(t)$  jest oknem biortogonalnym do  $g(t)$ .



Rys. 1.5. Funkcje bazowe analizy Gabora: a) dla kolejnych wartości współczynnika skali  $\lambda$ ; b) dla kolejnych wartości współczynnika przesunięcia  $b$

Krótkookresowa transformacja Fouriera jest alternatywną metodą analizy częstotliwościowej sygnałów niestacjonarnych. W odróżnieniu od transformacji Gabora, analiza i synteza sygnału wykorzystują tę samą funkcję okna. Wzory opisujące przekształcenie proste i odwrotne przypominają, że mamy do czynienia z modyfikacją transformacji Fouriera:

$$\text{STFT}_x(t, f) = \int_{-\infty}^{\infty} x(\tau) \cdot \gamma^*(t - \tau) \cdot e^{-j \cdot 2\pi f \tau} d\tau$$

$$x(t) = \frac{1}{\gamma(0)} \int_{-\infty}^{\infty} \text{STFT}_x \cdot e^{j \cdot 2\pi f t} df$$
(1.12)

Analogicznie do widma mocy, reprezentację udziału poszczególnych komponentów w całkowitej energii sygnału stanowi dwuwymiarowa funkcja zwana spektrogramem, zdefiniowana wzorem

$$S_x(t, f) = |\text{STFT}_x(t, f)|^2 \quad (1.13)$$

Spektrogram jest zwykle uzyskiwany z wykorzystaniem jednego ze standardowych okien: prostokątnego, Hamminga, Hana i innych. Wprowadzenie funkcji okna w dziedzinie czasu wiąże się nierozdzielnie z wystąpieniem większych lub mniejszych oscylacji w dziedzinie częstotliwości, a więc występuje zniekształcenie informacji zaburzające rekonstrukcję sygnału. Niemniej, dla wielu zastosowań analitycznych, a w szczególności do identyfikacji momentów wystąpień i energetycznego udziału podstawowych komponentów sygnału, metoda ta jest powszechnie i z powodzeniem stosowana.

Podstawową niedogodność opisanych dotąd transformacji czasowo-częstotliwościowych, oprócz redundancji lub strat informacji zakłócających rekonstrukcję sygnału, stanowiła stała, raz dobierana szerokość okna analizy.

Problem ma następujące dwa aspekty:

- dla częstotliwości odpowiadającej dolnej krawędzi prostokąta wyznaczającego zasięg analizy w dziedzinie czasowo-częstotliwościowej częstotliwość próbkowania jest optymalna; sygnały o wyższych częstotliwościach są lokalizowane w czasie z wciąż tą samą precyzją, niższą od optymalnej;
- dla wąskich w dziedzinie czasu okien analizy istotne staje się ograniczenie zakresu analizowanych częstotliwości od dołu opisane zasadą nieoznaczoności Heisenberga; znaczna ilość komponentów sygnału, a przeważnie także znaczna część jego energii, pozostaje poza zasięgiem analizy, gdyż okres odpowiadający charakteryzującej je częstotliwości przekracza długość okna wyznaczonego przez czasowy zasięg analizy.

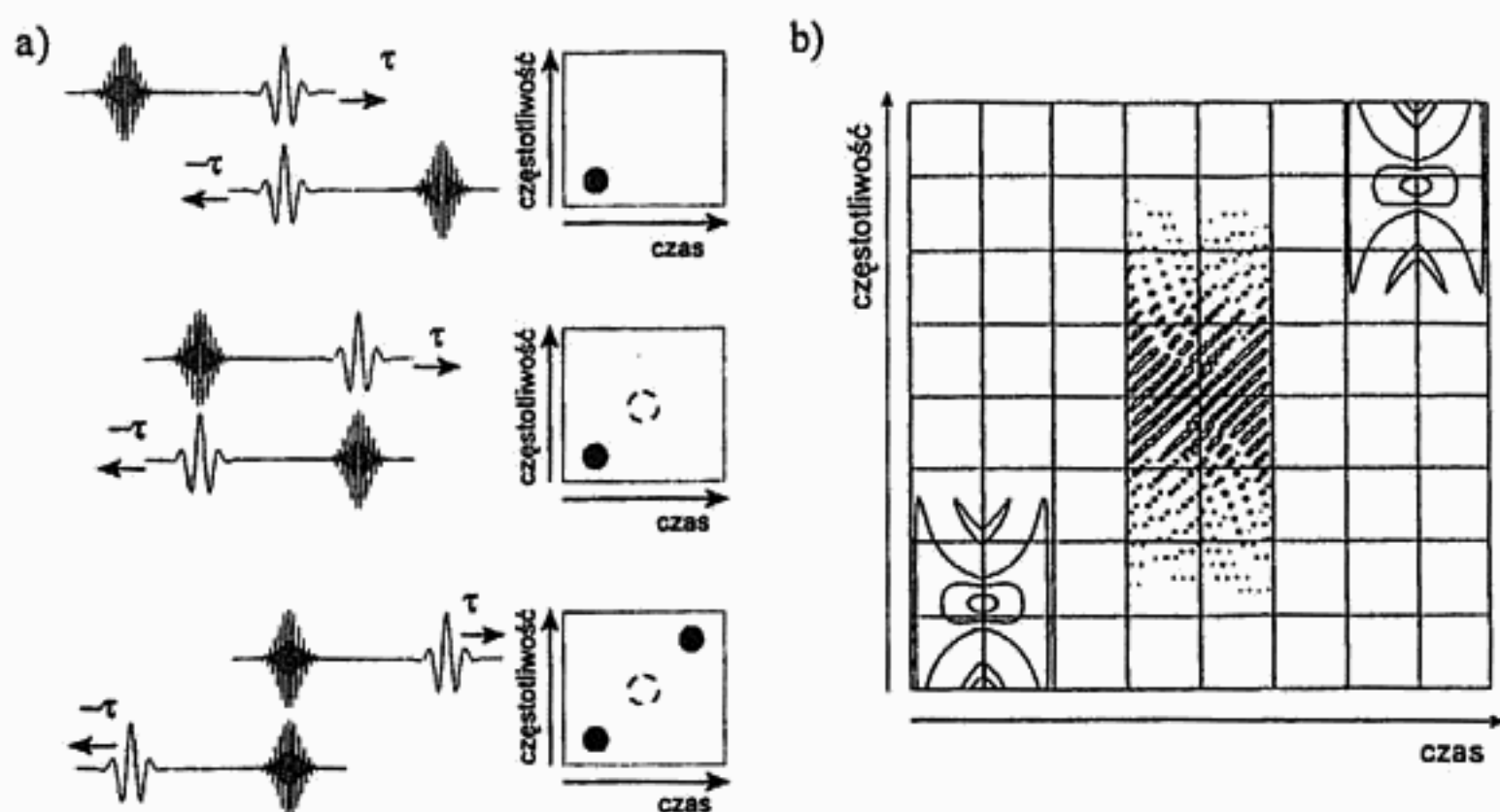
Problem poprawnego doboru rozmiarów okna zmuszał więc do kompromisów pomiędzy zbliżeniem precyzji analizy do teoretycznie możliwej wartości oraz objęciem analizą całego spektrum komponentów występujących (być może) w sygnale, stanowił podstawową wadę transformacji czasowo-częstotliwościowych do czasu zastosowania okna o zmiennej długości. Próbą adaptacyjnego doboru rozmiarów okna do lokalnych własności sygnału stanowi transformacja Wignera–Ville’a. Rolę okna w tej transformacji pełni... sam sygnał (rys. 1.6). Transformacja Wignera–Ville’a jest zdefiniowana wzorem

$$V_x(t, f) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) \cdot x^*\left(t - \frac{\tau}{2}\right) \cdot e^{-j2\pi f\tau} d\tau \quad (1.14)$$

Powyższy wzór prowadzi do wniosku, że jest to transformacja Fouriera sygnału zwanego jądrem Wignera–Ville’a

$$k(t, \tau) = x\left(t + \frac{\tau}{2}\right) \cdot x^*\left(t - \frac{\tau}{2}\right) \quad (1.15)$$

Reprezentacja Wignera–Ville’a rozwiązuje problem doboru długości okna analizy. Wszystkie komponenty występujące w sygnale są analizowane, każdy z nich jest zlokalizowany z optymalną precyzją czasową tak dokładnie, na ile tylko pozwala zasada nieoznaczoności Heisenberga.



Rys. 1.6. Zasada transformacji Wignera–Ville’a: a) koincydencja sygnału i jego lustrzanego odbicia dla kolejnych wartości  $\tau$  i ujawnianie poszczególnych składników sygnału; b) wynikowa reprezentacja czasowo-częstotliwościowa

Reprezentacja Wignera–Ville’a jest biliniowa, jej projekcja na oś czasu jest równa mocy chwilowej sygnału

$$\int_{-\infty}^{\infty} V_x(t, f) df = |x(t)|^2 \quad (1.16)$$

natomiast jej projekcja na oś częstotliwości jest równa widmowej gęstości mocy sygnału

$$\int_{-\infty}^{\infty} V_x(t, f) dt = |X(f)|^2 \quad (1.17)$$

Chociaż transformacja Wignera–Ville’a zapewnia teoretycznie najlepszą precyzję reprezentacji sygnału w dziedzinie czasowo-częstotliwościowej, jej stosowanie do sygnałów analizowanych w praktyce jest ograniczone przez interferencje, które znacznie utrudniają interpretację. Są one spowodowane przez wzajemne oddziaływania pomiędzy różnymi składowymi własnymi widma

$$V_{x+y} = V_x + V_y + 2 \cdot \text{Re}(V_{xy}) \quad (1.18)$$

gdzie

$$V_{xy} = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) \cdot y^*\left(t - \frac{\tau}{2}\right) \cdot e^{-j \cdot 2\pi f t} dt \quad (1.19)$$

jest komponentem reprezentującym interferencje pasożytnicze.

Kompromisem pomiędzy obniżeniem dokładności reprezentacji czasowo-częstotliwościowej a tłumieniem interferencji pasożytniczych jest wygładzanie reprezentacji Wignera–Ville’a. Prowadzi ono do ułatwienia interpretacji czasowo-częstotliwościowego rozkładu energii kosztem utraty niektórych własności transformacji.

## 1.3. Analiza sygnałów w dziedzinie czasowo-częstotliwościowej

### 1.3.1. Transformacja falkowa

Reprezentacja falkowa sygnału jest funkcją dwu zmiennych, stanowiącą kombinację liniową skalowalnej (parametr  $m$ ) i przesuwalnej (parametr  $n$ ) funkcji zwanej falką-matką  $\psi$  (ang. *mother wavelet*)

$$f(x) = \sum_m \sum_n d_{m,n} \psi(2^{-m}x - n) \quad (1.19)$$

Współczynniki  $d_{m,n}$  oblicza się analogicznie jak w przypadku transformaty Fouriera

$$d_{m,n} = 2^m \int f(x) \psi(2^{-m}x - n) dx \quad (1.20)$$

Wzór można interpretować również jako filtrację sygnału analizowanego  $f(x)$  przez sygnał analizujący  $\psi(x, m, n)$ .

Falka, na podstawie której za pomocą skalowania i translacji tworzy się zespół ortogonalnych funkcji bazowych, charakteryzuje się specyficznymi właściwościami: dobrą lokalizacją w czasie (przestrzeni), czyli zwartym nośnikiem funkcji, oraz ograniczonym pasmem w dziedzinie częstotliwości. Dobre umiejscowienie i selektywność w dziedzinie czasu powoduje, że falki doskonale nadają się do aproksymacji sygnałów zmiennych w czasie. Inną istotną cechą reprezentacji falkowej jest to, iż zawiera ona w sobie informacje o sygnale na różnych poziomach szczegółowości. Własność ta, określana mianem hierarchicznej rozdzielczości, jest szczególnie użyteczna w aplikacjach rozproszonych, gdyż umożliwia w naturalny sposób progresywną transmisję polegającą na przesyłaniu najpierw bardzo zgrubnego sygnału, a następnie dosyłaniu jego szczegółów. Poważnym atutem zastosowania transformacji falkowej do analizy sygnałów jest prostota obliczeniowa. Praktyczne wykorzystanie tego mechanizmu sprowadza się do obliczania splotów dwóch wektorów, z których jeden reprezentuje



wartości próbek sygnału badanego, a drugi jest kilku- lub kilkunastoelementowym wektorem reprezentującym filtr cyfrowy. Wspomniane własności falki powodują, że stała się ona „naturalnym” aproksymatorem sygnałów fizycznych. Podczas transformacji znaczna część współczynników  $d_{m,n}$  jest bardzo mała w porównaniu z pozostałymi i może być przyjęta za zero bez zauważalnego zniekształcenia aproksymowanej funkcji. Fakt ten decyduje o dużym znaczeniu transformaty w kompresji sygnałów i obrazów. Także w dziedzinie filtracji szumów falki mają olbrzymie praktyczne zastosowanie. Odpowiednie dobranie falki-matki (służącej do generacji zespołu falek bazowych) zgodnie z charakterem badanego sygnału, pozwala precyzyjnie wyodrębnić właściwy sygnał spośród przypadkowych zakłóceń, które po transformacji będą reprezentowane przez współczynniki o małych wartościach, a te z kolei zostaną usunięte przez funkcję progową.

Transformacja falkowa jest stosunkowo nowym i ciągle jeszcze odkrywanym narzędziem do obróbki sygnałów. Dobra aproksymacja oraz niski nakład obliczeniowy to główne atuty, które spowodowały, że już teraz reprezentacja falkowa jest coraz szerzej wykorzystywana w dziedzinach aplikacji takich, jak: przetwarzanie sygnałów i obrazów, analiza numeryczna, astronomia, chemia oraz wiele innych.

Ciągła transformata falkowa sygnału  $x(t)$  jest zdefiniowana w następujący sposób w dziedzinie czasu i częstotliwości:

$$\text{CWT}_x^T(t, a) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(\tau) \gamma^* \left( \frac{\tau - t}{a} \right) d\tau \quad (1.21a)$$

$$\text{CWT}_x^F(t, a) = \sqrt{|a|} \int_{-\infty}^{+\infty} X(f) A^*(af) e^{j \cdot 2\pi f t} df \quad (1.21b)$$

zaś transformacja do niej odwrotna jest określona zależnością

$$x(t) = \frac{2}{C_y} \int_0^{+\infty} \left[ \int_{-\infty}^{+\infty} \text{CWT}(\tau, a) g \left( \frac{t - \tau}{a} \right) d\tau \right] \frac{da}{a^2} \quad (1.22)$$

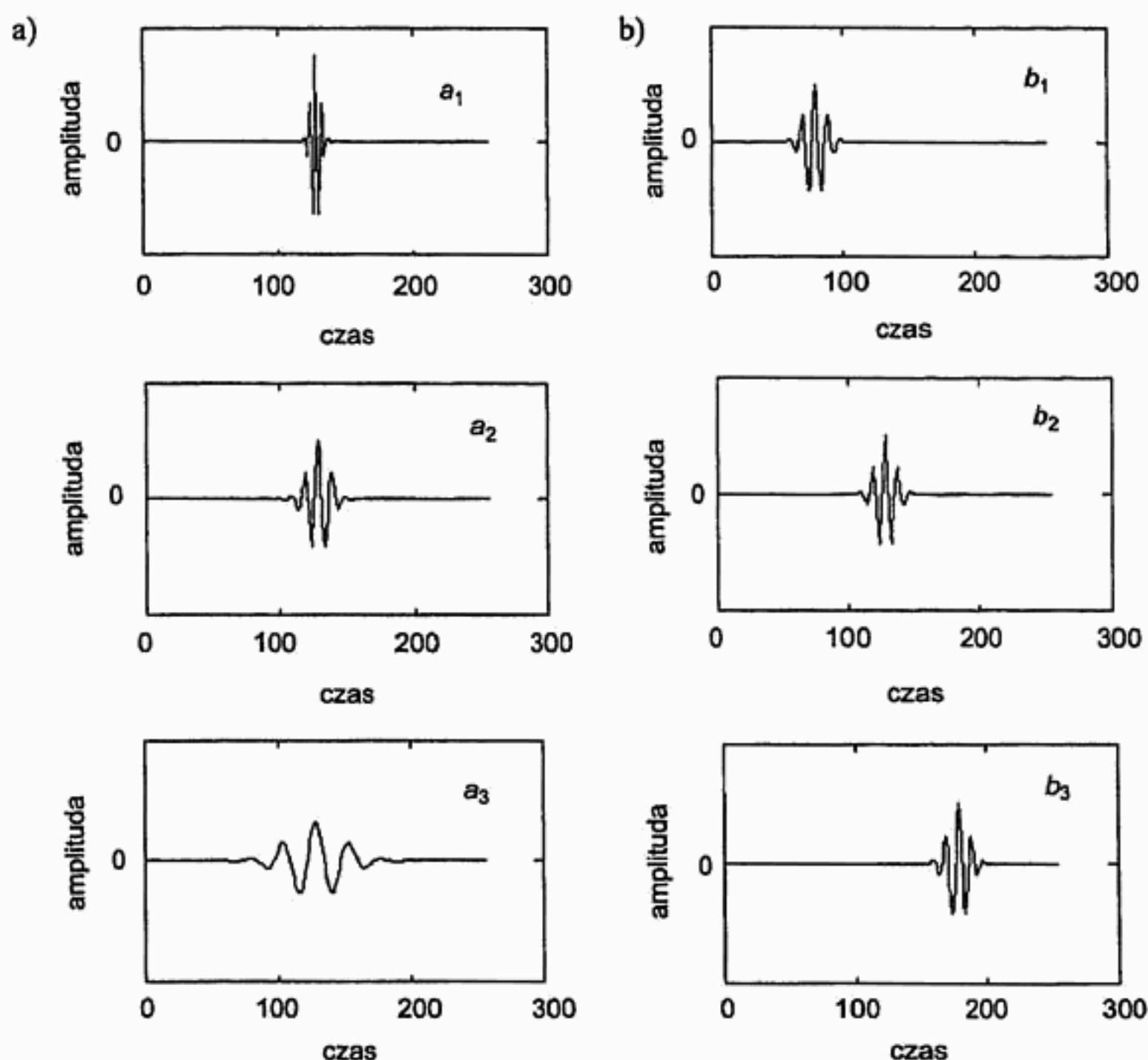
gdzie:

$$C_y = \int_{-\infty}^{+\infty} \frac{|\Gamma(f)|^2}{|f|} df < \infty, \quad \Gamma(0) = 0.$$

Funkcja  $g(t)$  oznacza falkę syntezującą,  $\gamma(t)$  dualną do niej falkę analizującą, zaś  $G(f)$  i  $\Gamma(f)$  są ich widmami Fouriera.

Równanie (1.21a) reprezentuje filtrację sygnału analizowanego  $x(t)$  przez sygnał analizujący  $\gamma(t)$ , przeskalowany w dziedzinie czasu współczynnikiem  $a$  („rozciągany” dla  $a > 1$  oraz „ściskany” dla  $a < 1$ ), natomiast równoważne mu równanie (1.21b) przedstawia odwrotne przekształcenie Fouriera iloczynu widma sygnału i przeskalowanego widma falki  $\gamma(t)$ .





**Rys. 1.7.** Falka Morleta: a) dla kolejnych wartości współczynnika skali  $a$ ;  
b) dla kolejnych wartości współczynnika przesunięcia  $b$

Falka-matka zastosowana po raz pierwszy przez J. Morleta do analizy odpowiedzi podłoża na sygnał sejsmiczny podczas poszukiwania ropy naftowej miała postać (rys. 1.7)

$$\gamma(t) = \exp(j\omega_0 t) \cdot \exp\left(-\frac{t^2}{2}\right) \quad (1.23)$$

gdzie  $\omega_0 = 5,33$  rd/s.

Chociaż w równaniu Morleta bardzo wyraźnie wyodrębnić można czynnik oscylacji i czynnik obwiedni, co pozornie czyni falkę-matkę podobną do funkcji analizujących w transformacji Gabora, należy zwrócić uwagę na wspólny parametr  $t$ , którego zmiana powoduje przeskalowanie zarówno „częstotliwości”, jak i „długości” okna. To właśnie skalowanie, czyli zmiana aspektu czasowego, jest najbardziej charakterystycz-

ną cechą sposobu konstruowania rodziny falek analizujących na podstawie falki-matki. Najbardziej charakterystyczną, ale też i implikującą wiele własności ciągłego przekształcenia falkowego.

Pytaniem najczęściej nasuwającym się podczas prezentacji falki Morleta jest możliwość użycia innej falki-matki lub, bardziej ogólnie, zagadnienie zasad konstrukcji innych falek-matek.

Warunki stawiane przed funkcją kandydującą do roli falki-matki pozwalają zaprojektować ją dość dowolnie:

po pierwsze 
$$\int_{-\infty}^{\infty} \frac{|\hat{\psi}(f)|^2}{|f|} df = K < \infty \quad (1.24)$$

co jest często formułowane jako wymaganie, aby falka-matka była równa zero prawie wszędzie, tj. poza pojedynczym skończonym przedziałem, kiedy nie jest; zwraca uwagę fakt, że wymagany brak składowej stałej implikuje  $\hat{\psi}(0) = 0$ ;

po drugie 
$$\|\psi\|_2 = 1 \quad (1.25)$$

co jest warunkiem normalizacji energii.

Prostym przykładem funkcji mogącej również pełnić rolę falki-matki jest druga pochodna funkcji Gaussa dana wzorem

$$\psi(t) = \frac{2}{\sqrt{3} \cdot \Pi^{\frac{1}{4}}} \cdot (1 - t^2) \cdot \exp\left(\frac{-t^2}{2}\right) \quad (1.26)$$

Interesujące jest także dokonanie Y. Meyera, który studiując szybkość zbieżności widma falki do zera, wprowadził pojęcia rzędu falki i momentu falki, a następnie wykazał istnienie ortogonalnej bazy falek elementarnych. Poszukiwanie takiej bazy było podyktowane faktem, że na skutek znacznej nadmiarowości informacji dotychczas stosowane falki nie zapewniały perfekcyjnej rekonstrukcji sygnału.

Za zbieżność widma falki odpowiada jej pierwszy niezerowy moment

$$M_l = \int_{-\infty}^{\infty} x^l \cdot \hat{\psi}(x) dx \quad (1.27)$$

czyli składnik jej rozwinięcia w szereg Taylora wokół początku układu współrzędnych

$$u_n = \frac{f'(0)}{n^2} \cdot \frac{M_1}{1!} + \frac{f''(0)}{n^3} \cdot \frac{M_{21}}{2!} + \dots + \frac{f^q(0)}{n^{q+1}} \cdot \frac{M_{q1}}{q!} + \dots \quad (1.28)$$

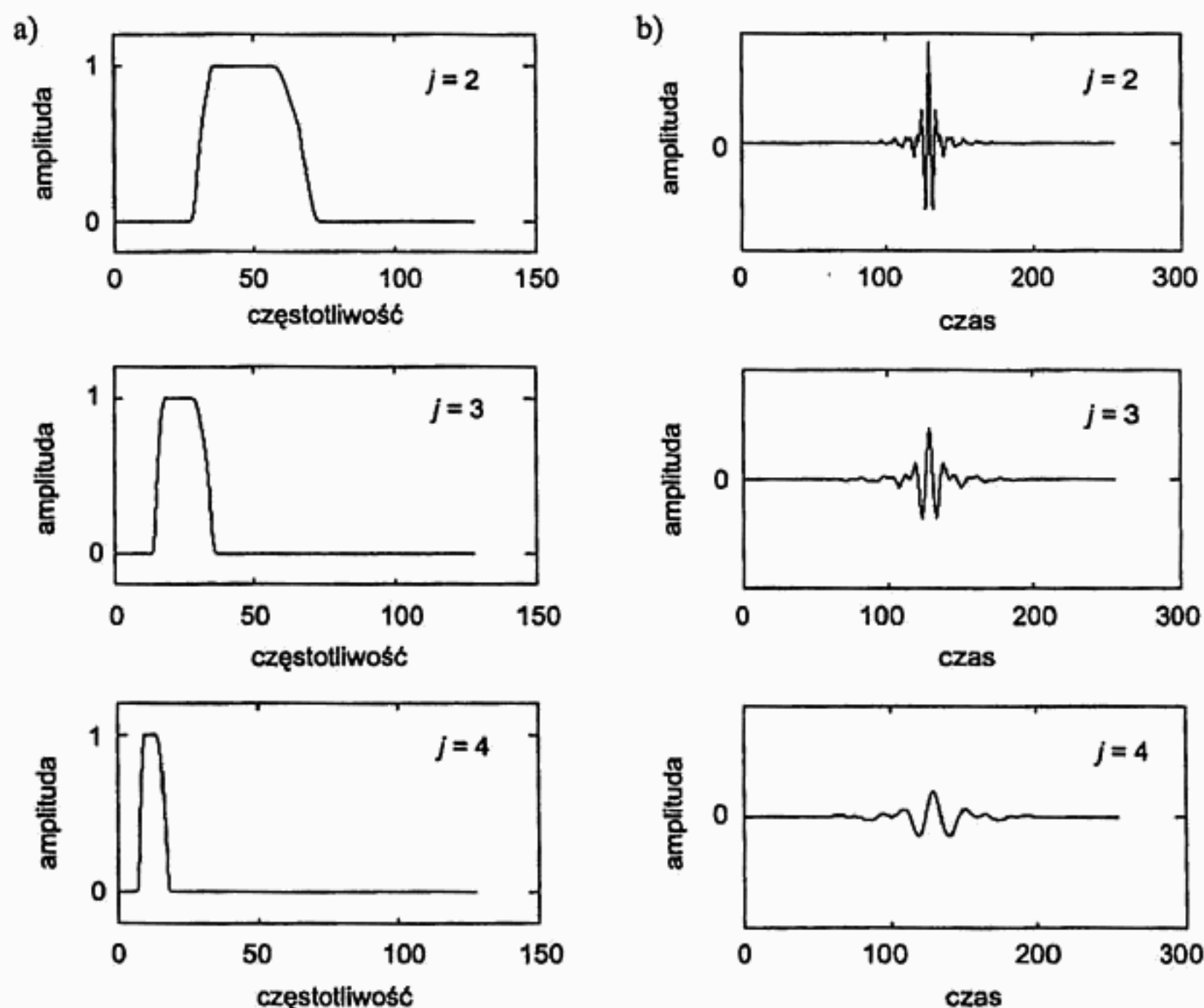
Rzędem falki nazwano liczbę jej pierwszych momentów równych zero. Falka jest rzędu  $r$ , jeśli:

$$\hat{\psi}^{(q)}(0) = 0, \quad q = 1, 2, \dots, r \quad (1.29)$$

Tak więc dla argumentu  $\xi = 0$ ,  $\hat{\psi}(\xi)$  funkcja ma miejsce zerowe rzędu  $r + 1$ .

Dążenie do wartości zerowej widma w otoczeniu punktu  $f = 0$  może zatem być zrealizowane z użyciem falki, której wszystkie momenty będą zerowe. Falka taka należy do klasy określanej jako  $C^\infty$ .

Cechą charakterystyczną falki Meyera jest zwarty nośnik jej transformaty Fouriera (rys. 1.8). Postać analityczna falki oraz stowarzyszonej z nią funkcji skalującej została zdefiniowana w dziedzinie częstotliwości. Żadna z tych funkcji nie posiada natomiast bezpośredniej definicji w dziedzinie czasu, możliwe jest jedynie uzyskanie jej za pomocą odwrotnej transformacji Fouriera.



Rys. 1.8. Falka Meyera dla kolejnych wartości współczynnika skali  $j$ :  
a) w dziedzinie częstotliwości; b) w dziedzinie czasu

Wychodząc z definicji ortonormalnej bazy dekompozycji

$$\psi_{j,k}(t) = 2^{\frac{j}{2}} \cdot \psi(2^j \cdot t - k) \quad (1.30)$$

lub lepiej jej wersji widmowej

$$\hat{\psi}_{j,k}(f) = 2^{-\frac{j}{2}} \cdot \exp(-i \cdot 2\pi f \cdot 2^{-j} \cdot k) \cdot \hat{\psi}(2^{-j} f) \quad (1.31)$$

falkę-matkę Meyera można skonstruować z użyciem dodatkowej funkcji  $\Theta(f)$

$$\hat{\psi}(f) = \exp(-i\pi f) \cdot \Theta(f) \quad (1.32)$$

Funkcja  $\Theta(f)$  musi z kolei spełniać następujące warunki:

$$\begin{aligned} \Theta(f) &= \Theta(-f) \quad \forall f \\ \Theta(f) &\neq 0 \quad \forall f : |f| \in \left[ \frac{1}{2} - \varepsilon, 1 + 2 \cdot \varepsilon \right] \\ \Theta(f) &= 1 \quad \forall f : |f| \in \left[ \frac{1}{2} + \varepsilon, 1 - 2 \cdot \varepsilon \right] \\ \Theta^2(f) + \Theta^2(1-f) &= 1 \quad \forall f : |f| \in \left[ \frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon \right] \\ \Theta^2(2 \cdot f) &= \Theta(1-f) \quad \forall f : |f| \in \left[ \frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon \right] \\ 0 < \varepsilon &\leq \frac{1}{6} \end{aligned} \quad (1.33)$$

Definicja funkcji  $\Theta(f)$  pozostawia pewną dowolność wyboru współczynnika  $\varepsilon$  oraz przebiegu samej funkcji w otoczeniu punktu 0,5. Najczęściej spotykane jest użycie w tej roli funkcji  $\delta(f)$  zdefiniowanej następująco:

$$\begin{aligned} \delta(f) &= \gamma(f) \quad \forall f : 0 \leq f \leq \varepsilon \\ \delta(f) &= \sqrt{1 - \gamma(f)} \quad \forall f : -\varepsilon \leq f < 0 \end{aligned} \quad (1.34)$$

z użyciem pomocniczej funkcji  $\gamma(f)$

$$\gamma(f) = 1 - \exp\left(\frac{\alpha \cdot \varepsilon^2}{(f - \varepsilon)^2}\right) \quad \alpha = \log\left(1 - \frac{1}{\sqrt{2}}\right) \quad (1.35)$$

Znacznie bardziej istotny jest wybór współczynnika  $\varepsilon$ , gdyż wpływa on bezpośrednio na własności konstruowanej falki. Dla wartości  $\varepsilon$  bliskich zero zbocza widma falki są bardzo strome, czego skutkiem są oscylacje czasowego przebiegu falki i powolna jej zbieżność do zera. Przeciwnie, wartości  $\varepsilon$  bliskie  $1/6$  pozwalają na uzyskanie bardziej łagodnych zboczy widma, a zatem bardziej zwartej postaci czasowej falki. Pozwala to uzyskać lepszą koncentrację energii atomu reprezentacji czasowo-częstotliwościowej, ograniczoną w teorii do wartości  $1/2\pi$  przez zasadę nieoznaczoności.

Innym aspektem wyboru współczynnika  $\varepsilon$  jest możliwość dyskretyzacji falki Meyera z zachowaniem wszystkich jej własności. Widmo falki w zakresie od 0 do 1 musi być przedstawialne za pomocą  $2^k$  próbek, w przeciwnym przypadku dokładne odwzorowanie falek dla niskich częstotliwości nie byłoby możliwe przy użyciu całkowitej liczby prążków. Wynikają stąd ograniczenia na współczynnik  $\varepsilon$ , mogący przyjąć w przypadku najczęściej praktycznie spotykanych sygnałów zaledwie kilka wartości z przedziału  $(0 \div 1/6)$ .

Falka Meyera była pierwszym praktycznie stosowanym systemem ortogonalnej dekompozycji sygnałów do dziedziny czasowo-częstotliwościowej. Posiada ona własność dokładnej rekonstrukcji sygnału, co otwierało nowe zakresy zastosowań, jak np. przetwarzanie sygnałów w dziedzinie czasowo-częstotliwościowej. Wadami tej transformacji są: wolno wygasający nośnik w dziedzinie czasu, co utrudnia przetwarzanie krótkich sygnałów, rozszerzając efekt brzegowy, oraz niedokładne reprezentowanie falki w widmie o niskich częstotliwościach, przez co analiza w tych zakresach traci lokalizację czasową.

Ciągła czasowo-częstotliwościowa reprezentacja falkowa zawiera wiele nadmiarowych informacji. W wyniku spróbkowania jej parametrów, tzn. czasu  $t$  i współczynnika skali  $a$ , otrzymuje się współczynniki szeregu falkowego, analogiczne do współczynników szeregu Fouriera. Po podstawieniu:

$$t = n2^{-m}, \quad a = 2^{-m} \quad (1.36)$$

uzyskuje się diadyczny szereg falkowy. Znalazł on obecnie najszersze praktyczne zastosowanie i nosi nazwę dyskretniej transformacji falkowej (DWT). Sygnał przedstawiony za pomocą DWT ma postać podwójnej sumy

$$x(t) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} d_{m,n} g_{m,n}(t) \quad (1.37)$$



w której współczynniki  $d_{m,n}$  oblicza się ze wzoru

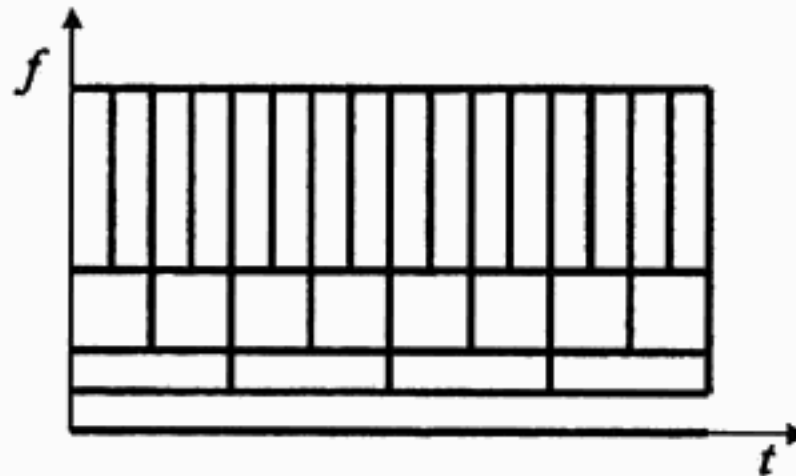
$$d_{m,n} = \int x(t) \gamma_{m,n}^*(t) dt \quad (1.38)$$

a funkcje bazowe  $g_{m,n}$  ze skalowania i translacji falki-matki:

$$g_{m,n}(t) = 2^{-m/2} \gamma(2^{-m}t - n) \quad (1.39)$$

Funkcja  $\gamma_{m,n}(t)$  jest funkcją dualną do  $g_{m,n}(t)$ . Przy założeniu ortogonalności bazy falkowej  $\gamma_{m,n}(t) = g_{m,n}(t)$ .

Cechą charakterystyczną dla dyskretnej transformacji falkowej jest stałość podstawowej komórki czasowo-częstotliwościowej tzw. atomu TF. Kiedy parametr skali  $m$  zwiększa się o 1, falki stają się dwa razy krótsze, ich widmo częstotliwościowe jest dwa razy szersze, czyli pole atomu TF pozostaje niezmiennie. W wyniku tego DWT realizuje schemat dekompozycji czasowo-częstotliwościowej przedstawiony na rysunku 1.9.



Rys. 1.9. Rozmieszczenie atomów w powierzchni czasowo-częstotliwościowej dekompozycji falkowej

### 1.3.2. Analiza wielorozdzielcza

Transformacja falkowa jest ściśle związana z teorią analizy wielorozdzielczej. W analizie tej rozpatruje się ciąg przestrzeni funkcyjnych, w których dane sygnały cyfrowe aproksymowane są z różną dokładnością i wyrażane za pomocą różnej liczby funkcji bazowych. Owe przestrzenie funkcyjne, oznaczane przez  $V_m$ ,  $m \in \mathbb{Z}$ , powinny spełniać następujące założenia:

- 1)  $\dots \subset V_2 \subset V_1 \subset V_0 \dots$
- 2)  $\bigcap_{m \in \mathbb{Z}} V_m = \{0\}$ ,  $\overline{\bigcup_{m \in \mathbb{Z}} V_m} = L^2(R)$
- 3)  $f(x) \in V_m$  wtedy i tylko wtedy, gdy  $f(2^m x) \in V_0$
- 4) jeżeli  $f(x) \in V_0$  to  $f(x-n) \in V_0$  dla wszystkich  $n \in \mathbb{Z}$

(1.40)

Pierwszy z aksjomatów wyraża wzajemne zawieranie się kolejnych przestrzeni. Oznacza to, że przestrzeń  $V_{m-1}$  zawiera wszystkie funkcje z przestrzeni  $V_m$  oraz pewne dodatkowe funkcje, których w  $V_m$  nie ma. Aksjomat drugi świadczy o tym, iż nie istnieje funkcja, która należałaby do wszystkich przestrzeni i że domknięcie wszystkich przestrzeni daje przestrzeń funkcji o „ograniczonej” energii. Warunek trzeci wyraża fakt, że wszystkie przestrzenie są skalowaną wersją jednej przestrzeni (aspekt wielorozdzielczy). Czwarte założenie stwierdza natomiast, iż dowolna funkcja po operacji translacji pozostaje w tej samej przestrzeni funkcyjnej, do której należała pierwotnie.

Dla przestrzeni o podanych wyżej własnościach można zdefiniować funkcję  $\varphi(x) \in V_0$  taką, że  $\varphi_{0,n}(x) = \varphi(x - n)$  stanowi bazę ortonormalną przestrzeni  $V_0$ . Funkcja  $\varphi(x)$  nazywana jest funkcją skalującą i odgrywa kluczową rolę w transformacji falkowej.

Przykładowo sygnał  $x(t)$  może być przedstawiony w przestrzeni  $V_m$  na  $m$ -tym poziomie aproksymacji przy pomocy sumy poprzesuwanych funkcji bazowych  $\varphi_{m,n}(t)$ .

$$x(t) \approx \sum_n c_{m,n} \varphi(2^{-m}t - n), \quad c_{m,n} = \int x(t) \varphi^*(2^{-m}t - n) dt \quad (1.41)$$

Ten sam sygnał można również przedstawić z dwukrotnie mniejszą rozdzielczością w przestrzeni  $V_{m+1}$  (gorsza aproksymacja – mniej szczegółowa)

$$x(t) \approx \sum_n c_{m+1,n} \varphi(2^{-(m+1)}t - n) = \sum_n c_{m+1,n} \varphi\left(\frac{2^{-m}t}{2} - n\right) \quad (1.42)$$

$$c_{m+1,n} = \int x(t) \varphi^*\left(\frac{2^{-m}t}{2} - n\right) dt$$

Przy przejściu z przestrzeni  $V_m$  (bardziej szczegółowej) do przestrzeni  $V_{m+1}$  (mniej szczegółowej) część informacji zostaje utracona. Odrzucone szczegóły można umownie zdefiniować w przestrzeni  $W_{m+1}$ , która stanowi dopełnienie przestrzeni  $V_{m+1}$  do  $V_m$ , czyli

$$V_m = V_{m+1} + W_{m+1} \quad (1.43)$$

Jeżeli przestrzeń  $V_{m+1}$  jest rozpinana przez funkcje  $\varphi_{m+1,n}(t)$ , to w przypadku przestrzeni  $W_{m+1}$  rolę tę pełnią funkcje  $\psi_{m+1,n}(x)$ , nazywane falkami, i są one ortogonalne do  $\varphi_{m+1,n}(t)$ , a wspólnie natomiast rozpinają przestrzeń  $V_m$ . Wynika z tego fakt, że bez straty informacji sygnał z poziomu  $m$  (przestrzeni  $V_m$ ) można opisać za pomocą funkcji bazowych z przestrzeni niższego poziomu  $V_{m+1}$  i  $W_{m+1}$

$$x(t) \approx \sum_n c_{m,n} \varphi(2^{-m}t - n) = \sum_n c_{m+1,n} \varphi(2^{-(m+1)}t - n) + \sum_n d_{m+1,n} \psi(2^{-(m+1)}t - n) \quad (1.44)$$

Podobnie jak powyżej, przestrzeń  $V_{m+1}$  można zastąpić za pomocą  $V_{m+2}$  i  $W_{m+2}$ . Analogicznie można definiować kolejne przestrzenie aproksymujące  $V$  oraz ich dopełnienia (do wyższych w hierarchii poziomów)  $W$ . W efekcie otrzymujemy drzewiastą strukturę przestrzeni z kolejnych poziomów:

$$V_m = V_{m+k} + W_{m+k} + W_{m+k-1} + \dots + W_{m+2} + W_{m+1}$$

$$V_m = V_{m+k} + \sum_{\rho=m+1}^{m+k} W_\rho \quad (1.45)$$

Zależność (1.45) jest fundamentalnym równaniem wykorzystywanym w dekompozycji falkowej. Pozwala zapisać sygnał na różnych poziomach szczegółowości za pomocą zespołu wzajemnie ortogonalnych funkcji bazowych: falek  $\psi(t)$  oraz funkcji skalujących  $\varphi(t)$

$$x(t) = \sum_n c_{m_k,n} \varphi(2^{-m_k} t - n) + \sum_{m=m_0}^{m_k} \sum_n d_{m,n} \psi(2^{-m} t - n) \quad (1.46)$$

gdzie współczynniki

$$c_{m_k,n} = \int x(t) \varphi^*(2^{-m_k} t - n) dt \quad (1.47a)$$

reprezentują aproksymację sygnału  $x(t)$  na  $k$ -tym poziomie, natomiast współczynniki

$$d_{m,n} = \int x(t) \psi^*(2^{-m} t - n) dt \quad (1.47b)$$

stanowią detale (szczegóły) z kolejnych poziomów przestrzeni funkcyjnych  $W$ .

Na funkcję skalującą oraz falekę narzuca się następujące warunki:

$$\int \varphi(t) dt = 1,$$

$$\int \psi(t) dt = 0 \quad (1.48)$$

Są one niezwykle istotne w przypadku falki. Oznaczają zwarty nośnik funkcji  $\psi(t)$ , czyli wartość zerową w całym zakresie czasowym, z wyjątkiem kilku skupionych punktów różnych od zera.

Otrzymanie współczynników dekompozycji falkowej  $c_{m,n}$  i  $d_{m,n}$  na podstawie definicji (wzory (1.47a) i (1.47b)) jest dosyć kłopotliwe, dlatego w praktycznej realizacji numerycznej dyskretnej transformacji falkowej nie wykorzystuje się falek (i funkcji skalujących), tylko związane z nimi filtry. Zagadnienie to wyjaśnić można dwoma drogami: za pomocą metod cyfrowego przetwarzania sygnałów oraz teorii wielorozdzielczej aproksymacji sygnału.

Oparcie się na teorii wielorozdzielczej aproksymacji sygnału pozwoli wyprowadzić rekurencyjną zależność między współczynnikami dekompozycji falkowej  $c_{m,n}$  i  $d_{m,n}$  z dwóch kolejnych poziomów rozdzielczości. Punktem wyjścia jest zależność między przestrzeniami funkcyjnymi

$$V_m = V_{m+1} + W_{m+1} \quad (1.49)$$

Jeżeli mamy pewną przestrzeń funkcyjną aproksymacji  $V_0$  oraz przestrzeń szczegółów  $W_0$ , to możemy zdefiniować przestrzeń „z wyższego poziomu”  $V_{-1}$ , która będzie „nadprzestrzenią”  $V_0$  i  $W_0$ . Skoro więc:

$$\begin{aligned} V_0 &\subset V_{-1}, \\ W_0 &\subset V_{-1} \end{aligned} \quad (1.50)$$

to każdą funkcję należącą do przestrzeni  $V_0$  i  $W_0$  możemy opisać za pomocą funkcji bazowych przestrzeni  $V_{-1}$ . Ponieważ funkcja skalująca  $\varphi_0(t)$  należy do  $V_0$ , a falka  $\psi_0(t)$  należy do  $W_0$  oraz  $\{\varphi_{-1}, n(t), n \in \mathbb{Z}\}$  jest bazą ortonormalną w  $V_{-1}$ , to:

$$\varphi(2^0 t) = \sum_n h(n) \varphi(2^1 t - n) \quad (1.51a)$$

$$\psi(2^0 t) = \sum_n g(n) \varphi(2^1 t - n) \quad (1.51b)$$

Z równania (1.51a) wynika, że funkcja skalująca na poziomie niższym jest liniową kombinacją kilku funkcji skalujących z poziomu wyższego. Ponieważ na poziomie niższym aproksymacja jest mniej szczegółowa (bardziej niskoczęstotliwościowa), to współczynniki wagowe  $h(n)$  pełnią rolę filtru dolnoprzepustowego. Analogicznie falka na poziomie niższym jest liniową kombinacją kilku funkcji skalujących z poziomu wyższego (1.51b), tym razem jednak reprezentuje szczegóły. Zatem współczynniki  $g(n)$  reprezentują filtr górnoprzepustowy.

Znając zależności między funkcjami bazowymi z dwóch kolejnych poziomów przestrzeni funkcyjnych (1.51a), (1.51b) oraz wykorzystując równania definicyjne współczynników  $c_{m,n}$  (1.47a) i  $d_{m,n}$  (1.47b), można wyprowadzić zależność pomiędzy współczynnikami rozwinięcia na tych poziomach:

$$c_m(n) = \sum_k h(n - 2k) c_{m-1}(k) \quad (1.52a)$$

$$d_m(n) = \sum_k g(n - 2k) c_{m-1}(k) \quad (1.52b)$$

Są to kluczowe wzory definiujące związek pomiędzy współczynnikami z dwóch kolejnych poziomów rozwinięcia falkowego. Właśnie te zależności mają obecnie zastosowanie w praktycznych algorytmach transformacji falkowych sygnałów. Mają one swoją praktyczną interpretację: Współczynniki  $c_{m,n}$  poziomu niższego otrzymuje się drogą filtracji dolnoprzepustowej współczynników  $c_{m-1,n}$  z poziomu wyższego za pomocą filtru  $h(n)$  oraz następującej po niej decymacji rzędu drugiego tzw. „down-samplingu” (usunięcie co drugiej wartości), zaś współczynniki  $d_{m,n}$  w wyniku filtracji górnoprzepustowej tych samych współczynników  $c_{m-1,n}$  – za pomocą filtru  $g(n)$  i analogicznej decymacji. Sam proces filtracji, lub inaczej: splotu dwóch sygnałów, zdefiniowany jest następująco

$$c'_m(n) = \sum_k h(n-k) c_{m-1}(k) \quad (1.53)$$

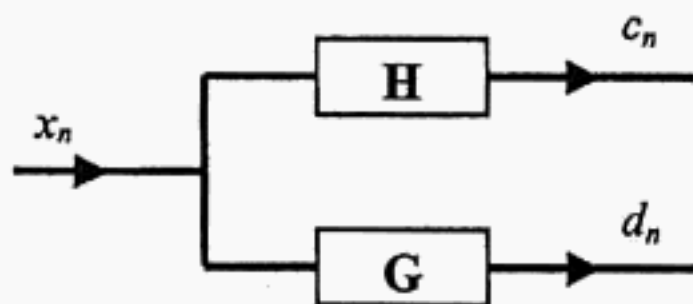
Decymacja sygnału jest natomiast zrealizowana przez

$$c_m(n) = c'_m(2n) \quad (1.54)$$

Wzór (1.52) jest zwięźlejszym zapisem dwóch powyższych operacji.

### 1.3.3. Filtracja pasmowa sygnału

Wykorzystując metody cyfrowego przetwarzania sygnałów do wyznaczenia współczynników  $c_{m,n}$  i  $d_{m,n}$  uwzględnia się fakt, iż funkcje skalujące  $\varphi(t)$  posiadają widmo częstotliwościowe odpowiadające filtrowi dolnoprzepustowemu, natomiast widmo falki  $\psi(t)$  przypomina widmo filtru pasmowoprzepustowego (górnoprzepustowego). W tym celu należy zaprojektować odpowiednie filtry, których pasma przepuszczania będą odpowiadały widmom funkcji bazowych ( $\varphi(t)$  i  $\psi(t)$ ). Przefiltrowanie sygnału dyskretnego  $x_n$  przez zespół „filtrów bazowych” da w efekcie ciąg współczynników  $c_{m,n}$  i  $d_{m,n}$ . Pojedynczy cykl filtracji (na jednym poziomie dekompozycji  $m$ ) jest zaprezentowany na rysunku 1.10.



Rys. 1.10. Filtracja sygnału  $x(t)$  filtrami dolnoprzepustowym  $H$  i górnoprzepustowym  $G$

Filtr  $H$  jest filtrem dolnoprzepustowym o transmitancji  $H(\omega)$  odpowiadającej funkcji  $\varphi(t)$ , natomiast  $G$  jest filtrem górnoprzepustowym o transmitancji  $G(\omega)$  odpowiadającym funkcji  $\psi(t)$ . Pojedynczy etap filtracji przedstawiony na rysunku dokonuje



rozszczerzenia sygnału na dwa sygnały: dolnopasmowy i górnopasmowy, czyli wyodrębnienia „szkieletu” i szczegóły sygnału. Z numerycznego punktu widzenia operacja filtracji sygnału nie nastręcza żadnych trudności. W dziedzinie częstotliwości transformaty sygnału na wyjściu jest równa iloczynowi transmitancji filtru i transformaty sygnału wejściowego:

$$\begin{aligned} C(\omega) &= H(\omega) \cdot X(\omega) \\ D(\omega) &= G(\omega) \cdot X(\omega) \end{aligned} \quad (1.55)$$

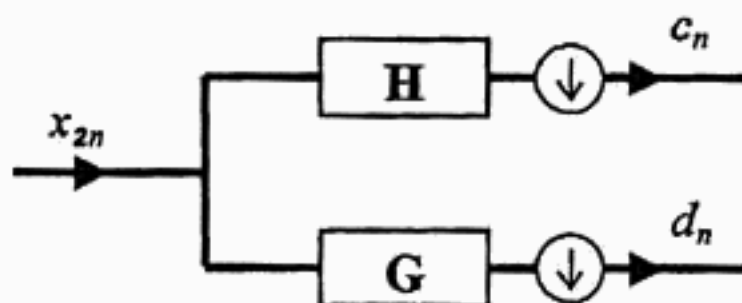
W dziedzinie czasu natomiast operacja ta jest splotem (konwolucją) sygnału wejściowego i odpowiedzi impulsowej filtru  $h(n)$ ,  $g(n)$  (sygnałem otrzymanym przy wymuszeniu filtru „delta Diraca”  $\delta(t)$ ):

$$c(n) = \sum_k h(n-k)x(k) \quad (1.56b)$$

$$d(n) = \sum_k g(n-k)x(k) \quad (1.56a)$$

Powyższe zależności są analogiczne do wzorów (1.52a) i (1.52b) wyprowadzonych w oparciu o teorię wielorozdzielczej aproksymacji sygnału. Brak w nich jedynie mechanizmu decymacji, czyli usuwania co drugiej próbki.

Po uwzględnieniu decymatorów i założeniu, że sygnał wejściowy  $x(t)$  jest równy współczynnikom aproksymacji z wyższego poziomu  $c(n)$ , postać zależności (1.56a) i (1.56b) jest już identyczna ze wzorami (1.52a), (1.52b) wyprowadzonymi na podstawie teorii wielorozdzielczej analizy sygnałów. Natomiast układ filtracji z rysunku 1.10 zmieni postać na przedstawioną na rysunku 1.11 i będzie realizował proces pojedynczego poziomu dekompozycji polegający na uzyskaniu współczynników niższego poziomu  $m+1$  na podstawie współczynników z wyższego poziomu.



Rys. 1.11. Pojedynczy etap dekompozycji falkowej

Należy zauważyć, że dzięki operacji „downsamplingu” powyższy algorytm nie wprowadza zmian długości obrabianego sygnału. Sam proces filtracji spowodowałby, że długość sygnału  $c_{m+1}$  oraz długość  $d_{m+1}$  byłaby równa długości sygnału  $x_m$ , co dwukrotnie zwiększyłoby ilość danych potrzebnych do magazynowania tej samej informa-

cji. Usunięcie co drugiej próbki nie wprowadza żadnych zniekształceń przenoszonej informacji (umożliwia pełną rekonstrukcję sygnału) oraz pozwala zachować pierwotną długość sygnału.

Przedstawiony powyżej układ filtracji i „downsamplingu” stanowi pojedynczy poziom dekompozycji falkowej (oznaczony jako „H-G”), która w swojej idei ma za zadanie rozłożyć sygnał na wiele sygnałów niosących informacje o sygnale pierwotnym na różnych poziomach szczegółowości. Ogólny schemat obliczania transformaty falkowej przedstawia się według poniższego algorytmu

$$c_0 = [c_{0,0} \quad c_{0,1} \quad \dots \quad c_{0,N-1}]_N \quad (1.57)$$

Sygnał poddawany transformacji falkowej zawierający  $N$  próbek wygodnie jest zapisać za pomocą wektora (indeks  $N$  za nawiasem oznacza długość wektora)

$$[c_{1,0} \quad c_{1,1} \quad \dots \quad c_{1,N/2-1}]_{N/2} [d_{1,0} \quad d_{1,1} \quad \dots \quad d_{1,N/2-1}]_{N/2} \quad (1.58)$$

Po pierwszym kroku dekompozycji transformata falkowa tego sygnału jest wektorem o tej samej długości, tzn. zawiera  $N/2$  współczynników  $c_{1,k}$  i  $N/2$  współczynników  $d_{1,k}$ .

W następnej iteracji transformata falkowa będzie miała elementy:

$$[c_{2,n}]_{N/4}, [d_{2,n}]_{N/4}, [d_{1,n}]_{N/2} \quad (1.59)$$

Po  $l$ -krotnym zastosowaniu tej procedury otrzymuje się transformatę falkową wektora  $[c_{0,n}]_N$  w postaci wektora:

$$[c_{l,n}]_{N/(2^l)}, [d_{l,n}]_{N/(2^l)}, \dots, [d_{2,n}]_{N/4}, [d_{1,n}]_{N/2} \quad (1.60)$$

Jeżeli  $N$  jest potęgą 2, to  $l$  może osiągnąć wartość  $\log_2 N$ . Na dole piramidy (rys. 1.12) będzie pojedynczy współczynnik reprezentujący aproksymację (od funkcji skalującej), powyżej zaś – same współczynniki reprezentujące szczegóły (od falek).

Kolejne wartości współczynników wyznacza się z zależności:

$$c_{m,n} = \sum_k h_k c_{m-1,2n-k}, \quad d_{m,n} = \sum_k g_k c_{m-1,2n-k} \quad (1.61)$$

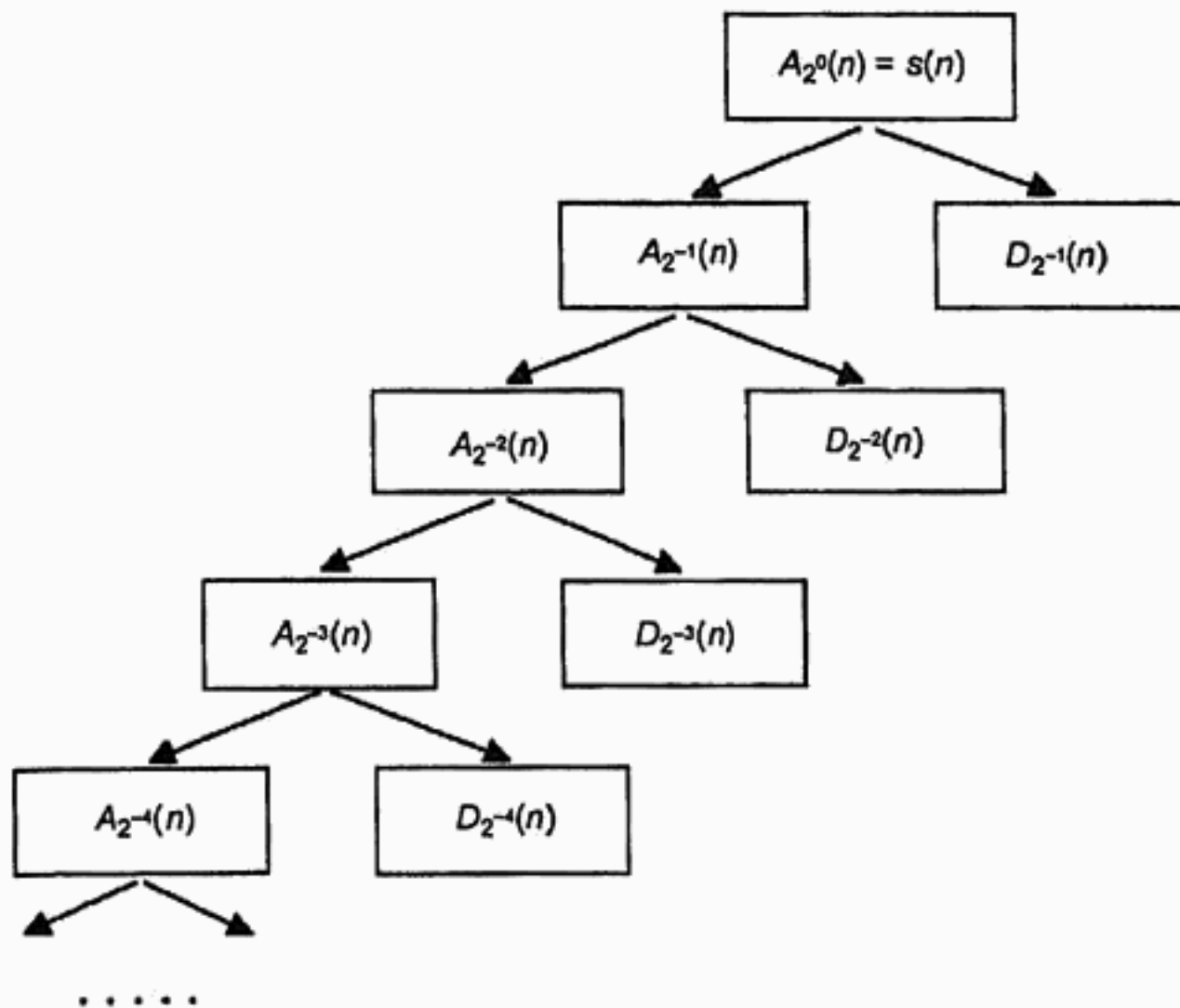
gdzie:

$h_k = [h_{k,0}, h_{k,1}, \dots, h_{k,L-1}]$  – odpowiedź impulsowa filtru dolnoprzepustowego,

$g_k = [g_{k,0}, g_{k,1}, \dots, g_{k,L-1}]$  – odpowiedź impulsowa filtru górnoprzepustowego,

$n = 0, 1, \dots, N/2-1,$

$\max(0, 2n - L + 1) \leq k \leq \min(2n, L - 1).$



Rys. 1.12. Schemat wielopoziomowej dekompozycji piramidowej

Użyte wyżej filtry muszą charakteryzować się określonymi właściwościami:

- 1) muszą mieć skończoną odpowiedź impulsową FIR (ang. *finite impulse response*),
- 2) powinny zapewniać dokładną rekonstrukcję sygnału, tj. spełniać równanie

$$c_{m-1,n} = \sum_k [h' 2_{k-n} c_{m,k} + g' 2_{k-n} d_{m,k}] \quad (1.62)$$

gdzie:

$$h_k = [h_{k,L-1}, h_{k,L-2}, \dots, h_{k,1}, h_{k,0}],$$

$$g_k = [g_{k,L-1}, g_{k,L-2}, \dots, g_{k,1}, g_{k,0}],$$

$$n = 0, 1, \dots, N-1,$$

$$n/2 \leq k \leq \min(N/2-1, n+L-1);$$

- 3) powinny być filtrami liniowo-fazowymi (ang. *linear-phase*), tj. takimi, których transmitancja wyraża się wzorem:

$$H(e^{j\omega}) = A(\omega) \exp[-j\Phi(\omega)], \quad \Phi(\omega) = [k\pi + (L-1)\omega]/2$$

$$k = 0 \text{ dla filtru symetrycznego } (h_{L-n-1} = h_n, n = 0, 1, \dots, L-1), \quad (1.63)$$

$$k = 1 \text{ dla filtru antysymetrycznego } (h_{L-n-1} = -h_n, n = 0, 1, \dots, L-1).$$

W rzeczywistości wszystkie trzy warunki spełnione są tylko dla tzw. falki Haara, która jest przypadkiem trywialnym i nie ma w praktyce większego zastosowania. Pro-

projektowane obecnie filtry spełniają dwa pierwsze warunki i są to tzw. filtry kwadraturowe FIR-QMF (ang. *finite impulse response-quadrature mirror filter*). Jest to zespół czterech filtrów (dwóch stosowanych do analizy –  $h$ ,  $g$ , i dwóch do syntezy –  $h'$ ,  $g'$ ) ściśle powiązanych ze sobą i odpowiadających jednej falce.

Z praktycznego punktu widzenia projektowanie falek zdeterminowane jest stworzeniem zespołu filtrów kwadraturowych o żądanych właściwościach, z których jednoznacznie wynika kształt falki. Same filtry natomiast tworzy się według poniższego algorytmu.

1. Ustala się odpowiedź impulsową filtru dolnoprzepustowego analizy  $h$  (pochodzącego od funkcji skalującej) spełniającą następujące ograniczenia:

$$\sum_{n=0}^{L-1} (h_n)^2 = 1 \quad (1.64a)$$

$$\sum_{n=0}^{L-1} h_{2k+n} h_n = 0 \quad (1.64b)$$

Pierwsze ograniczenie wynika z „unormowania” odpowiedzi filtru (równości norm funkcji skalujących dla różnych rozdzielczości). Często przekształca się je do prostszej postaci

$$\sum_{n=0}^{L-1} h_n = \sqrt{2} \quad (1.65)$$

Drugi warunek wynika natomiast z ortogonalności translacji funkcji skalującej. Wymnożenie filtru „przez siebie” z przesunięciem o parzystą liczbę próbek daje wartość zerową.

Powyższe dwa ograniczenia z reguły nie określają jednoznacznie postaci filtru, lecz pozostawiają jeszcze kilka stopni swobody, które wykorzystuje się w celu uzyskania żądanych, specyficznych właściwości falki.

2. Pozostałe zaś filtry: górnoprzepustowy analizy  $g$  i obydwa filtry syntezy  $h'$  i  $g'$ , zdefiniowane są za pomocą filtru  $h$ :

$$\begin{aligned} g_n &= (-1)^n h_{L-1-n} \\ h'_n &= h_{L-1-n} \\ g'_n &= g_{L-1-n} = (-1)^{n+1} h_n \end{aligned} \quad (1.66)$$

Jeżeli więc zaprojektowany filtr  $h$  ma odpowiedź impulsową

$$(h_n) = (h_0, h_1, h_2, h_3) \quad (1.67)$$

to postać pozostałych filtrów będzie następująca:

$$\begin{aligned}(g_n) &= (h_3, -h_2, h_1, -h_0) \\ (h'_n) &= (h_3, h_2, h_1, h_0) \\ (g'_n) &= (-h_0, h_1, -h_2, h_3)\end{aligned}\tag{1.68}$$

### 1.3.4. Rekonstrukcja

Rekonstrukcja sygnału jest procesem odwrotnym do dekompozycji. Stanowi wyodrębnienie sygnału pierwotnego z szeregu współczynników falkowych otrzymanych w wyniku dekompozycji. Do rekonstrukcji wykorzystuje się dwa filtry syntezy z zestawu filtrów kwadraturowych: dolnoprzepustowy  $h'$  i górnoprzepustowy  $g'$ .

Pojedynczy cykl procesu rekonstrukcji odtwarza sygnał zawierający się w przestrzeni  $V_m$  z sygnału wygładzonego z przestrzeni  $V_{m+1}$  i z sygnału szczegółowego należącego do przestrzeni  $W_{m+1}$ . Odbywa się to zgodnie ze wzorem

$$\begin{aligned}c_{m,n} &= \sum_k k [h'_{2k-n} c_{m+1,k} + g'_{2k-n} d_{m+1,k}] \\ &= \sum_k k h'_{2k-n} c_{m+1,k} + \sum_k k g'_{2k-n} d_{m+1,k}\end{aligned}\tag{1.69}$$

Realizacja numeryczna wzoru (1.69) jest dosyć kłopotliwa, dlatego też w praktyce stosuje się przekształcenie według poniższego algorytmu:

$$\sum_k k h'_{2k-n} c_{m+1,k} = \sum_k k h'_{k-n} c'_{m+1,k}\tag{1.70a}$$

gdzie:

$$c'_m = c_{2m},$$

$$\sum_k k g'_{2k-n} d_{m+1,k} = \sum_k k g'_{k-n} d'_{m+1,k}\tag{1.70b}$$

gdzie:

$$d'_m = d_{2m},$$

wobec czego wzór (1.69) zmieni się do postaci

$$c_{m,n} = \sum_k k h'_{k-n} c'_{m+1,k} + \sum_k k g'_{k-n} d'_{m+1,k}\tag{1.71}$$

Zamiast sumowania po parzystych indeksach odpowiedzi impulsowej filtru i oryginalnego sygnału (jak sugeruje wzór (1.69)), powyższe sumowanie rozciąga się na wszystkie elementy filtru i sygnału przekształconego poprzez dopróbkowanie za po-

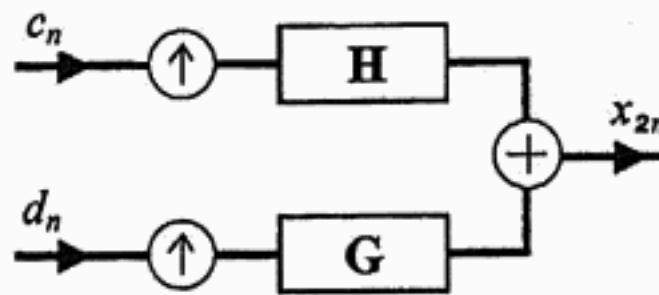


mocą tzw. ekspanderów. Powodują one dwukrotne wydłużenie sygnału, wstawiając pomiędzy oryginalnymi próbkami dodatkowe próbki zerowe:

$$\begin{aligned} c'_{m+1,k} &= (c_{m+1,0}, 0, c_{m+1,1}, 0, c_{m+1,2}, 0, c_{m+1,3}, 0, c_{m+1,4}, \dots) \\ d'_{m+1,k} &= (d_{m+1,0}, 0, d_{m+1,1}, 0, d_{m+1,2}, 0, d_{m+1,3}, 0, d_{m+1,4}, \dots) \end{aligned} \quad (1.72)$$

Wyrażenie (1.71) jest równoważne definicji rekonstrukcji falkowej (1.69). Określa ono praktyczny algorytm obliczania próbek sygnału z współczynników falkowych.

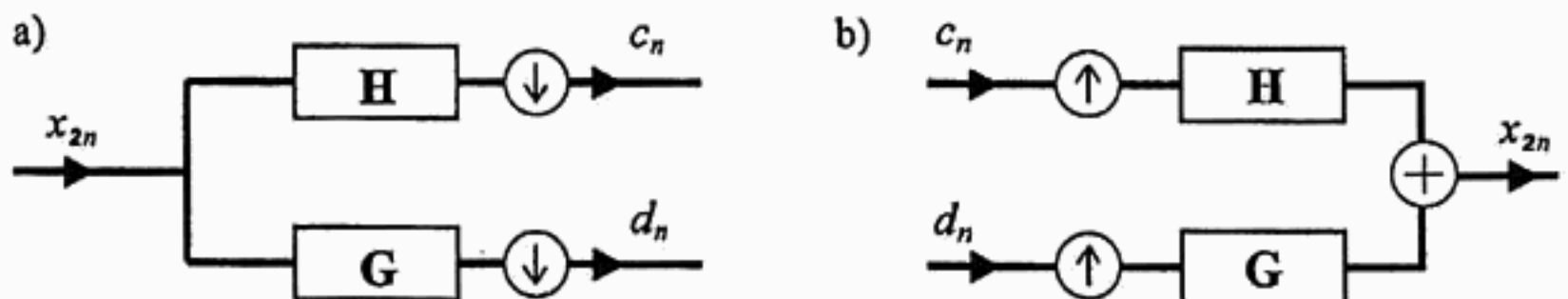
Schemat pojedynczego poziomu rekonstrukcji przedstawia rysunek 1.13.



Rys. 1.13. Pojedynczy etap rekonstrukcji sygnału

Obydwa sygnały: wygładzony  $c_n$  i reprezentujący szczegóły  $d_n$ , są podawane na ekspandery (oznaczone kółkiem ze strzałką w górę), które dokonują procesu „upsamplingu”, czyli dwukrotnego wydłużenia sygnału poprzez dołożenie między wszystkimi próbkami sygnału pierwotnego próbek zerowych. Powstaną wtedy ciągi próbek mające postać (1.72). Następnie otrzymane sygnały są filtrowane odpowiednio filtrami syntezy dolnoprzepustowym  $h'$  i górnoprzepustowym  $g'$ , po czym z kolei następuje operacja sumowania. Na wyjściu sumatora pojawia się zrekonstruowany sygnał  $x_{2n}$  należący do przestrzeni  $V_m$ .

Z dotychczasowych rozważań wynika, że transformacja falkowa jest transformacją bezstratną, tzn. można sygnał poddać procesowi dekompozycji, a następnie współczynniki falkowe (będące próbkami sygnałów z przestrzeni o niższych rozdzielczościach) poddać procesowi rekonstrukcji, by w efekcie otrzymać sygnał identyczny z oryginalnym. Pojedynczy poziom dekompozycji i rekonstrukcji falkowej jest przedstawiony na rysunku 1.14.



Rys. 1.14. Pojedynczy etap dekompozycji (a) i rekonstrukcji (b) pokazujący prostą i odwrotną transformację falkową bez straty informacji

### 1.3.5. Zasada zachowania energii

Transformacja falkowa jest przekształceniem bezstratnym. Podczas przejścia sygnału z dziedziny czasu do postaci współczynników falkowych cała energia zostaje zachowana. Właściwość tę precyzyjnie określa wzór

$$\sum_{n=0}^{N-1} |s(n)|^2 = \frac{1}{N} \sum_{v=-N/2+1}^{N/2} |\tilde{s}(v)|^2 = N^2 |S_{0,0}|^2 + \sum_{j=-p}^{-1} \sum_{k=0}^{2^j N-1} |S_{j,k}|^2 \quad (1.73)$$

gdzie:

$s(n)$  – wartość  $n$ -tej próbki sygnału,

$N$  – długość sygnału,

$S_{j,k}$  –  $k$ -ty współczynnik detali z  $j$ -tego poziomu dekompozycji,

$S_{0,0}$  – współczynnik aproksymacji stojący na piramidzie współczynników falkowych po  $p$ -krotnej dekompozycji falkowej (tuż obok odpowiadającego mu współczynnika detali  $S_{-p,0}$ ),

$j$  – indeks skalowania,

$k$  – indeks translacji.

Zależność ta definiuje zasadę zachowania energii sygnału podczas transformacji falkowej prostej (dekompozycji) i odwrotnej (rekonstrukcji). Mówi ona o tym, że cała energia sygnału będąca sumą kwadratów wartości wszystkich próbek jest równa sumie kwadratów wszystkich współczynników uzyskanych w procesie dekompozycji falkowej.

## 1.4. Falki

### 1.4.1. Funkcja Haara

Wykorzystywane w praktyce falki nie mają prostych postaci analitycznych.

Stosunkowo nieskomplikowaną falką jest funkcja Haara. Nie ma ona jednak większego zastosowania praktycznego, obecnie wykorzystuje się ją raczej do celów dydaktycznych. Jej definicja wyraża się następująco

$$\psi(x) = \begin{cases} 1 & 0 \leq x < 0,5 \\ -1 & 0,5 \leq x < 1 \\ 0 & \end{cases} \quad (1.74)$$

Jest jedyną falką, która posiadając własność ortogonalności, posiada również oś antysymetrii. Oznacza to, że odpowiadający jej filtr nie wprowadza nieliniowego prze-

sunięcia fazowego pomiędzy sygnałem na wejściu a sygnałem na wyjściu. Daje to między innymi możliwość kaskadowego łączenia filtrów bez konieczności kompensacji fazy sygnału.

### 1.4.2. Falka Daubechies

Falka Daubechies jest obecnie jedną z najpopularniejszych i najchętniej wykorzystywanych falek. Jej główne atuty to zwarty nośnik o szerokości  $2p-1$ , dokładna aproksymacja funkcji stałych i liniowych oraz stosunkowo prosta postać.

Współczynniki odpowiedzi impulsowej filtrów generujących falę 'db2' o nośniku od 0 do 3 ( $h_0, h_1, h_2, h_3$ ) można wyznaczyć za pomocą zależności (1.75) i (1.76).

Z warunku (1.65) współczynniki muszą spełniać

$$h_0 + h_1 + h_2 + h_3 = \sqrt{2} \quad (1.75)$$

Z warunku ortogonalności wynika

$$h_0 h_2 + h_1 h_3 = 0 \quad (1.76)$$

Otrzymano układ dwóch równań z czterema niewiadomymi. Aby „wypełnić” pozostałe dwa stopnie swobody, należy stworzyć stowarzyszony z ciągiem  $(h_n)$  ciąg  $(g_n)$ , który ma postać  $(h_3, -h_2, h_1, -h_0)$ . Przy założeniu, że będzie on ortogonalny do ciągu stałego  $(1, 1, 1, 1)$  oraz do ciągu liniowego  $(1, 2, 3, 4)$  otrzymamy brakujące dwa równania:

$$h_3 - h_2 + h_1 - h_0 = 0 \quad (1.77)$$

$$h_3 - 2h_2 + 3h_1 - 4h_0 = 0 \quad (1.78)$$

Rozwiązaniem tych równań są współczynniki odpowiedzi impulsowych filtrów kwadraturowych generujące falę 'db2':

$$h_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, \quad h_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, \quad h_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, \quad h_3 = \frac{1-\sqrt{3}}{4\sqrt{2}} \quad (1.79)$$

### 1.4.3. Falki Coiflets

Zaproponowana przez I. Daubechies dekompozycja i rekonstrukcja falkowa charakteryzowała się tym, że wszystkie cztery filtry pochodziły od jednego filtru  $H$ . Filtr ten był filtrem niesymetrycznym i niesymetryczna była odpowiadająca mu falka. Filtry niesymetryczne wprowadzają nieliniowe przesunięcie fazowe pomiędzy sygnałem na wejściu i sygnałem na wyjściu, co w przypadku pewnych zastosowań może być wadą.

Falka Coiflets była pierwszą próbą eliminacji tego zjawiska. Dzięki założeniu, że funkcje skalujące mają posiadać znikające momenty postaci:

$$\int x^l \varphi(x) dx = 0, \text{ dla } l = 1, 2, \dots, p-1 \quad (1.80)$$

falka Coiflets posiada stosunkowo małą asymetrię spośród ortogonalnych falek o zwartym nośniku. Zostało to okupione wydłużeniem nośnika falki, który ma szerokość  $3p-1$ .

#### 1.4.4. Falki biortogonalne

W falkach biortogonalnych definitywnie rozwiązany został problem asymetrii. Tradycyjne podejście do projektowania filtrów zostało uogólnione na przypadek dwóch różnych (niezależnych) filtrów dolnoprzepustowych analizy (dekompozycji)  $H$  i syntezy (rekonstrukcji)  $H'$ , w taki sposób, że filtry mogą być symetryczne i tym samym o liniowym przesunięciu fazowym. Osiąga się to za cenę rezygnacji z ortogonalności falek odpowiadających filtrowi  $H$  (równoważnej warunkowi  $\sum_n h_{2k+n} h_n = 0$ ) na rzecz tzw. biortogonalności, określonej warunkiem

$$\sum_n h'_n h_{2k+n} = 0 \quad (1.81)$$

W konsekwencji przetwarzany sygnał wyrażany jest poprzez transformatę falkową jako

$$f = \sum_m \sum_n \langle f, \psi_{m,n} \rangle \psi'_{m,n} \quad (1.82)$$

Oznacza to, że analiza sygnału dokonywana jest w jednej bazie falkowej, natomiast synteza w drugiej. Obie bazy nie są ortogonalne, ale pary falek, jedna z bazy analizy, druga z bazy syntezy, są względem siebie ortogonalne. Filtry górnoprzepustowe oblicza się ze wzorów:

$$g'_n = (-1)^{n+1} h_{-n+1} \quad (1.83a)$$

$$g_n = (-1)^{n+1} h'_{-n+1} \quad (1.83b)$$

Filtry analizy  $h$  i  $g$  mogą mieć jedną z następujących postaci:

- obydwa filtry są symetryczne i nieparzystej długości i długości ich różnią się o liczbę stanowiącą iloczyn liczby nieparzystej i 2;
- jeden filtr jest symetryczny, a drugi niesymetryczny; długości obydwu są parzyste i takie same lub różnią się o liczbę stanowiącą iloczyn liczby parzystej i 2;
- jeden filtr ma długość nieparzystą, drugi parzystą; obydwa mają wszystkie zera na okręgu jednostkowym; obydwa są symetryczne lub jeden jest symetryczny, a drugi niesymetryczny.

### 1.4.5. Porównanie parametrów falek

W tabeli 1.1 zebrano podstawowe informacje na temat poszczególnych rodzin falek.

**Tabela 1.1**  
Podstawowe parametry falek

Rodzina	Haar	Daubechies	Symlets	Coiflets	Biorthogonal
Skrócona nazwa	'haar'	'db'	'sym'	'coif'	'bior'
Indeks	–	$N$	$N$	$N$	$N_d, N_r$
Przykład	haar	db2	sym4	coif3	bior3.7
Ortogonalność	tak	tak	tak	tak	nie
Biortogonalność	tak	tak	tak	tak	tak
Długość filtru	2	$2N$	$2N$	$6N$	$\max(2N_d, 2N_r)+2$
Symetria	tak	silna asymetria	niewielka asymetria	niewielka asymetria	tak
Moment zaniku falki	1	$N$	$N$	$2N$	$N_r-1$
Moment zaniku funkcji skalującej	–	–	–	$2N-1$	–



## ROZDZIAŁ 2

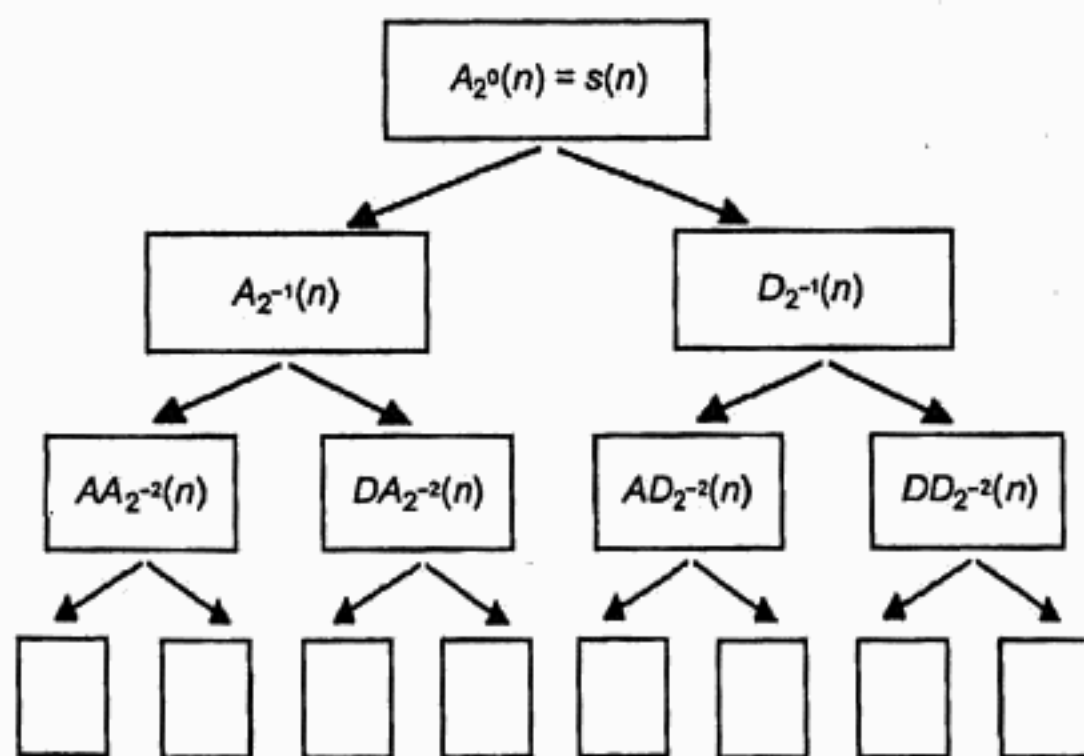
# Rozwinięcia metod dekompozycji sygnału

W rozdziale tym zostaną omówione metody przetwarzania sygnałów wywodzące się z klasycznej transformacji falkowej. Te najbliższe spokrewnione z klasyczną transformacją (pakiety falkowe) polegają na drobnej zmianie piramidowego schematu dekompozycji, co jednak zdecydowanie zmienia właściwości reprezentacji sygnału i otwiera szereg nowych obszarów zastosowań. Z kolei techniki poszukiwania dopasowań, które można nazwać dalszymi krewnymi transformacji falkowej, wykorzystują wprawdzie ideę skalowania funkcji bazowych (nazywanych w tym zastosowaniu elementami słownika), ale pozostawiają znaczną swobodę ich wyboru.

### 2.1. Transformacja *Wavelet Packets*

#### 2.1.1. Dekompozycja sygnału z użyciem pakietów falkowych

Schemat dekompozycji sygnału z użyciem pakietów falkowych wykorzystuje pełne drzewo binarne w miejsce stosowanego w tradycyjnej transformacji falkowej drzewa piramidowego (rys. 2.1).



Rys. 2.1. Schemat pełnej dekompozycji z użyciem pakietów falkowych

Przedmiotem kolejnych etapów dekompozycji są więc nie tylko aproksymacje sygnałów na coraz niższych poziomach rozdzielczości, ale także odpowiadające im sygnały detaliczne.

Podobnie jak w dyskretnej transformacji falkowej, w dekompozycji z użyciem pakietów falkowych wykorzystywane są wyłącznie filtry ortogonalne. Zapewnia to kompletność informacji na wszystkich poziomach dekompozycji, a także jej rozłączność w odpowiadających sobie szeregach aproksymacji i detali. Powstające w dekompozycji falkowej sygnały detali są dyskretnymi szeregami czasowymi zawierającymi informacje o zawartości górnej połowy dekomponowanego pasma ( $f_p/4 \div f_p/2$ ).

Podczas dalszej analizy tych sygnałów warto pamiętać, że ich widmo jest przesunięte o  $f_p/4$  przez operację podpróbkowania.

Transformacja sygnału do dziedziny czasowo-częstotliwościowej jest w tej metodzie dokonywana za pomocą zestawu funkcji zdefiniowanego następująco

$$\begin{cases} W_{2n}(x) = \sqrt{2} \cdot \sum_{k=0}^{2N-1} h(k) \cdot W_n(2x - k) \\ W_{2n+1}(x) = \sqrt{2} \cdot \sum_{k=0}^{2N-1} g(k) \cdot W_n(2x - k) \end{cases} \quad (2.1)$$

gdzie  $W_0(x) = \phi(x)$  oznacza funkcję skalującą,  $W_1(x) = \psi(x)$  – odpowiadającą jej falke, natomiast  $g$  i  $h$  są filtrami lustrzanymi o długości  $2N$  odpowiednio górno- i dolnoprzepustowym.

Pojedynczy etap dekompozycji sygnału detalicznego  $s_D$  przebiega identycznie jak w przypadku aproksymacji. Powstają więc dwa wzajemnie ortogonalne szeregi czasowe, z których jeden  $s_{DA}$  zawiera informację o sygnale w zakresie ( $f_p/4 \div 3 \cdot f_p/8$ ) natomiast drugi  $s_{DD}$  – jej uzupełnienie w zakresie ( $3 \cdot f_p/8 \div f_p/2$ ). Liczba próbek tworzących te szeregi jest połową liczby próbek pierwotnego sygnału detalicznego, tak więc wzrost precyzji analizy w dziedzinie częstotliwości ma miejsce kosztem obniżenia precyzji lokalizacji czasowej. Zgodnie z zasadą nieoznaczoności pole czasowo-częstotliwościowe atomu sygnału pozostaje stałe.

Dekompozycja prowadzona z wykorzystaniem pełnego drzewa binarnego prowadzi do uzyskania  $n$  współczynników (gdzie  $n$  jest długością sygnału), reprezentujących energię w poszczególnych zakresach częstotliwościowych. Ponieważ osiągnięto maksymalną selektywność w dziedzinie częstotliwości, długość wszystkich atomów dekompozycji jest równa długości sygnału, aspekt rozdzielczości czasowej został utracony i nie można już mówić o analizie czasowo-częstotliwościowej. Posortowanie współczynników pełnej dekompozycji narastająco według reprezentowanej częstotliwości prowadzi do otrzymania dyskretnego szeregu widmowej gęstości mocy.

## 2.1.2. Poszukiwanie najlepszego drzewa dekompozycji

Jednym z wniosków z rozważań przeprowadzonych w podrozdziale 2.1.1 jest możliwość otrzymania szeregu widmowej gęstości mocy w wyniku pełnej dekompozycji sygnału z zastosowaniem drzewa binarnego. Z praktycznego punktu widzenia możliwość taka jest mało interesująca, gdyż ten sam rezultat można otrzymać prościej za pomocą dyskretnej transformacji Fouriera. Niemniej, rezultat pełnej dekompozycji wskazuje na istnienie „płynnego przejścia” (w rzeczywistości jest ono ziarniste) z dziedziny czasu do dziedziny częstotliwości. Tworzą je reprezentacje sygnału w dziedzinie czasowo-częstotliwościowej będące rezultatem częściowej dekompozycji.

Ponieważ szeregi reprezentacji sygnałów są wzajemnie ortogonalne (również te, na różnych poziomach dekompozycji), zatrzymanie dekompozycji w dowolnym węźle drzewa binarnego nie powoduje w konsekwencji zmian treści informacyjnej sygnału. Taką samą zawartość posiada więc sygnał w dziedzinie czasu, sygnał całkowicie zdekomponowany oraz każda z jego wersji częściowo zdekomponowanych. Jednym z rezultatów własności dokładnej rekonstrukcji jest zachowanie całkowitej energii sygnału w każdej jego wersji.

Dla dyskretnego ciągu czasowego o długości  $n$  binarne drzewo pełnej dekompozycji zawiera  $n-1$  węzłów. W przypadku dopuszczenia niepełnego drzewa dekompozycji otrzymujemy  $2^n$  wersji reprezentacji sygnału w dziedzinie czasowo-częstotliwościowej. Krańcowymi przykładami są:

- szereg czasowy – „zerowe” drzewo (brak) dekompozycji,
- szereg widmowej gęstości mocy – pełna dekompozycja.

Wszystkie wersje reprezentacji sygnału charakteryzuje identyczna zawartość informacyjna i energetyczna.

Spośród tak wielu równoważnych wersji reprezentacji sygnału można wybrać reprezentację optymalną. Ponieważ przeważnie celem dekompozycji sygnału jest dążenie do przedstawienia go za pomocą jak najmniejszej liczby współczynników, jako kryterium optymalizacji zwykle wybierana jest funkcja opisująca koncentrację energii. Szczęśliwie, sumaryczna energia wszystkich reprezentacji jest identyczna, a więc porównanie energii  $0 < k < n$  największych współczynników różnych reprezentacji pozwala wybrać tę najlepszą.

Wybór najlepszego drzewa dekompozycji dokonuje się automatycznie przez maksymalizację funkcji koncentracji, a więc struktura drzewa dekompozycji dopasowuje się do analizowanego sygnału i jego lokalnych własności. Dlatego połączenie techniki pełnej dekompozycji oraz optymalizacji reprezentacji sygnału nosi nazwę adaptacyjnych pakietów falkowych.

Naturalną dziedziną zastosowań adaptacyjnych pakietów falkowych jest kompresja sygnału. Spośród posortowanych współczynników optymalnej reprezentacji zostają zatrzymane tylko te, zwykle nieliczne, reprezentujące określony procent energii sygnału (np. 95%). Pozostała, zwykle znacząca liczba współczynników jest odrzuca-  
na, a ich brak w procesie rekonstrukcji skutkuje niewielkimi tylko zniekształceniami

sygnału (5%). Niestety, pomimo że reprezentacja sygnału jest optymalna, rzeczywista efektywność kompresji jest znacznie ograniczona przez konieczność przechowywania struktury drzewa dekompozycji. Porządek współczynników nie jest bowiem naturalny jak w szeregu czasowym lub widmowym i wraz z wartościami współczynników dekompozycji należy przechowywać ich współrzędne.

### 2.1.3. Pojęcie entropii

Wybór najlepszego drzewa dekompozycji wymaga zdefiniowania funkcji koncentracji energii. Wartości tej funkcji powinny być duże dla dużej koncentracji energii (niewielki odsetek współczynników ma znaczące wartości), a małe gdy energia jest rozłożona równomiernie w zbiorze współczynników. W praktyce w roli funkcji koncentracji energii występuje entropia, będąca funkcją odwrotną. Tak więc minimalizacja entropii jest metodą wyznaczania reprezentacji optymalnej.

Entropia  $E$  musi spełniać następujące warunki:

$$\begin{aligned} E(0) &= 0 \\ E(x) &= \sum_i E(x_i) \end{aligned} \tag{2.2}$$

gdzie  $x$  jest ciągiem współczynników reprezentacji, a  $i$  numerem współczynnika.

Najczęściej stosowane definicje entropii wymieniono poniżej.

— Entropia Shannona

$$E(x) = -\sum_i x_i^2 \cdot \log(x_i^2), \quad 0 \log(0) = 0 \tag{2.3}$$

— Koncentracja w  $l^p$

$$E(x) = \sum_i |x_i|^p, \quad 1 \leq p \leq 2 \tag{2.4}$$

— Logarytm energii

$$E(x) = \sum_i \log(x_i^2), \quad \log(0) = 0 \tag{2.5}$$

— Entropia progowa

$$E(x) = \sum_i E(x_i), \quad E(x_i) = \begin{cases} 1 & \forall x_i : |x_i| > \epsilon \\ 0 & \forall x_i : |x_i| \leq \epsilon \end{cases} \tag{2.6}$$

Konstruowanie wszystkich możliwych reprezentacji sygnału, aby później z zastosowaniem entropii wybrać reprezentację optymalną, jest bardzo złożone numerycznie.



Dlatego w praktyce minimalizacja entropii jest dokonywana już na etapie dekompozycji poprzez decyzje o dekompozycji (lub nie) kolejnych węzłów drzewa. W tym ujęciu wybór najlepszej reprezentacji jest dokonywany za każdym razem według poniższego schematu, począwszy od najniższego poziomu dekompozycji ( $j = 0$ , rys. 2.2).

1. Jeżeli suma entropii szeregów składowych na poziomie  $j$  jest wyższa od entropii szeregu na poziomie  $j + 1$

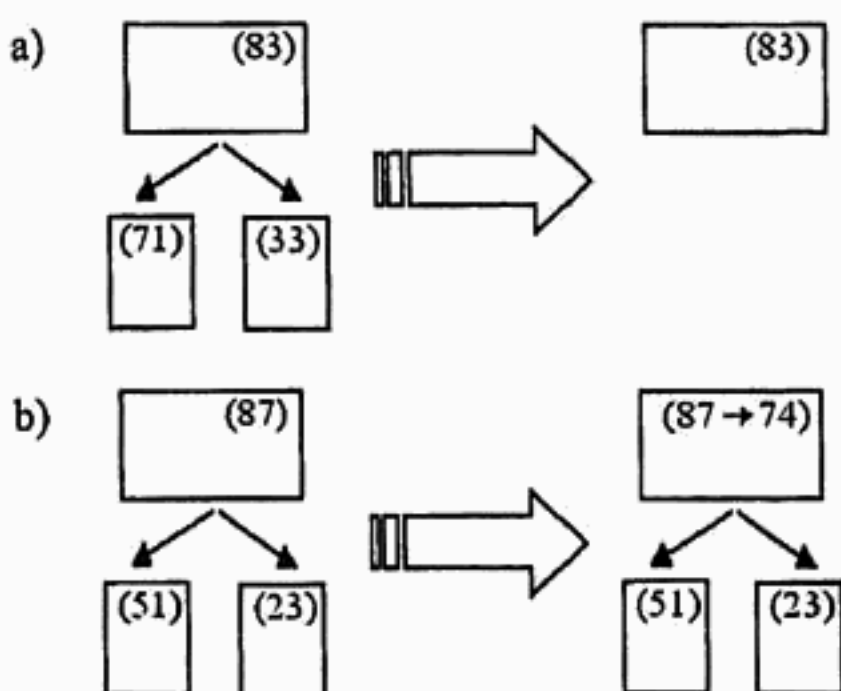
$$E(x, A_{j+1,n}) < E(x, A_{j,2n}) + E(x, A_{j,2n+1}) \quad (2.7)$$

to szeregi składowe na poziomie  $j$  są zastępowane przez szereg na poziomie  $j+1$ .

2. W przeciwnym przypadku entropia szeregu na poziomie  $j+1$  jest zastępowana sumą entropii szeregów składowych na poziomie  $j$

$$E(x, A_{j+1,n}) = E(x, A_{j,2n}) + E(x, A_{j,2n+1}) \quad (2.8)$$

W tym przypadku szeregi składowe na poziomie  $j$  są dołączane do reprezentacji sygnału.

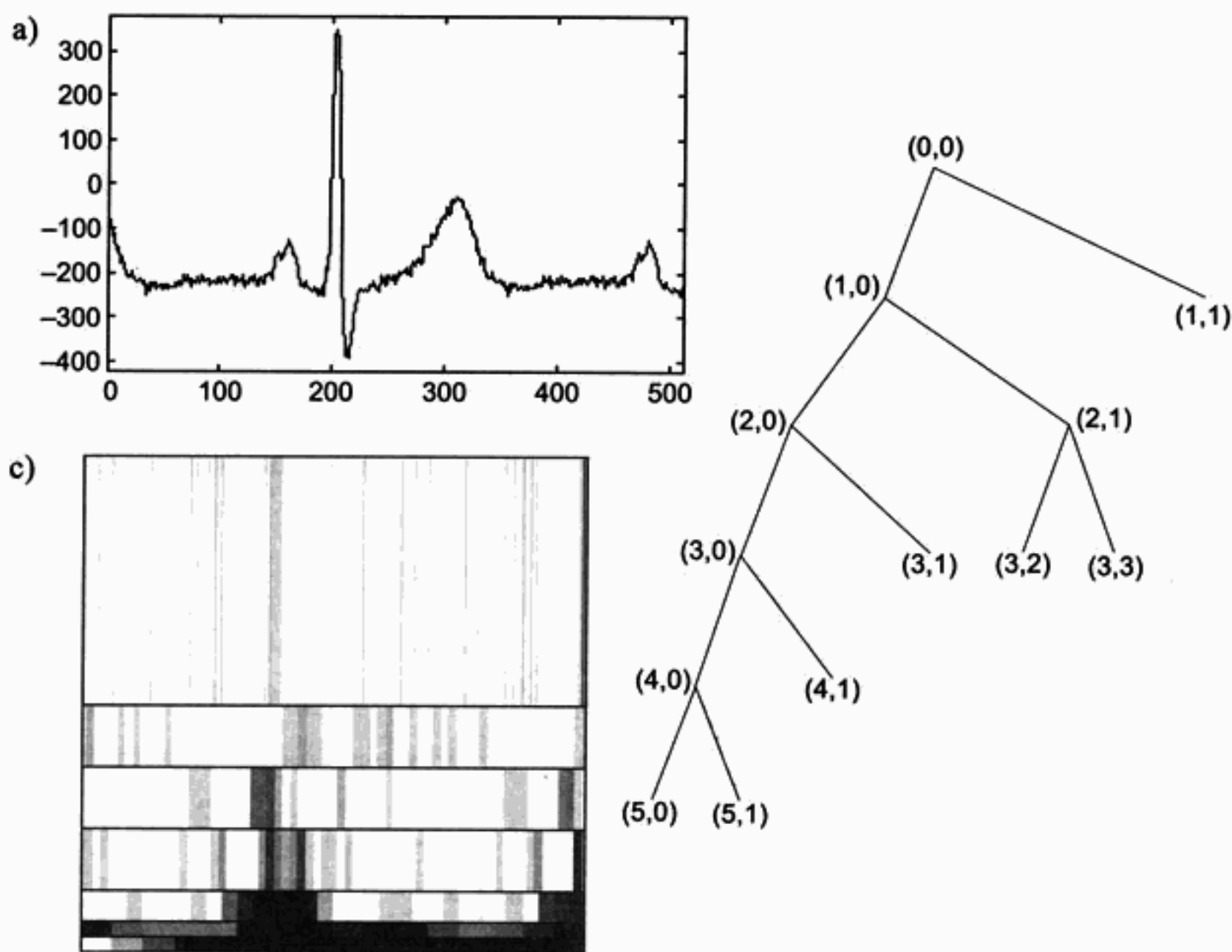


Rys. 2.2. Schemat optymalizacji węzła dekompozycji: a) przypadek zastąpienia reprezentacji składowych przez reprezentację wyższego poziomu o entropii (83) niższej niż suma entropii składowych ( $71+33=104$ ); b) przypadek pozostawienia reprezentacji składowych, których suma entropii ( $51+23=74$ ) jest niższa niż entropia węzła wyższego poziomu (87)

Podany przykład wykorzystuje funkcję entropii zastosowaną do optymalizacji czasowo-częstotliwościowej reprezentacji sygnału, gdyż celem optymalizacji było zawarcie maksymalnie dużego odsetka energii w możliwie jak najmniejszej liczbie atomów (rys. 2.3). Jest to użycie typowe dla kompresji sygnału, a „energetyczny” charakter funkcji entropii odpowiada metodzie estymacji zniekształceń sygnału skompresowanego. Używana w tym celu funkcja różnicy średniokwadratowej sygnału oryginalnego i odtworzonego z kompresji w rzeczywistości reprezentuje procentowy udział energii zniekształceń w energii sygnału.



Możliwe jest odmienne zdefiniowanie kryterium optymalizacji reprezentacji, wówczas identycznie przebiegający proces wyboru najlepszego drzewa wskaże inną z  $2^N$  reprezentacji czasowo-częstotliwościowych jako optymalną.



**Rys. 2.3.** Poszukiwanie optymalnego drzewa dekompozycji: a) fragment sygnału EKG; b) optymalne dla niego drzewo dekompozycji *Wavelet Packets*; c) reprezentacja na powierzchni czasowo-częstotliwościowej

## 2.2. Techniki poszukiwania dopasowań

### 2.2.1. Wstęp

Techniki poszukiwania dopasowań (ang. *matching pursuit*) są rezultatem rezygnacji z ortogonalnej bazy dekompozycji stosowanych przy transformacji falkowej na rzecz funkcji lepiej reprezentujących komponenty poszukiwane w sygnale analizowanym. Ponieważ celem analizy jest wyrażenie zawartości sygnału w jak najmniejszej liczbie współczynników, korzystne jest użycie funkcji analizujących, których kształt odpowiada elementom poszukiwanym. W szczególnie szczęśliwym przypadku dekompozycja może doprowadzić do powierzchni zawierającej wartości zerowe dla wszyst-

kich współczynników z wyjątkiem pojedynczych, znacząco większych, przypadających dokładnie w momencie wystąpienia poszukiwanych komponentów. Analogiczny przypadek mógł wprawdzie mieć miejsce także w analizie Fouriera, gdy analizowany sygnał zawierał np. wyłącznie sinusoidę reprezentowaną przez jeden prążek widma, jednak w technice poszukiwania dopasowań nawet skomplikowane komponenty mogą być reprezentowane w ten sposób.

Naturalnym ograniczeniem klasycznej transformacji falkowej w zastosowaniach biomedycznych jest ograniczony zbiór funkcji analizujących. Choć znane są próby projektowania falek-matek tak, aby swym kształtem przypominały podstawowe komponenty sygnału, będące w użyciu w przeważającej większości przypadków ortogonalna transformacja falkowa lub transformacja z użyciem pakietów falkowych wykorzystują bazę dekompozycji zawierającą minimalną liczbę funkcji bazowych. W efekcie reprezentacja pojedynczego komponentu sygnału składa się z kilku, a niekiedy nawet ze znacznej liczby komponentów o istotnych wartościach. Taka reprezentacja nie jest oczywiście optymalna z punktu widzenia przechowywania oraz interpretacji, natomiast próby pominięcia niektórych atomów o mniejszej energii mogą pozbawić sygnał istotnych niuansów.

Techniki poszukiwania dopasowań zakładają istnienie redundancji w zbiorze funkcji analizujących (słowniku). Ułatwia to precyzyjną reprezentację złożonych sygnałów charakteryzujących się bogactwem kształtów.

Reprezentacja sygnału względem funkcji analizujących tworzących słownik wymaga wyszukania podzbioru takich jego elementów, które zapewniają najlepsze dopasowanie w sensie minimalnoodległościowym. Ponieważ minimum odległości sygnału i funkcji analizującej zostało przyjęte jako miara jakości dopasowania, iteracyjny algorytm poszukiwania dopasowań w każdym kroku ma za zadanie znalezienie takiego elementu słownika, dla którego korelacja z nie zakodowaną częścią sygnału jest maksymalna.

Anegdotyczne porównanie dekompozycji falkowej i poszukiwania dopasowania jest następujące: założmy, że na skraju puszczy poszukujemy słońa. Metoda analogiczna do analizy falkowej polega na przedstawieniu obserwowanej sceny za pomocą podstawowych figur geometrycznych w różnych pozycjach i skalach, a następnie sprawdzeniu, czy figury o istotnie większym zaczernieniu (zwykle jest ich kilka, jeśli nie kilkadziesiąt) układają się w kształt zwierzęcia. Metoda poszukiwania dopasowań jest w tym zastosowaniu znacznie bardziej naturalna: należy wyobrazić sobie kształt słońa, a następnie przeskalowując go i przesuwając stwierdzić dopasowanie z obserwowaną sceną. Znaczna większość otrzymanych współczynników będzie zerowa, z wyjątkiem jednego (jest jeden słoń) lub co najwyżej kilku (w przypadku stada słońi) o znacznych wartościach. Istotne współczynniki mogą pojawić się dla różnych wartości przesunięć lub dla różnych skal (stado słońi z małymi).

### 2.2.2. Algorytm poszukiwania dopasowań

Algorytm poszukiwania dopasowań jest algorytmem iteracyjnym realizującym liniowe rozwinięcie sygnału  $f$  względem zbioru funkcji analizujących  $g$  będących elementami słownika. Pierwszym krokiem jest wyszukanie elementu  $g_{j_0}$ , na który rzut

ortogonalny sygnał jest maksymalny. Następnie sygnał  $f$  zostaje przedstawiony jako suma produktu rzutu  $f$  na  $g_{\gamma 0}$  i  $g_{\gamma 0}$  oraz pozostałej zawartości sygnału (residuum rzędu pierwszego).

$$f = \langle f, g_{\gamma 0} \rangle g_{\gamma 0} + R^1 f \quad (2.9)$$

W kolejnej iteracji residuum  $R^1$  zajmuje miejsce sygnału  $f$  i proces poszukiwania elementu  $g_{\gamma 1}$  jest powtarzany. W ogólnym przypadku residuum rzędu  $n+1$  jest wyznaczane za pomocą dekompozycji residuum rzędu  $n$  względem najlepiej dopasowanego elementu słownika  $g_{\gamma n}$

$$R^n f = \langle R^n f, g_{\gamma n} \rangle g_{\gamma n} + R^{n+1} f \quad (2.10)$$

Po  $m$  iteracjach wyjściowy sygnał można przedstawić w postaci

$$f = \sum_{n=0}^{m-1} \langle R^n f, g_{\gamma n} \rangle g_{\gamma n} + R^m f \quad (2.11)$$

lub równoważnej

$$f = \sum_{n=0}^{m-1} \langle R^n f, g_{\gamma n} \rangle g_{\gamma n} + R^m f \quad (2.12)$$

Wynikiem algorytmu jest reprezentacja sygnału  $f$  będąca sumą składowych rozwinięcia względem elementów słownika, wybranych z uwzględnieniem najlepszego dopasowania do jego residuów. Można wykazać, że pomimo nieliniowości rozwinięcie takie zachowuje energię sygnału.

$$\|f\|^2 = \sum_{n=0}^{m-1} \left| \langle R^n f, g_{\gamma n} \rangle \right|^2 + \|R^m f\|^2 \quad (2.13)$$

Ponieważ w miarę dekompozycji poszczególne wzorce są wyczerpywane z sygnału, residua kolejnych rzędów zawierają coraz mniej energii

$$\lim_{m \rightarrow \infty} \|R^m f\| = 0 \quad (2.14)$$

a więc reprezentacja sygnału za pomocą elementów słownika jest coraz bliższa sygnałowi oryginalnemu.

W praktycznych zastosowaniach ograniczeniem metody poszukiwania dopasowań jest zwolnienie zbieżności energii do zera dla kolejnych residuów. W zależności od

zawartości sygnału i funkcji analizujących tworzących słownik, na pewnym etapie dekompozycji poszukiwanie kolejnych wzorców skutkuje niewielkimi współczynnikami dopasowania, a energia residuum kolejnego rzędu jest nieznacznie tylko mniejsza niż w poprzedniej iteracji. Sytuacja taka zachodzi wtedy, gdy wszystkie komponenty sygnału znajdujące swą reprezentację w słowniku zostały już „wyjawione”, ale w sygnale pozostają jeszcze składowe niepodobne do żadnego z elementów słownika. W celu ograniczenia liczby bezproduktywnych iteracji warunek końca dla algorytmu dopasowania definiuje się jako przekroczenie założonej wartości progowej (np. 0,1) przez współczynnik stosunku energii  $\lambda(n)$

$$\lambda(n) = \frac{\langle R^n f, g_{\gamma n} \rangle}{\|R^n f\|} \quad (2.15)$$

Po zatrzymaniu iteracyjnego algorytmu poszukiwania dopasowań, pozostałe residuum rzędu  $m$  jest zwane błędem aproksymacji. Zawiera ono komponenty sygnału nie znajdujące reprezentacji wśród funkcji analizujących słownika.

### 2.2.3. Dobór elementów słownika

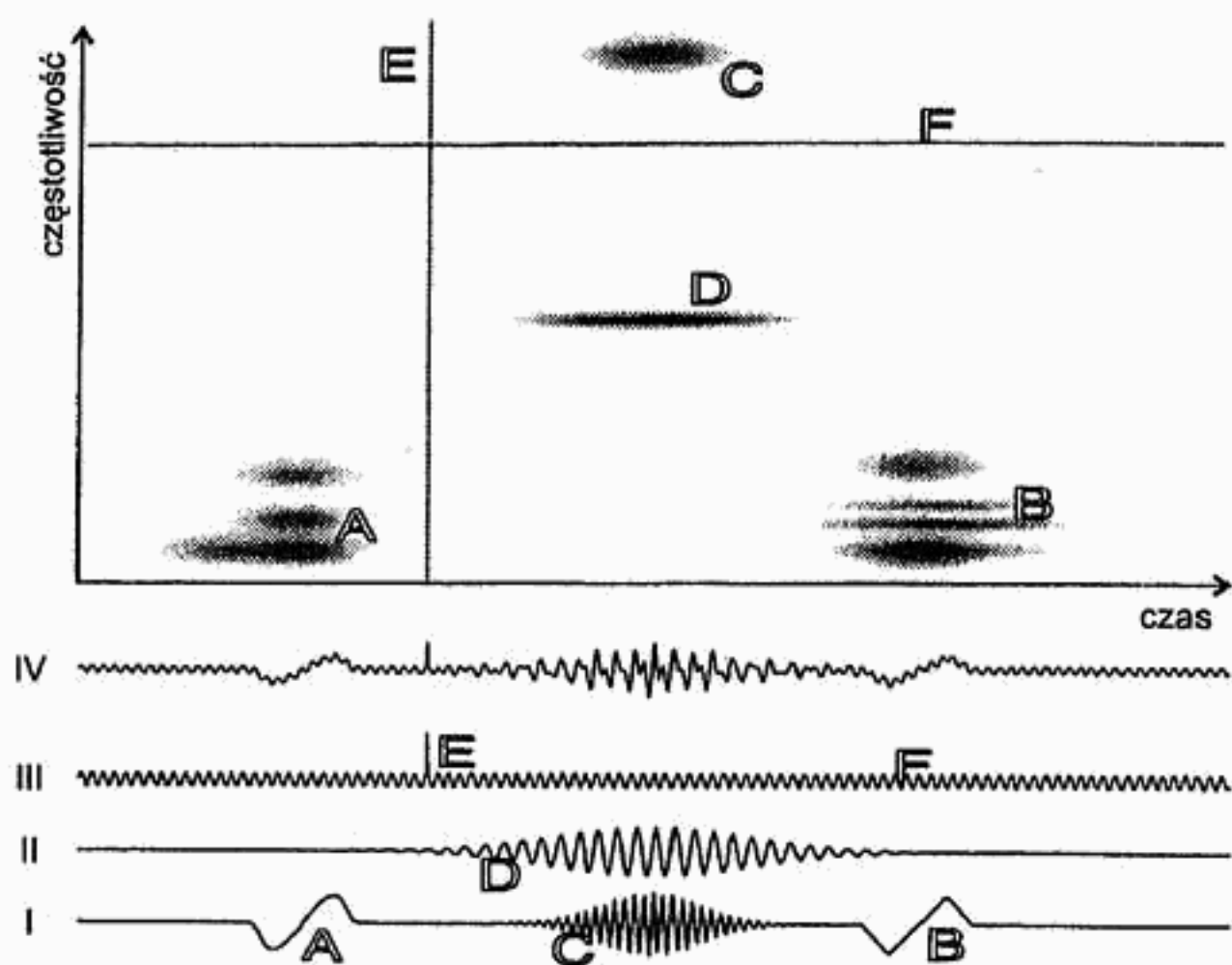
Powodzenie algorytmu poszukiwania dopasowań i końcowy błąd aproksymacji silnie zależą od zawartości słownika. Niewłaściwy wybór funkcji słownika może spowodować, że zaskakująco duża porcja sygnału pozostaje błędem aproksymacji. Jego elementy, np.  $N = 10^6$  funkcji, generowane są automatycznie. Zwykle jednak końcową reprezentację sygnału ogranicza się do znacznie mniejszej ilości funkcji analizujących (np.  $M = 100$ ), które wybrane są tak, że opisują największy procent energii spośród wszystkich podzbiorów funkcji słownika. Wybór optymalnego zestawu funkcji tworzących reprezentację sygnału jest bardzo skomplikowany. Może się on opierać na minimalizacji różnicy sygnału i sumy jego rzutów na elementy słownika, jednak wykazano, że złożoność obliczeniowa takiego algorytmu jest narastająca szybciej niż każdego wielomianu (jest to tzw. problem NP-trudny).

Pozostaje więc intuicyjny dobór elementów słownika w aspekcie analizowanego sygnału lub stosowanie słowników Gabora. Swą nazwę zawdzięczają one temu, że elementami słownika są funkcje Gabora – gaussoida modulowana przebiegiem sinusoidalnym. W tym przypadku upada jednak podstawowa zaleta technik poszukiwania dopasowań, jaką jest reprezentacja sygnałów o skomplikowanych kształtach za pomocą niewielkiej liczby współczynników.

Z powyższych rozważań wynika wniosek, że techniki poszukiwania dopasowań nie oferują dekompozycji w pełni ortogonalnej, ortogonalne są tylko kolejne rzuty dzięki czemu spełniona jest zasada zachowania energii. Próba złożenia otrzymanych reprezentacji nie doprowadzi do dokładnej rekonstrukcji sygnału oryginalnego. Najbardziej trafnym zastosowaniem tych technik jest więc poszukiwanie konkretnych,



znanych *a priori* komponentów w sygnale, określanie ich udziału energetycznego i momentów wystąpień. W przypadku analizy sygnałów nieznanych, słowniki Gabora, wprowadzie zespolone, są najlepszym z dostępnych rozwiązań, gdyż funkcje analizujące minimalizują nieoznaczoność wynikającą z zasady Heisenberga (rys. 2.4).



Rys. 2.4. Analiza sygnału testowego za pomocą poszukiwania dopasowań: a) reprezentacja sygnału na powierzchni czasowo-częstotliwościowej; b) sygnał testowy (IV) i jego komponenty (I–III)



## ROZDZIAŁ 3

# Analiza falkowa – od teorii do implementacji

### 3.1. Falki o nośniku zwartym

Analiza sygnału za pomocą funkcji należących do rodziny falek analizujących, zróżnicowanych w aspekcie skalowym i czasowym, następuje poprzez stwierdzenie stopnia lokalnego podobieństwa sygnału do każdej z tych funkcji. Odbywa się to przy wykorzystaniu operacji splotu sygnału z falką (konwolucji) w dziedzinie czasu lub mnożenia ich widm.

$$y(n) = \sum_{k=1}^N h(k) \cdot x(n-k) \quad (3.1a)$$

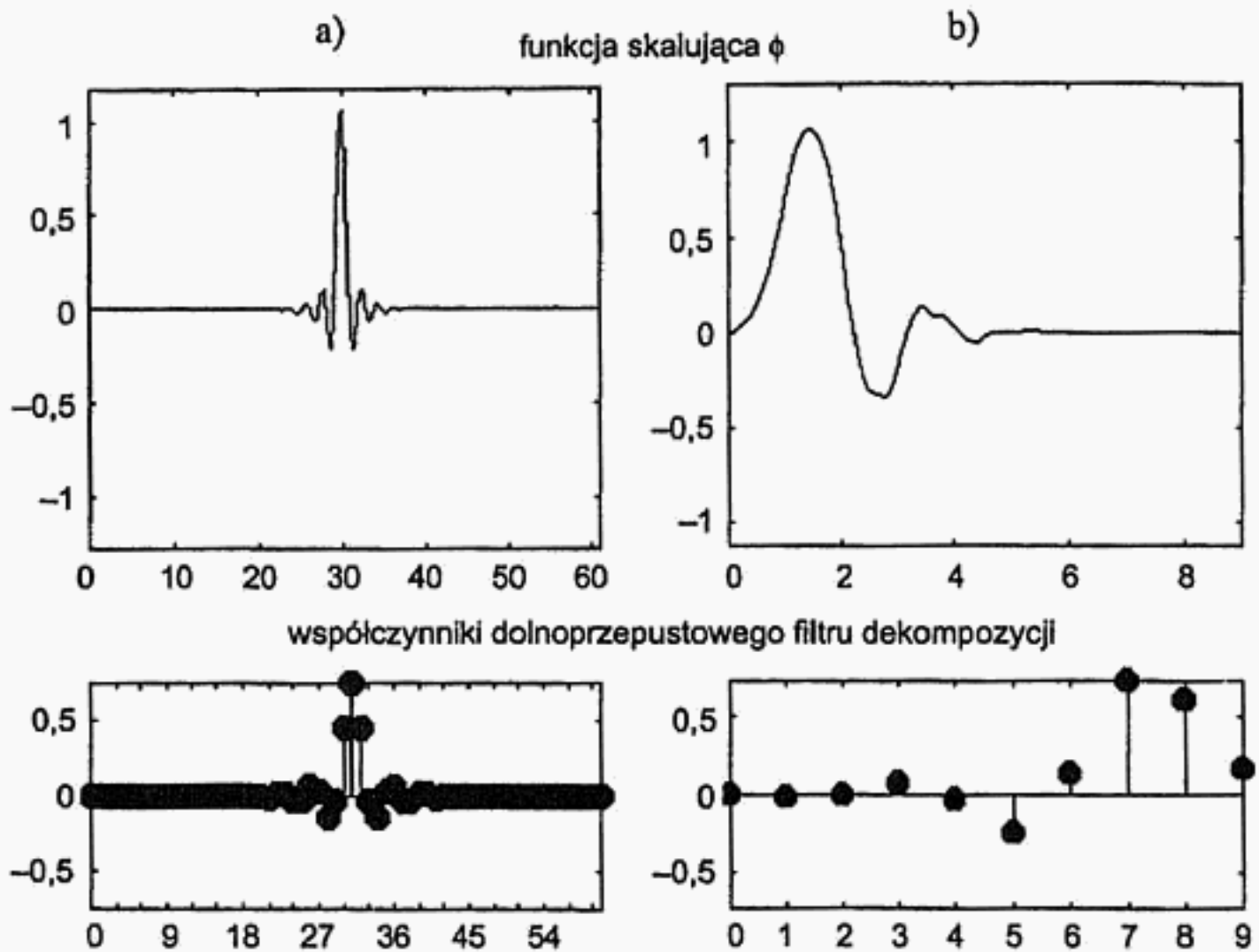
$$Y(\omega) = H(\omega) \cdot X(\omega) \quad (3.1b)$$

Niezależnie od metody, zakładana jest nieskończona długość obu porównywanych ciągów, co spełnione jest tylko przez falkę, natomiast odnośnie sygnału analizowanego przyjmowane jest dodatkowe założenie. Podobnie jak w przypadku transformacji Fouriera, posiadany odcinek sygnału uważany jest za okres podstawowy nieskończonego ciągu okresowego. Założenie to wiąże się ze „sklejeniem” obu brzegów sygnału (bezpośrednio po zakończeniu porcja sygnału zaczyna się od nowa), czego następstwem jest tzw. efekt brzegowy będący rezultatem nieciągłości sygnału (lub jego pochodnych) w miejscu sklejenia.

Aspekt lokalnego podobieństwa wymaga precyzyjnego określenia zasięgu analizy, a więc czasowego i częstotliwościowego zakresu, w którym falka analizująca reprezentuje sygnał. Jest to konieczne do poprawnego „podziału ról” pomiędzy poszczególne falki. Niestety, matematyczna definicja sformułowana na użytek ciągłej transformacji falkowej wymagająca od falki-matki, aby była równa zero, prawie wszędzie nie precyzuje zakresu istotnych wartości, stąd najczęściej do wyznaczenia granic atomów czasowo-częstotliwościowych przyjmowany jest zakres wyznaczany przez 3 dB spadek obwiedni falki w dziedzinie czasu i częstotliwości.

Istnienie precyzyjnych granic zakresu analizy falki jest szczególnie istotne w dyskretnych systemach ortogonalnych, w których zakłada się wykorzystanie własności

idealnej rekonstrukcji sygnału. W tym przypadku falka w dziedzinie czasu oraz jej widmo muszą być reprezentowane przez skończoną liczbę niezerowych współczynników, przy założeniu zerowych wartości pozostałych, definiujących falkę w całej dziedzinie. Zaokrąglenia bardzo małych wartości brzegowych falki w każdej z dziedzin powodują utratę danych, a więc w przypadku oczekiwania idealnej rekonstrukcji sygnału nie są dopuszczalne. Tak więc główną zaletą falek o nośniku zwartym jest pewność, że poza skończoną liczbą wartości, pozostałe są dokładnie zerowe, a zatem ich pominięcie nie wpłynie na utratę własności dokładnej rekonstrukcji sygnału (rys. 3.1).

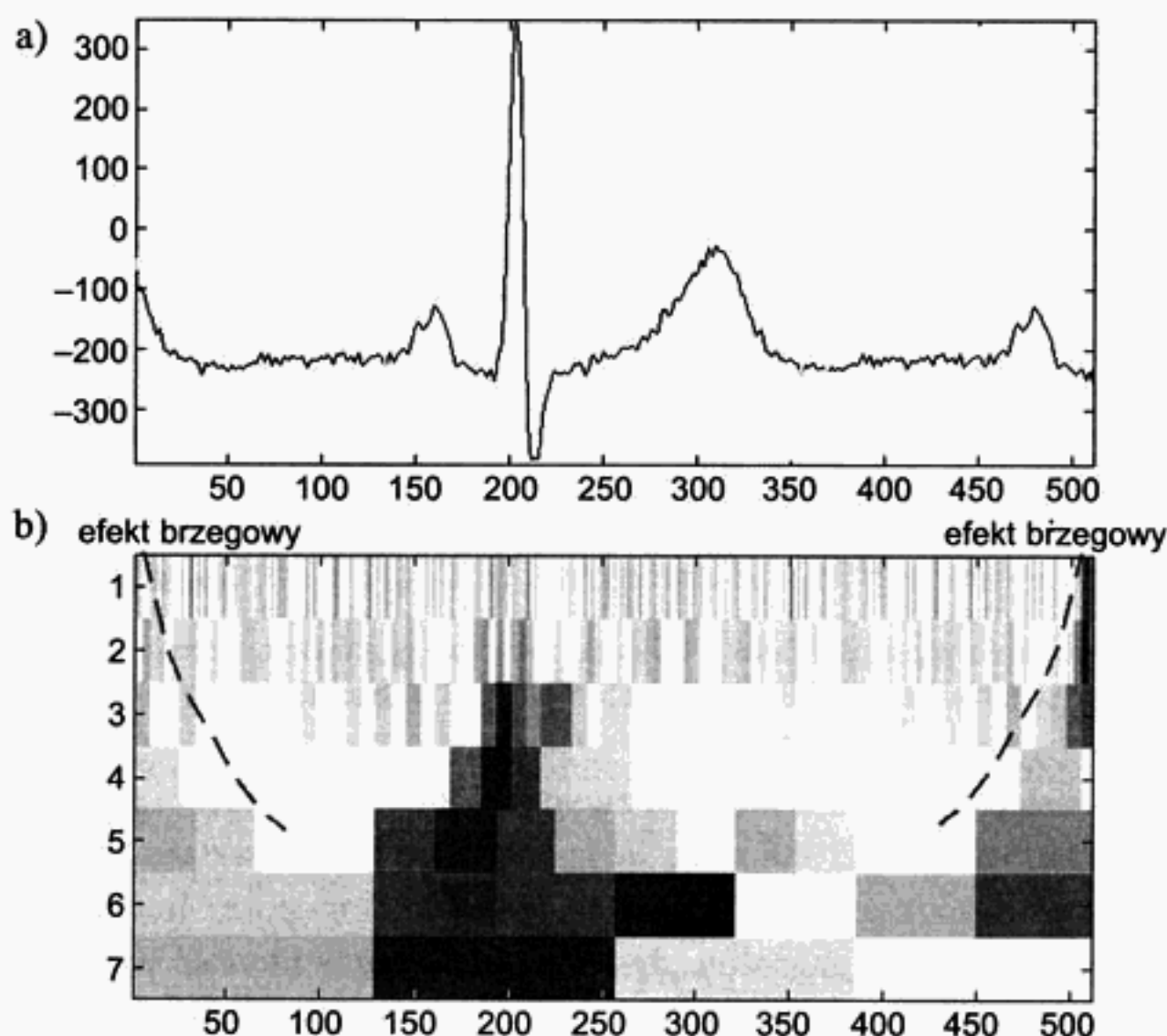


**Rys. 3.1.** Porównanie nośników falek: a) falka Meyera o nośniku rozmytym w dziedzinie czasu; b) falka Daubechies o nośniku zwartym

Niezerowe wartości reprezentacji falki w każdej z dziedzin stanowią jej nośnik. O falce-matce posiadającej wszystkie współczynniki równe dokładnie zero z wyjątkiem  $N$  kolejnych niezerowych wartości powiemy, że jest falką o nośniku zwartym o długości  $N$ . Długość nośnika, wyrażona liczbami naturalnymi, precyzyjnie wyznacza zakres analizy czasowej i częstotliwościowej falki. Zgodnie z zasadą nieoznaczoności długość nośnika nie może być dowolnie mała jednocześnie w dziedzinie czasu i częstotliwości.

Efekt brzegowy (rys. 3.2) zniekształca rezultaty analizy zawsze, gdy nośnik falki pokrywa się z punktem uciąglenia („sklejenia”) sygnału. W zasięg analizy jednej falki wchodzi wtedy zarówno punkty końcowe porcji sygnału, jak i punkty początkowe, a więc próbki bez wzajemnego naturalnego następstwa w sygnale analizowanym. Dla-

tego długość nośnika falki nabiera istotnego znaczenia w przypadku analizy sygnałów o niewielkiej długości, porównywalnej z długością nośnika. Zwykle podawanym przykładem jest linia obrazu (składająca się np. z 256 pikseli), ale problem ma bardziej ogólną naturę, gdyż w przypadku pełnej dekompozycji piramidowej aproksymata na ostatnich poziomach dekompozycji zawsze ma niewielką długość.

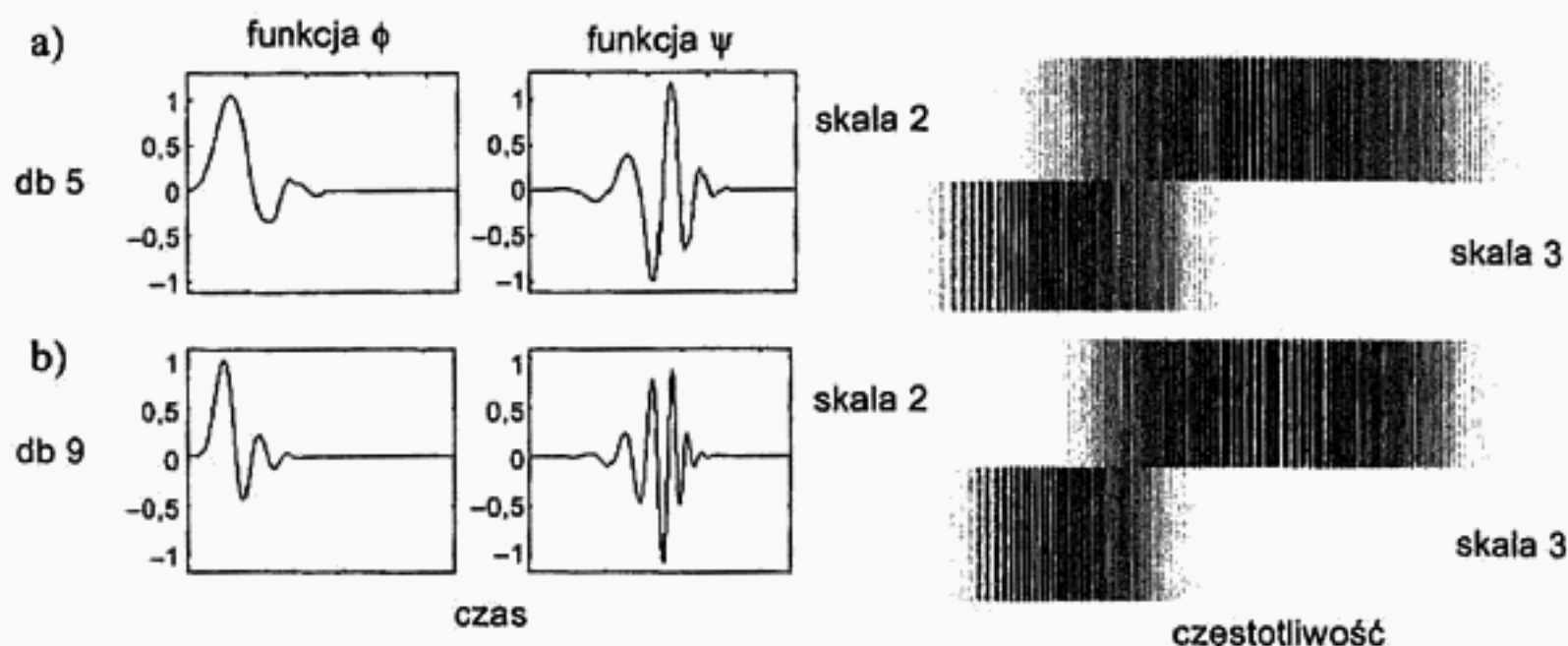


**Rys. 3.2.** Ilustracja efektu brzegowego na przykładzie czasowo-częstotliwościowej reprezentacji fragmentu zapisu EKG: a) sygnał w dziedzinie czasu; b) powierzchnia czasowo-częstotliwościowa

Tak więc niektóre zastosowania transformacji falkowej implikują użycie falek o bardzo krótkim nośniku (2–5 współczynników). Ponieważ rozdzielczość częstotliwościowa takich falek jest bardzo rozmyta, komponenty o znacznej przewadze częstotliwości podstawowej lub nawet złożone tylko z jednej częstotliwości są reprezentowane w sąsiednich pasmach. Jednakże nawet wtedy, dzięki ortogonalności funkcji bazowych, reprezentacje te są komplementarne, co gwarantuje możliwość dokładnej rekonstrukcji sygnału oryginalnego. Kosztem utraty własności rozdzielczych w dziedzinie częstotliwości osiągnąć można znaczne ograniczenie zniekształceń analizy przez efekt brzegowy.

Jeżeli sygnał jest długi, w aspekcie czasowym interesujące obserwatora komponenty są bardziej oddalone od zniekształconych brzegów, ale są także w aspekcie częstotliwościowym bardziej odległe od niskich pasm, dla których długość aproksymacji jest porównywalna z długością nośnika falki. W tych przypadkach stosowanie falki

o krótkim nośniku nie jest już koniecznością, a zatem powinny zostać użyte funkcje o dłuższym nośniku, zapewniające lepszą separację częstotliwościowych komponentów sygnału (rys. 3.3). Nośnik tych falek nadal pozostaje zwarty gwarantując możliwość dokładnej rekonstrukcji sygnału oryginalnego i składa się z kilkunastu współczynników.



**Rys. 3.3.** Porównanie własności falek o nośniku zwartym w dziedzinie czasu i częstotliwości:  
a) falka Daubechies rzędu 5; b) falka Daubechies rzędu 9

Sposób konstrukcji falek o nośniku zwartym dowolnej długości został zaproponowany przez I. Daubechies w 1988 roku. W następstwie tego powstało kilka rodzin falek o nośniku zwartym i różnych własnościach. Ze względu na coraz liczniejsze praktyczne zastosowania w analizie sygnałów dyskretnych, a także łatwość projektowania stosownie do potrzeb, falki o nośniku zwartym są obecnie używane znacznie chętniej niż zdyskretyzowane wersje ciągłych transformacji falkowych (np. Morleta czy Meyera).

## 3.2. Projektowanie własnych falek

W przeciwieństwie do wcześniejszych transformacji sygnału, transformacja falkowa używa otwartego zbioru funkcji analizujących. Już zastosowana przez Morleta falka-matka była wynikiem pewnego arbitralnego wyboru podyktowanego zastosowaniem, a dopiero rozpowszechnienie metod analizy falkowej i ewidentna możliwość wykorzystania w tej roli różnych funkcji doprowadziły do sformułowania ogólnych reguł konstruowania falek. Fakt, że falka-matka zaproponowana przez Morleta niedokładnie spełnia podstawowy postulat o zerowej energii w całej dziedzinie, nie przeszkadzał w efektywnym jej zastosowaniu do analizy sygnałów sejsmicznych.

Ponieważ zagadnienia analizy czasowo-częstotliwościowej i filtracji sygnałów dyskretnych są pokrewne, projektowanie własnych falek analizujących wykorzystuje techniki projektowania filtrów. Rodzina falek analizujących może więc pochodzić



z banku filtrów. W szczególności interesujące są filtry dzielące pasmo sygnału na połowę w sposób ortogonalny, ponieważ zapewnia to dokładną rekonstrukcję sygnału oraz umożliwia wykorzystanie piramidowego algorytmu dekompozycji. Reprezentacja falek analizujących za pomocą filtrów ma dodatkowo jeszcze i tę zaletę, że falka-matka może być opisana współczynnikami odpowiedzi impulsowej filtru. W przypadku falki o nośniku zwartym jest to zaledwie kilka niezerowych wartości.

Warunki, jakie muszą spełniać filtry, aby zachować perfekcyjną rekonstrukcję oraz ortogonalność funkcji bazowych, są zdefiniowane zależnościami (3.2) i (3.3), a ponadto ilość współczynników filtru musi być liczbą parzystą.

$$\sum_{n=1}^N h(n) = \sqrt{2} \quad (3.2)$$

$$\sum_{n=1}^N h(n) \cdot h(n - 2k) = \begin{cases} 1 & \text{dla } k = 0 \\ 0 & \text{w pozostałych przypadkach} \end{cases} \quad (3.3)$$

Współczynniki stowarzyszonego filtru  $g(n)$  można obliczyć ze wzoru

$$g(n) = (-1)^n \cdot h(N - n) \quad (3.4)$$

System biortogonalny, będący uogólnieniem systemu ortogonalnego, umożliwia projektowanie symetrycznych falek i funkcji skalujących. Mimo dość poważnej wady, jaką jest niespełnienie zasady zachowania energii, systemy biortogonalne są chętniej projektowane, gdyż umożliwiają łatwiejszy i bardziej elastyczny wybór parametrów analizy. System biortogonalny tworzą cztery kwadraturowe filtry lustrzane, z których para  $g(n) - h(n)$  używana jest do syntezy, a para  $\tilde{g}(n) - \tilde{h}(n)$  do analizy sygnału.

Filtry  $h(n)$  i  $\tilde{h}(n)$  powinny spełniać następujące zależności (3.5) i (3.6), a ponadto różnica ilości współczynników filtrów musi być liczbą parzystą.

$$\sum_{n=1}^N h(n) = \sum_{n=1}^N \tilde{h}(n) = \sqrt{2} \quad (3.5)$$

$$\sum_{n=1}^N \tilde{h}(n) \cdot h(n - 2k) = \begin{cases} 1 & \text{dla } k = 0 \\ 0 & \text{w pozostałych przypadkach} \end{cases} \quad (3.6)$$

Współczynniki stowarzyszonych filtrów  $g(n)$  oraz  $\tilde{g}(n)$  można obliczyć ze wzoru

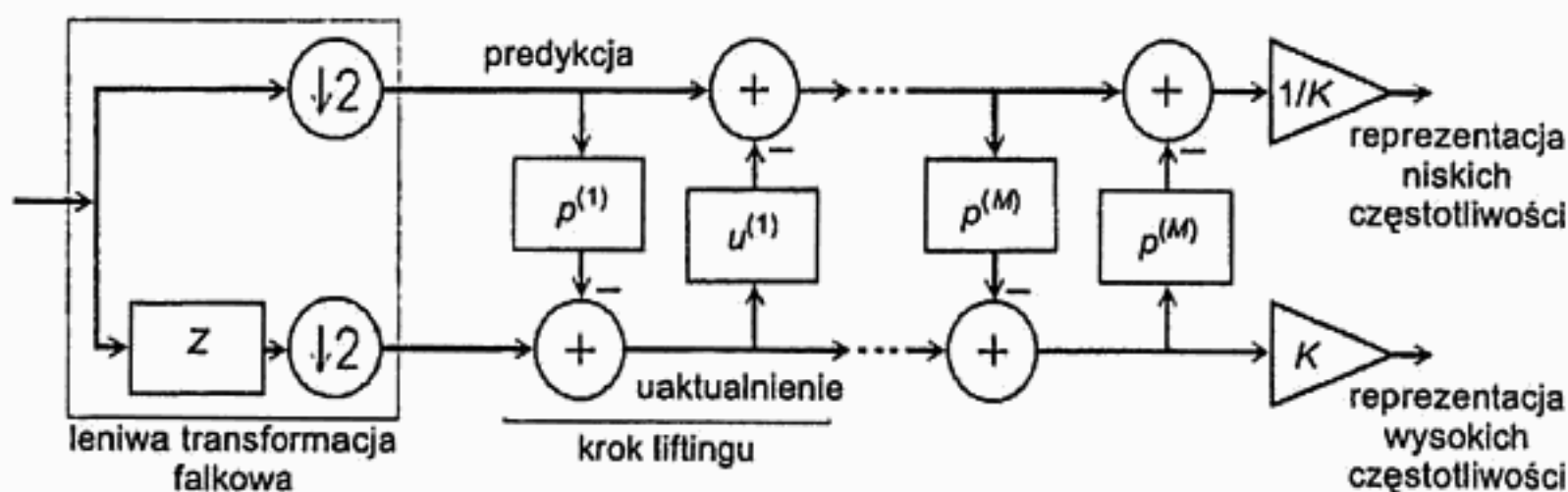
$$\begin{cases} \tilde{g}(n) = (-1)^n \cdot h(1 - n) \\ g(n) = (-1)^n \cdot \tilde{h}(1 - n) \end{cases} \quad (3.7)$$



Spełnienie powyższych warunków pozostawia pewną ilość stopni swobody w doborze współczynników zależną od długości nośnika. W systemach ortogonalnych można je wykorzystać do osiągnięcia założonej regularności falki, a w systemach biortogonalnych – do uzyskania symetrii funkcji analizujących.

### 3.3. Predykcyjna transformacja falkowa

Projektowanie własnych systemów dekompozycji falkowej może wykorzystywać także schemat predykcyjny. W schemacie tym rozdzielenie ciągów reprezentacji aproksymacji i detali na każdym poziomie dekompozycji następuje z użyciem prostych filtrów, a następnie własności widmowe tych ciągów są poprawiane za pomocą używanych naprzemiennie operacji predykcji i uaktualnień (rys. 3.4). Każdy bank filtrów używany w klasycznej transformacji falkowej może zostać przekształcony do postaci predykcyjnej, tak więc predykcyjna transformacja falkowa stanowi uogólnienie dotychczasowych doświadczeń w zakresie ortogonalnej dekompozycji sygnału.



Rys. 3.4. Schemat jednego etapu transformacji falkowej w wersji predykcyjnej

Wprowadzenie predykcyjnej transformacji falkowej wygodnie jest rozpocząć od własności filtru dolnoprzepustowego  $h$ . Ponieważ jest to filtr o skończonej odpowiedzi impulsowej (typu FIR), tylko skończona liczba definiujących go współczynników  $k_2 - k_1$  jest różna od zera. Jego reprezentacja wielomianowa w dziedzinie transformaty  $Z$

$$h(z) = \sum_{k=k_1}^{k_2} h_k \cdot z^{-k} \quad (3.8)$$

może być przedstawiona w postaci polifazowej

$$h(z) = h_p(z^2) + z^{-1} \cdot h_n(z^2) \quad (3.9)$$

gdzie  $h_p(z)$  zawiera parzyste, a  $h_n(z)$  nieparzyste współczynniki filtru  $h(z)$ :

$$h_p(z) = \sum_{k=k_1}^{k_2} h_{2k} \cdot z^{-k} \quad (3.10a)$$

$$h_n(z) = \sum_{k=k_1}^{k_2} h_{2k+1} \cdot z^{-k} \quad (3.10b)$$

Definiując analogicznie postać polifazową stowarzyszonego filtru górnoprzepustowego  $g(n)$ , można otrzymać macierz polifazową postaci

$$P(z) = \begin{bmatrix} h_p(z) & g_p(z) \\ h_n(z) & g_n(z) \end{bmatrix} \quad (3.11)$$

Zwykle oczekuje się, aby filtry stowarzyszone posiadały własność dokładnej rekonstrukcji, którą dla banku filtrów wyraża się następująco:

$$h(z) \cdot \tilde{h}(z^{-1}) + g(z) \cdot \tilde{g}(z^{-1}) = 2 \quad (3.12)$$

$$h(z) \cdot \tilde{h}(-z^{-1}) + g(z) \cdot \tilde{g}(-z^{-1}) = 0 \quad (3.12b)$$

W przypadku macierzy polifazowych  $P(z)$  – używanej do syntezy, oraz  $\bar{P}(z)$  – używanej do analizy sygnału, warunek dokładnej rekonstrukcji przyjmie postać

$$P(z) \cdot \bar{P}(z^{-1})^t = I \quad (3.13)$$

co jest spełnione tylko wtedy, gdy wyznacznik macierzy  $P(z)$  jest równy 1. Parę filtrów  $h(z)$  i  $g(z)$  nazwiemy parą komplementarną. W konsekwencji również filtry stowarzyszone będą parą komplementarną.

Najprostszym sposobem spełnienia warunku dokładnej rekonstrukcji jest przypadek, gdy  $P(z) = I$ . Wówczas  $h(z) = \tilde{h}(z) = 1$  oraz  $g(z) = \tilde{g}(z) = z^{-1}$ , a transformacja dokonuje jedynie separacji próbek parzystych i nieparzystych do odrębnych ciągów nie posiadających żadnych własności widmowych. Przekształcenie to nazwane zostało leniwą transformacją falkową i jest zwykle używane jako operacja poprzedzająca filtrację, a więc rozdzielenie pasma za pomocą filtrów  $h_0(z)$  i  $g_0(z)$  pochodnych od filtrów  $h(z)$  i  $g(z)$ .

Na podstawie filtrów  $h(z)$  i  $g(z)$  możliwe jest konstruowanie filtrów pochodnych  $h_0(z)$  i  $g_0(z)$  bez utraty komplementarności. I tak:

$$g_0(z) = g(z) + h(z) \cdot s(z^2) \quad (3.14)$$

gdzie  $s(z)$  jest filtrem o skończonej odpowiedzi impulsowej (FIR), a więc

$$s(z) = \sum_k s_k \cdot z^{-k} \quad (3.15)$$

oraz analogicznie

$$h_0(z) = h(z) + g(z) \cdot t(z^2) \quad (3.16)$$

gdzie  $t(z)$  jest filtrem o skończonej odpowiedzi impulsowej, a więc

$$t(z) = \sum_k t_k \cdot z^{-k} \quad (3.17)$$

Każdy ze konstruowanych w ten sposób filtrów  $h_0(z)$  jest komplementarny do  $g(z)$ , a każdy z filtrów  $g_0(z)$  jest komplementarny do  $h(z)$ , właściwość dokładnej rekonstrukcji jest więc zachowana także dla „poprawionych” ciągów reprezentacji aproksymacji i detali.

Etapy naprzemiennego poprawiania własności widmowych tych ciągów, zwane odpowiednio liftingiem (konstrukcja filtru  $g_0(z)$ ) oraz liftingiem dualnym (konstrukcja filtru  $h_0(z)$ ) są realizowane z użyciem operacji predykcji i uaktualniania. Ponieważ dla każdej pary komplementarnej  $h(z)$  i  $g(z)$  istnieją wielomiany  $s(z)$  i  $t(z)$  oraz stała  $K$ , takie że

$$P(z) = \prod_{i=1}^m \begin{bmatrix} 1 & s_i(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ t_i(z) & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & \frac{1}{K} \end{bmatrix} \quad (3.18)$$

każdy falkowy bank filtrów może zostać przedstawiony za pomocą leniwej transformaty falkowej,  $m$  operacji naprzemiennego liftingu oraz skalowania.

Predykcyjna transformacja falkowa ma liczne zalety:

- jest zawsze odwracalna (nawet gdy predykcja i uaktualnienie są operacjami nieliniowymi),
- może być zrealizowana przez algorytm szybkiej transformacji falkowej o znacznie zredukowanej złożoności obliczeniowej,
- może być zrealizowana w systemach dopuszczających użycie wyłącznie reprezentacji całkowitoliczbowej.

### 3.4. Transformacja falkowa w reprezentacji całkowitoliczbowej

Z punktu widzenia zastosowań w rzeczywistych systemach przetwarzania sygnału, dotychczas rozważane transformacje mają zasadniczą wadę. Reprezentacja sygnału w dziedzinie czasu jest wyrażona liczbami całkowitymi (lub nawet naturalnymi), gdyż jej źródłem jest zwykle przetwornik analogowo-cyfrowy. Tymczasem reprezentacja czasowo-częstotliwościowa używa wartości zespolonych lub w przypadku czysto rzeczywistych rodzin falkowych – wartości rzeczywistych. Ponieważ istnieją zasadnicze różnice w sposobie przechowywania i obróbki wartości całkowitych i rzeczywistych w maszynie cyfrowej, transformacja falkowa związana ze zmianą formatu sygnału ze stałoprzecinkowego na zmiennoprzecinkowy implikuje szereg istotnych konsekwencji:

- zakres wartości formatu stałoprzecinkowego i zmiennoprzecinkowego jest inny;
- reprezentacja maszynowa obu formatów jest odmienna, w szczególności jedna próbka sygnału w reprezentacji zmiennoprzecinkowej zajmuje więcej pamięci;
- operacje arytmetyczne w przypadku formatu stałoprzecinkowego są realizowane przez inne procedury niż w przypadku formatu zmiennoprzecinkowego;
- format zmiennoprzecinkowy jest związany z zaokrągleniem i nie daje gwarancji uzyskania dokładnych wartości całkowitych próbek sygnału po zastosowaniu odwrotnej transformacji falkowej.

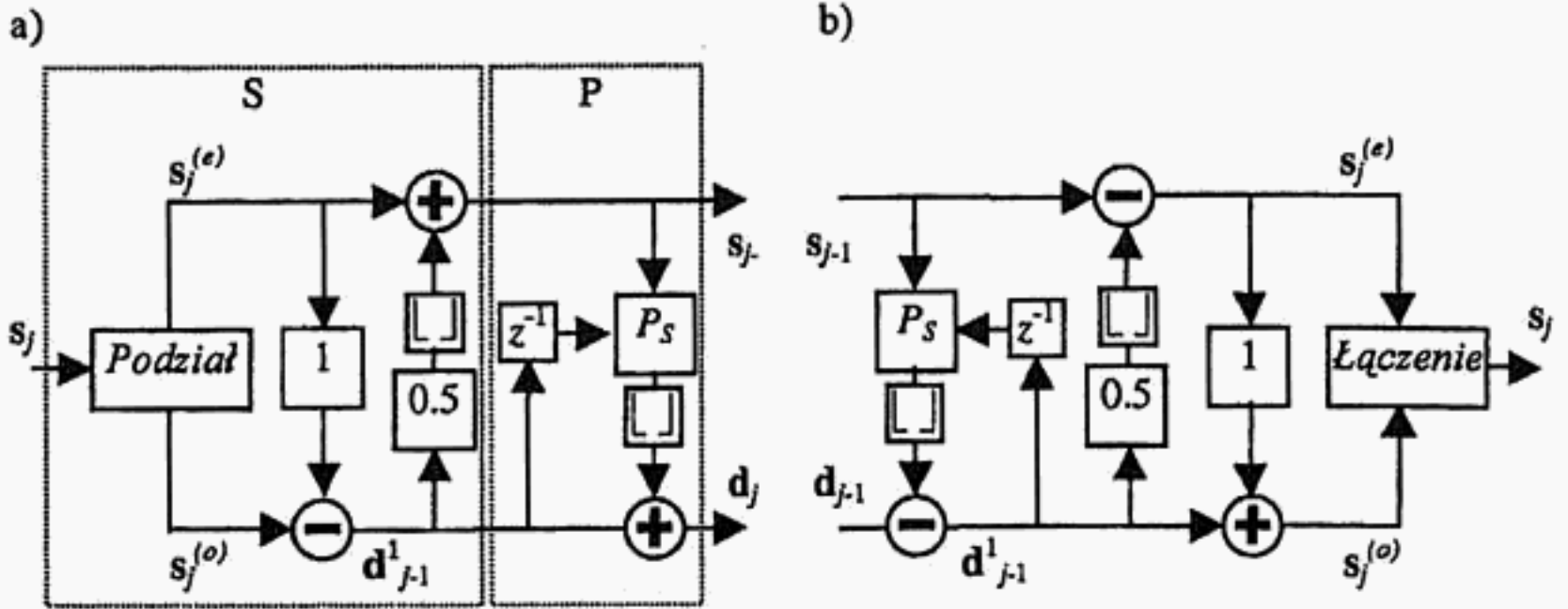
Powyższe mankamenty, a także możliwość realizacji transformacji falkowej z użyciem bardziej rozpowszechnionych i szybszych procesorów stałoprzecinkowych, stały się motywacją do wynalezienia transformacji falkowej używającej współczynników całkowitych.

Przed wynalezieniem predykccyjnej transformacji falkowej umożliwiającej realizację dowolnego systemu falkowego w dziedzinie całkowitoliczbowej rolę najbardziej zaawansowanej metody dekompozycji sygnału używającej wyłącznie reprezentacji stałoprzecinkowej pełniła transformacja S oraz jej udoskonalona wersja znana jako transformacja S+P (rys. 3.5). W wyniku transformacji S ciąg próbek  $d$  zawiera reprezentację wyższej połowy pasma, natomiast ciąg  $s$  jest aproksymacją przy niższym poziomie rozdzielczości.

$$s_{1,l} = (s_{0,2l+1} + s_{0,2l}) / 2 \quad (3.19a)$$

$$d_{1,l} = s_{0,2l+1} - s_{0,2l} \quad (3.19b)$$

Przy zaokrągleniu wartości ciągu  $s$  wykorzystuje się własność liczb całkowitych, dla których, dowolnych dwóch, suma i różnica mają tę samą parzystość (obie parzyste albo obie nieparzyste). Znając parzystość różnicy (a więc próbki ciągu  $d$ ), można więc określić kierunek zaokrąglenia średniej (próbki ciągu  $s$ ).



Rys. 3.5. Schemat blokowy transformacji S+P: a) moduł analizy; b) moduł syntezy

Pewną poprawę miernych własności widmowych ciągów generowanych przez transformację S osiągnięto przez dołączenie operacji predykcyjnej używającej nie tylko wartości poprzedzających próbek aproksymaty  $s$ , ale także wyliczonych wartości detali  $d$ :

$$d_{1,l}^{(1)} = s_{0,2l+1} - s_{0,2l} \quad (3.20a)$$

$$s_{1,l} = s_{0,2l+1} + d_{1,l}^{(1)} / 2 \quad (3.20b)$$

$$d_{1,l} = d_{1,l}^{(1)} + \alpha \cdot (s_{1,l-2} - s_{1,l-1}) + \beta \cdot (s_{1,l-1} - s_{1,l}) + \gamma \cdot (s_{1,l} - s_{1,l+1}) - \delta \cdot d_{1,l}^{(1)} \quad (3.21)$$

Współczynniki wagowe  $\alpha$ ,  $\beta$ ,  $\gamma$  i  $\delta$  przyjmują różne wartości w zależności od zastosowań.

Predykcyjna transformacja falkowa jest w istocie także uogólnieniem transformacji S oraz S+P. Transformacja S jest bowiem przypadkiem, w którym współczynnik wagowy predyktora jest równy jedności, co wynika z założenia, że wartość próbki nieparzystej  $s_{0,2l+1}$  jest równa wartości próbki parzystej  $s_{0,2l}$ . Z wymagania, aby wartość średnia aproksymacji na kolejnych poziomach dekompozycji  $s_j$  oraz  $s_{j-1}$  była taka sama, można wyliczyć wartość wagi uaktualnienia jako  $1/2$ . Tak więc transformacja S jest szczególnym przypadkiem predykcyjnej transformacji falkowej, w której zastosowanie w roli predyktora i uaktualnienia wielomianów rzędu 1 (jednomianów) doprowadziło do dekompozycji sygnału, w której funkcje bazowe, tj. falka i funkcja skalująca, są funkcjami Haara. Jest to jedyny przykład dekompozycji ortogonalnej, ponieważ stosowanie predyktora i uaktualnienia wyższych rzędów odpowiada dekompozycji biortogonalnej.

Twierdzenie, że każda transformacja falkowa może być zapisana jako predykcyjna, jest równoznaczne z możliwością skonstruowania każdej transformacji falkowej



w wersji całkowitoliczbowej. W tym celu wystarczy zaokrąglić rezultat filtracji bezpośrednio przed operacją sumowania. W wersji stałoprzecinkowej operacja liftingu jest więc wyrażona wzorem

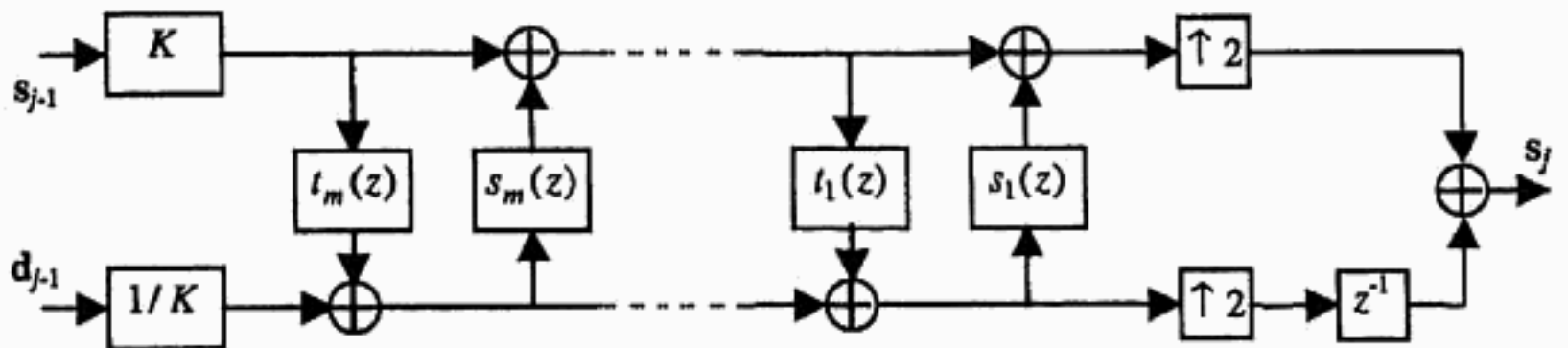
$$d_{1,l}^{(i)} = d_{1,l}^{(i-1)} - \left( \sum_k p_k^{(i)} \cdot s_{1,l-k}^{(i-1)} + \frac{1}{2} \right) \quad (3.22a)$$

natomiast operacja liftingu dualnego wzorem

$$s_{1,l}^{(i)} = s_{1,l}^{(i-1)} - \left( \sum_k u_k^{(i)} \cdot d_{1,l-k}^{(i)} + \frac{1}{2} \right) \quad (3.22b)$$

co oczywiście prowadzi do transformacji ciągu liczb całkowitych w ciąg liczb całkowitych. Jest to transformacja odwracalna, posiadająca własność dokładnej rekonstrukcji sygnału. Zarówno w dziedzinie czasu, jak i w dziedzinie czasowo-częstotliwościowej sygnał jest kodowany w identycznej reprezentacji maszynowej i przetwarzany za pomocą tych samych procedur.

Całkowitoliczbową predycyjną odwrotną transformację falkową otrzymać można przez proste odwrócenie kolejności operacji i odwrócenie znaków sumowania (rys. 3.6).



Rys. 3.6. Schemat blokowy predycyjnej odwrotnej transformacji falkowej w wersji całkowitoliczbowej

Ostatnim zagadnieniem wymagającym wyjaśnienia jest współczynnik skalujący  $K$ .

Dwa rozwiązania są stosowane w praktyce.

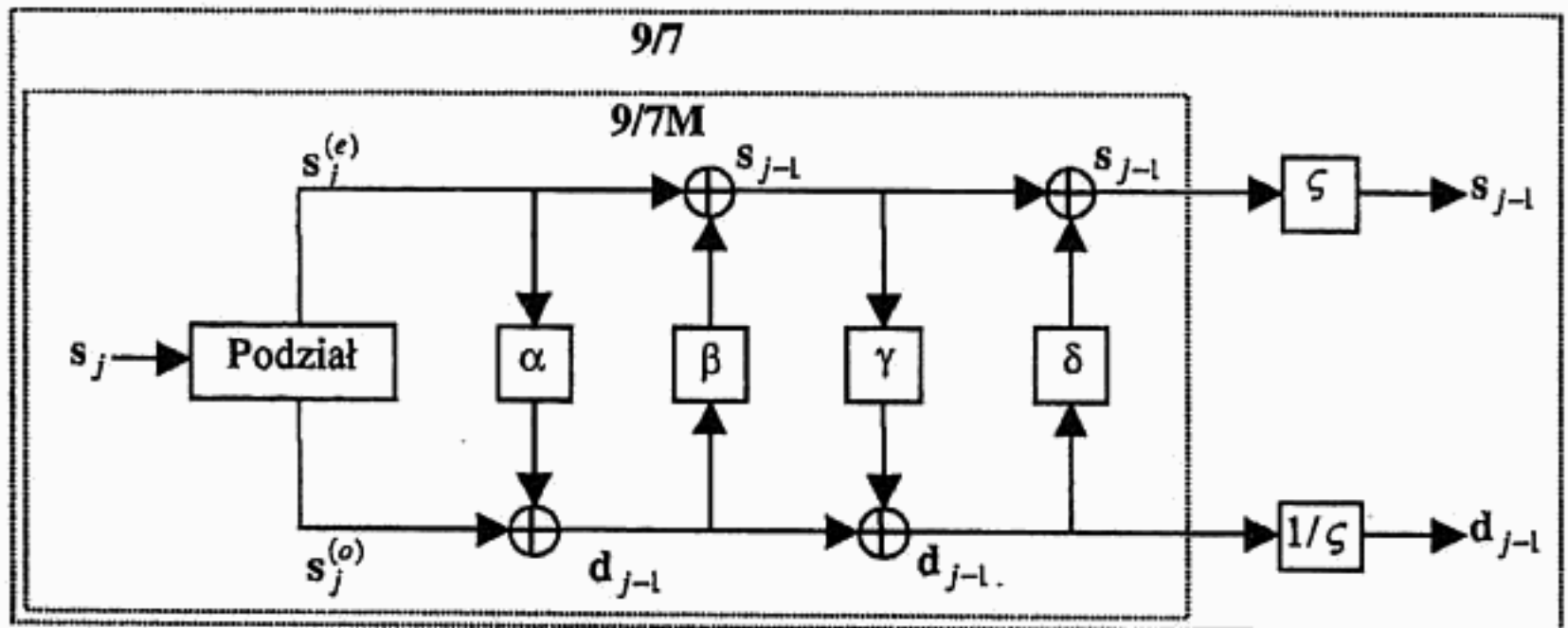
- 1) Zaniedbanie współczynnika  $K$ , co wymaga pamiętania, że wynikowy szereg aproksymacji zawiera wartości  $K$ -krotnie większe, natomiast wartości szeregu detali są  $K$ -krotnie mniejsze. W takim przypadku korzystne jest takie zaprojektowanie filtrów predykcji i uaktualnienia, aby wartość współczynnika  $K$  była jak najbliższa jedności. Można to osiągnąć wykorzystując niejednoznaczność przekształcenia banku filtrów do postaci liftingowej (istnieje więcej niż jeden sposób odwzorowania).

- 2) Zastosowanie trzech dodatkowych operacji liftingowych, możliwych do zaprojektowania w wersji stałoprzecinkowej zamiast skalowania. Wykorzystywana jest przy tym jedna z następujących zależności:

$$\begin{bmatrix} K & 0 \\ 0 & \frac{1}{K} \end{bmatrix} = \begin{bmatrix} 1 & K - K^2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{K} & 1 \end{bmatrix} \begin{bmatrix} 1 & K - 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad (3.23a)$$

$$\begin{bmatrix} K & 0 \\ 0 & \frac{1}{K} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 - \frac{1}{K} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ K & 1 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{K} - \frac{1}{K} \\ 0 & 1 \end{bmatrix} \quad (3.23b)$$

Jednym z pierwszych zastosowań całkowitoliczbowej predykcyjnej transformacji falkowej jest bezstratna kompresja obrazów. Stosuje ona dwie operacje liftingu na przemian z dwoma operacjami liftingu dualnego. Filtr analizy  $\tilde{h}(z)$  ma dziewięć współczynników, a filtr syntezy  $h(z)$  – siedem, dlatego bank filtrów nazwany jest 9/7 (rys. 3.7). Współczynnik skalujący dla tego filtru jest równy 1,1496 i jest zaniedbywany.



$$\alpha = -\frac{203}{128}, \quad \beta = -\frac{217}{4096}, \quad \gamma = \frac{113}{128}, \quad \delta = \frac{1817}{4096}, \quad \zeta \approx 1,149604398$$

Rys. 3.7. Schemat blokowy całkowitoliczbowej transformacji falkowej 9/7 i 9/7M

## ROZDZIAŁ 4

# Ekstrakcja informacji diagnostycznych z reprezentacji czasowo-częstotliwościowej

W rozdziale tym zostaną omówione przykłady pozyskiwania informacji diagnostycznych z czasowo-częstotliwościowych reprezentacji sygnałów biomedycznych. Cechą wspólną omawianych tu zastosowań jest charakter informacji diagnostycznych określonych na przecięciu zwartego przedziału czasowego i widmowego. Informacje takie, trudne lub niemożliwe do pozyskania z czasowej postaci sygnału, stają się łatwe do identyfikacji na płaszczyźnie czasowo-częstotliwościowej.

Najprostsze metody analizy nie wymagają stosowania kolejnych automatycznych metod rozpoznawania, a ponieważ transformacja jest dokonywana w jedną stronę, jej odwracalność, a więc ortogonalność bazy dekompozycji, pozostaje bez znaczenia. Bardziej zaawansowane zastosowania tej grupy algorytmów stosują automatyczne metody rozpoznawania i klasyfikacji cech, a postać czasowo-częstotliwościowa umożliwia łatwiejsze wyodrębnienie komponentów sygnału, których wartości stanowią zmienne wejściowe dla kolejnego algorytmu. Typowym przedstawicielem takiego agregatu metod jest złożenie transformacji falkowej z sieciami neuronowymi.

W innych zastosowaniach produkty dekompozycji traktowane są jak sygnały w dziedzinie czasu, o częstotliwości próbkowania zmienionej w stosunku do sygnału oryginalnego, zawierające interesujący wycinek pasma. Sygnały takie są przekształcane do dziedziny częstotliwości za pomocą transformacji Fouriera albo analizowane w naturalnej dziedzinie z wykorzystaniem odmiennych niż w dziedzinie czasu własności statystycznych.

### 4.1. Detekcja późnych potencjałów komorowych w dziedzinie czasowo-częstotliwościowej

#### 4.1.1. Definicja i znaczenie diagnostyczne mikropotencjałów

Nazwa „późne potencjały komorowe” (mikropotencjały komorowe, ang. *late potentials*) określa zjawiska obserwowane w elektrokardiologii jako patologia depolaryzacji komórek mięśnia roboczego komór. Komórki te charakteryzuje fizjologiczne

zjawisko tzw. przedłużonej refrakcji, dzięki któremu nie odbudowują one swego potencjału spoczynkowego, a zatem pozostają nieczułe na bodźce i nie są zdolne do powtórzenia skurczu przez okres ok. 270 ms. Wobec faktu, że komórki mięśnia roboczego komór serca mogą wzajemnie przekazywać sobie pobudzenie, zjawisko przedłużonej refrakcji ma fundamentalne znaczenie dla synchronizacji ich skurczu w czasie akcji komór. Dzięki niemu dochodzi bowiem do sytuacji, kiedy pobudzenie wygenerowane w układzie bodźcotwórczym obejmuje kolejne komórki w objętości mięśnia tak długo, aż wszystkie one zostaną zdepolaryzowane. Pod koniec fazy skurczu, kiedy brak już komórek gotowych do przyjęcia pobudzenia, wygasa ono do czasu ponownego zainicjowania skurczu przez układ bodźcotwórczy serca. W ten sposób kolejne skurcze są ściśle kontrolowane przez układ bodźcotwórczy (najczęściej przez węzeł zatokowo-przedsionkowy), a ich częstotliwość jest adaptowana do zmieniającego się zapotrzebowania organizmu na tlen.

Sytuacja patologiczna polega na utracie własności przedłużonej refrakcji przez niektóre komórki mięśnia roboczego komór serca. Komórki takie mogą być pobudzane wielokrotnie w czasie tej samej akcji skurczowej, a bodziec elektryczny może dzięki nim znajdować drogę powrotną lub okrężną. Bez wchodzenia w przyczyny zaburzeń przedłużonej refrakcji należy stwierdzić, że jest to degradacja elektrochemicznych właściwości błony komórkowej obejmująca początkowo pojedyncze komórki, lecz z czasem rozprzestrzeniająca się na komórki sąsiednie. W początkowym stadium, gdy komórki dotknięte patologią są jeszcze nieliczne, zmiany w funkcjonowaniu i wydolności serca są trudne do zauważenia. Najbardziej widocznym objawem są zjawiska elektryczne generowane przez włókna kurczące się wielokrotnie, których obraz elektrokardiograficzny uzasadnia nazwę „mikropotencjały”. Następują one bezpośrednio za głównym impulsem skurczowym (zespołem QRS), charakteryzują się niewielką energią (gdyż reprezentują niewielki odsetek komórek) oraz wysoką częstotliwością, co jest następstwem wielokrotnych skurczów. Zjawiska te są trudne do zauważenia w związku z ich lokalnym występowaniem oraz z bezpośrednim sąsiedztwem zespołu QRS o znacznie (do 100 razy) większej amplitudzie.

W stadium zaawansowanym istnienie drogi powrotnej lub okrężnej może powodować nawracanie pobudzeń mające poważne konsekwencje dla pracy serca:

- skurcze partii mięśnia najbardziej odległych na drodze okrężnej następują naprzemiennie, a nie jednocześnie, co upośledza pracę serca jako pompy krwi;
- zawsze znajdują się komórki zdolne przejąć pobudzenie, które nie wygasa, w rezultacie akcja serca wymyka się spod kontroli układu bodźcotwórczego.

Konsekwencją krążącego pobudzenia jest częstoskurcz komorowy, w czasie którego częstość akcji serca może osiągnąć wartość do 600 uderzeń na minutę. Ponieważ w tych warunkach wydajność hemodynamiczna serca jest znikoma, a jednocześnie mięsień wymaga wzmożonego ukrwienia, powstają lokalne ośrodki martwicy mięśnia sercowego, a ich rozszerzanie prowadzi do nieodwracalnego zniszczenia narządu i śmierci.



Z tych przyczyn wielkie znaczenie diagnostyczne mikropotencjałów polega na możliwości wczesnego wykrycia zaburzeń repolaryzacji oraz wskazania pacjentów o podwyższonym ryzyku tachykardii komorowej. Szczególną rolę odgrywa diagnostyka mikropotencjałów u pacjentów po świeżo przeżytym zawale serca.

#### 4.1.2. Detekcja mikropotencjałów w dziedzinie czasu

Pomimo braku standardowych kryteriów wystąpienia późnych potencjałów, powszechnie przyjętą metodą ich wykrywania jest spełnienie następujących warunków (rys. 4.1):

- czas trwania zespołu QRS (QRSd) filtrowanego górnoprzepustowym filtrem dwukierunkowym 40 Hz musi przekroczyć 114 ms;
- wartość napięcia średniokwadratowego końcowych 40 ms zespołu QRS (RMS40) nie przekracza  $20 \mu\text{V}$ ;
- czas, w którym wartość napięcia średniokwadratowego pozostaje poniżej  $40 \mu\text{V}$ , przekracza 38 ms.

Podstawowym problemem technicznym występującym przy analizie mikropotencjałów jest zapewnienie wysokiego stosunku sygnału do szumu oraz znacznej rozdzielczości i wyższej niż standardowa częstotliwości próbkowania. Dwa ostatnie postulaty są spełniane szczególnie przez elektrokardiografię wysokiej rozdzielczości HR-ECG (ang. *high resolution electrocardiography*), ale również standardowe rejestratory o częstotliwości próbkowania rzędu 400÷600 Hz i rozdzielczości 12 bitów dostarczają sygnał o parametrach umożliwiającym prowadzenie analizy późnych potencjałów.

Wysoką wartość stosunku sygnału do szumu można zapewnić przez uśrednianie sygnału kolejnych uderzeń (o tej samej morfologii). Ponieważ sygnał użyteczny  $x(n)$  i szum  $w(n)$  nie są skorelowane, zarejestrowany przebieg  $y(n)$  można przedstawić jako sumę

$$y(n) = x(n) + w(n) \quad (4.1)$$

Jeżeli kolejne uśredniane okna są odpowiednio zsynchronizowane, położenie punktu synchronizacji jest wspólne dla wszystkich sygnałów, więc kolejne sygnały są skorelowane ze sobą. Nie są w żaden sposób skorelowane składowe szumów w kolejnych oknach. Ponieważ uśrednianie dotyczy składowych sygnału i składowych szumu jednocześnie, wartość amplitudy sygnału (dzięki korelacji zapewnianej przez wspólny punkt synchronizacji) zostaje zachowana, natomiast wariancja uśrednionej składowej szumowej zmniejszy się  $N$ -krotnie

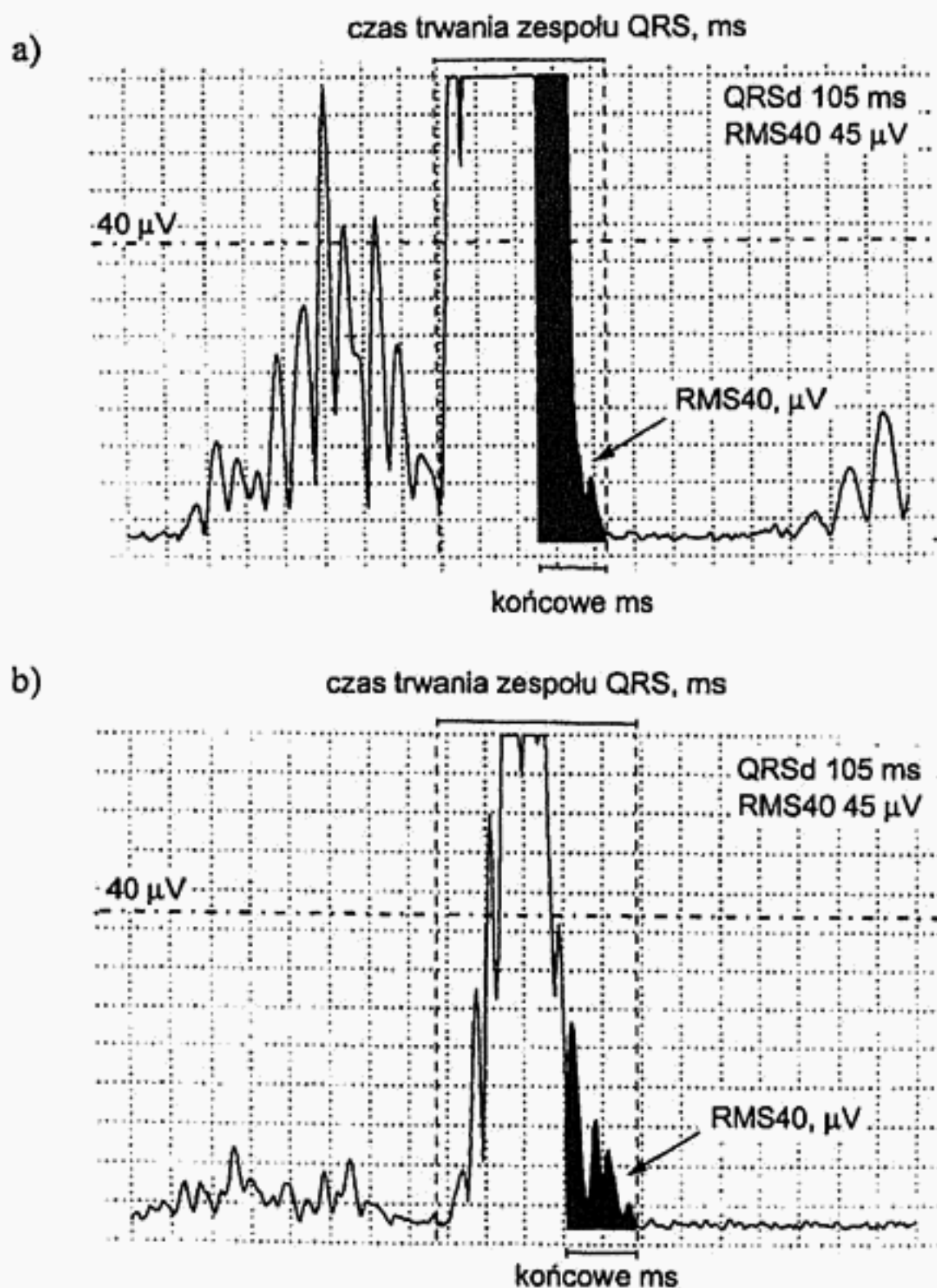
$$\delta_a = \frac{\delta}{N} \quad (4.2)$$

Obliczając poprawę stosunku sygnału do szumu w jednostkach odchylenia średniokwadratowego, otrzymać można

$$\text{SNR} = \text{SNR}_i \cdot \sqrt{N} \quad (4.3)$$



Ponieważ uśrednianie dyskryminuje reprezentację szybkozmiennych zjawisk w zakresie późnych potencjałów, liczba kolejnych uśrednionych zespołów rzadko przekracza 100, co jednak jest już wystarczające, aby 10-krotnie poprawić stosunek sygnału do szumu.



**Rys. 4.1.** Prawidłowy zapis EKG: a) zapis patologiczny;  
b) zawierający późne potencjały komorowe

Innym problemem jest zapewnienie odpowiedniej stabilności punktu centrującego. Wobec braku synchronizacji akcji serca z taktowaniem przetwornika analogowo-cyfrowego może się zdarzyć, że nawet identyczne zespoły QRS będą reprezentowane przez zdecydowanie różne ciągi próbek. Właściwe zsynchronizowanie uśrednianych zespołów można wówczas zapewnić poprzez:

- aproksymację szczytu fali R parabolą i przyjęcie maksimum paraboli jako punktu synchronizującego;

- minimalizację funkcji korelacji wzajemnej zespołu i sumy częściowej poprzedzającą sumowanie;
- ekstrakcję z użyciem dodatkowych metod rozpoznawania (sieci neuronowe, spektra wyższych rzędów, statystyka itp.).

#### **4.1.3. Detekcja mikropotencjałów w dziedzinie czasowo-częstotliwościowej**

Definicja mikropotencjałów sugeruje użycie analizy sygnału w dziedzinie czasowo-częstotliwościowej i pozwala oczekiwać wysokiej efektywności takiego rozpoznania:

- tylko komponenty występujące w określonym przedziale czasowym mogą być uważane za mikropotencjały, ze względów fizjologicznych występują one w bezpośrednim następstwie zespołu QRS (do 100 ms od wierzchołka);
- tylko komponenty o określonych własnościach widmowych mogą być uważane za mikropotencjały, zawierają one znaczną energię w przedziale 70÷250 Hz.

Dekompozycja czasowo-częstotliwościowa sygnału elektrokardiograficznego w celu detekcji mikropotencjałów jest wykonana za pomocą analizy falkowej Morleta. Metoda ta charakteryzuje się lepszą precyzją czasowej lokalizacji komponentów wysoko-częstotliwościowych niż spektrogram uzyskany za pomocą krótkookresowego przekształcenia Fouriera. Jednocześnie wobec sąsiedztwa wysokoenergetycznego zespołu QRS stosowanie transformacji Wignera–Ville’a do detekcji mikropotencjałów jest niewłaściwe, gdyż związane z tą metodą interferencje pasożytnicze znacznie przekraczają energię mikropotencjałów, skutecznie utrudniając ich detekcję.

W detekcji mikropotencjałów istotne znaczenie ma możliwość dowolnego wyboru współczynnika kontrakcji kolejnych wersji częstotliwościowych falek analizujących, dzięki czemu częstotliwości środkowe odpowiadających im filtrów mogą być ustawiane w sposób optymalny dla efektywności detekcji. Brak ortogonalności tak powstałej bazy dekompozycji, a zatem niemożliwość dokładnej rekonstrukcji sygnału, jest bez znaczenia i nie stanowi wady w kontekście zadania detekcji. Użyta falka Morleta jest ponadto prosta obliczeniowo i łatwa w implementacji. Również zespolony charakter tej falki nie stanowi wady, gdyż estymatorem energetycznego udziału interesujących komponentów w sygnale jest moduł wartości odpowiedniego współczynnika powierzchni czasowo-częstotliwościowej.

Warunki do wystąpienia mikroskurczów są analogiczne przy każdej ewolucji serca, dlatego uśrednianie sygnału kolejnych zespołów QRS przynosi poprawę odstępu sygnału mikropotencjałów od szumu wraz z liczbą uśrednionych zespołów, o ile tylko są one poprawnie zsynchronizowane. Ponieważ pomiędzy kolejnymi ewolucjami serca mogą występować różnice fazowe sygnału mikropotencjałów, uśrednianiu podlega moduł sygnału EKG. Uśrednianie sygnału może zostać zrealizowane także za pomocą wartości średnich odpowiadających sobie współczynników czasowo-częstotliwościowych kolejnych ewolucji serca.

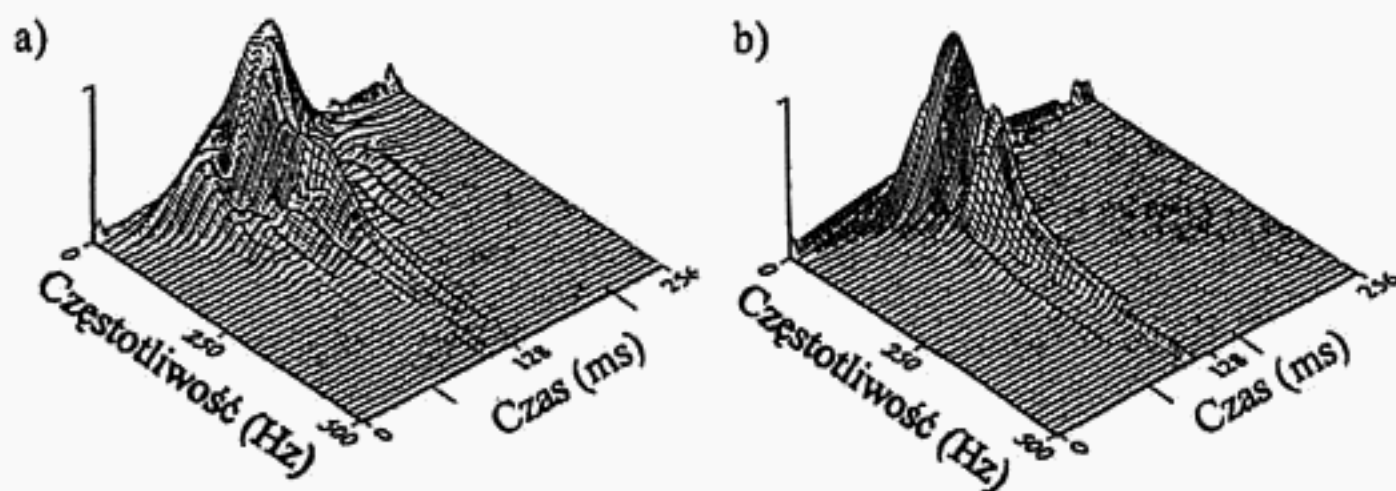
Analiza elektrokardiogramów wysokiej rozdzielczości pacjentów, u których stwierdzono wystąpienie mikropotencjałów z użyciem metody czasowej (podrozdz. 4.1.2), i porównanie z zapisami osób zdrowych doprowadziły do wskazania współczynników powierzchni czasowo-częstotliwościowej o największej zdolności separacji obu grup. Pozycja tych współczynników w dziedzinie częstotliwości jest określona przez numer pasma, natomiast pozycja w dziedzinie czasu jest określona względem punktów początku i końca zespołu QRS.

Branie pod uwagę punktu początkowego zespołu QRS jest uzasadnione, gdyż:

- reprezentuje on początek akcji skurczowej będącej bezpośrednią przyczyną mikropotencjałów;
- może być wyznaczony z większą dokładnością niż punkt końcowy zespołu QRS, zwłaszcza gdy detekcja punktu końcowego jest zakłócona przez mikropotencjały;
- może być wyznaczony z większym poziomem ufności niż punkt końcowy zespołu QRS.

Podobnie jak w przypadku detekcji w dziedzinie czasu, dla wybranych współczynników określono następnie wartości progowe, których przekroczenie jest równoznaczne z wykryciem mikropotencjałów. Ponieważ współczynników tych jest zwykle kilka, ostateczną decyzję przyjmuje się w wyniku ważenia decyzji cząstkowych.

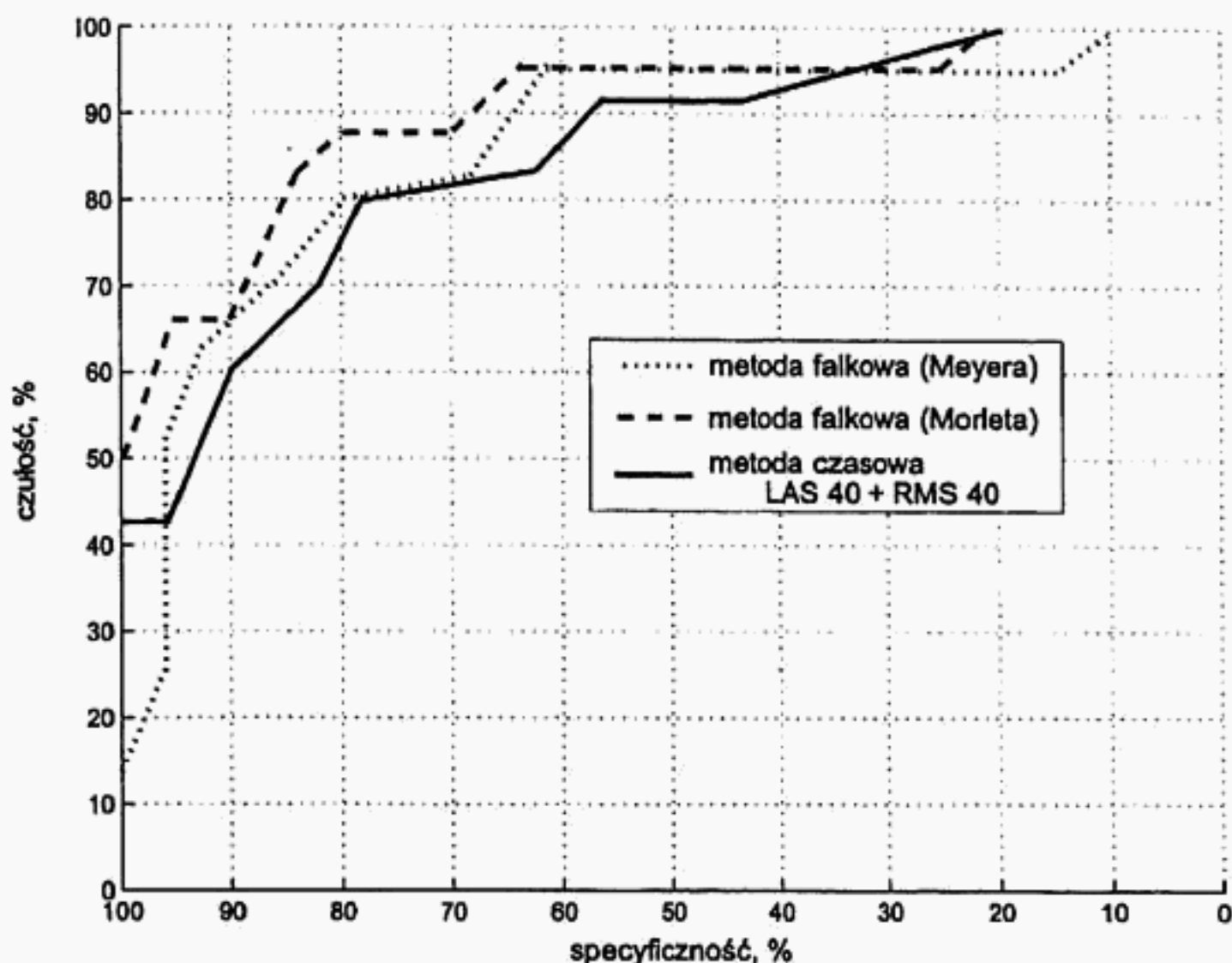
Alternatywnym rozwiązaniem jest zastosowanie logiki rozmytej albo systemu klasyfikacji z użyciem sieci neuronowych, którego wejściem jest zestaw wybranych współczynników czasowo-częstotliwościowej (uśrednionej) reprezentacji sygnału. W tych przypadkach odpowiedź systemu nie jest ograniczona do wartości binarnej (wykryto, nie wykryto), ale może zawierać wskazówki odnośnie prawdopodobieństwa lub skłonności do późnych potencjałów albo wskazania dotyczące dalszych testów.



**Rys. 4.2.** Przykładowe powierzchnie czasowo-częstotliwościowe uzyskane z elektrokardiogramu za pomocą analizy falkowej Morleta: a) dla przypadku występowania mikropotencjałów komorowych; b) dla przypadku fizjologicznego

Rysunek 4.2 przedstawia przykładowe powierzchnie czasowo-częstotliwościowe uzyskane z elektrokardiogramu za pomocą analizy falkowej Morleta dla przypadku występowania mikropotencjałów komorowych oraz dla przypadku fizjologicznego.

Rysunek 4.3 przedstawia porównanie skuteczności detekcji mikropotencjałów metodą tradycyjną w dziedzinie czasu oraz metodą czasowo-częstotliwościową z użyciem transformacji falkowej.



Rys. 4.3. Porównanie skuteczności detekcji mikropotencjałów metodą tradycyjną w dziedzinie czasu oraz metodą czasowo-częstotliwościową z użyciem transformacji falkowej

## 4.2. Detekcja zjawisk padaczkowych

### 4.2.1. Zastosowanie elektroencefalografii w detekcji padaczki

Elektroencefalografia (EEG) jest techniką rejestracji i analizy czynności elektrycznej mózgu. Niewielka amplituda sygnału, w najlepszym przypadku sięgająca  $100 \mu\text{V}$ , oraz konieczność specjalnego przygotowania skóry głowy sprawiają, że wykonanie zapisu EEG nie należy do zadań technicznie łatwych. Obecnie zapis elektroencefalograficzny jest drugim, po elektrokardiografii, powszechnie wykonywanym testem elektrodiagnostycznym. Również pod względem liczby metod automatycznego przetwarzania sygnału i wspomagania interpretacji badań komputerem ustępuje tylko elektrokardiografii.

Zapis elektroencefalograficzny jest standardowo wykonywany z użyciem 16-kanalowego rejestratora zapewniającego czułość ok.  $50 \mu\text{V}$  na zakres pomiarowy w pa-



śmie częstotliwości 1÷100 Hz. W powszechnym użyciu przy badaniach osób dorosłych jest tzw. System Międzynarodowy (System 10-20), zaproponowany w 1958 roku, składający się z 19 elektrod podstawowych.

Badanie elektroencefalograficzne wykonuje się zwykle jako spoczynkowe, a następnie aktywacyjne, co ma na celu wyodrębnienie i zlokalizowanie zmian elektrycznej czynności mózgu wywołanych określonym bodźcem.

Zapis spoczynkowy jest wykonywany w warunkach spokoju, bez dopływu bodźców słuchowych, wzrokowych i dotykowych o nadmiernym natężeniu. Osoba badana siedzi lub leży w wygodnej pozycji, w stanie relaksu – odprężenia fizycznego i psychicznego. W takich warunkach uwidaczniają się tzw. cechy podstawowe zapisu EEG, które przyjęto uważać za punkt odniesienia w badania stanów patologicznych oraz dla potrzeb wyodrębnienia składowych związanych z bodźcem aktywacyjnym.

Zapis aktywacyjny polega na zastosowaniu bodźca w celu wywołania zmian reprezentacji elektrycznej aktywności mózgu, a następnie diagnozowaniu poprawności pracy mózgu w zależności od zaobserwowanych zmian zapisu. Do rutynowo stosowanych aktywacji należą: zamknięcie oczu, hiperwentylacja, sen fizjologiczny lub farmakologiczny oraz aktywacje chemiczne.

Zapis EEG zarejestrowany w pamięci jako wielowymiarowy sygnał dyskretny może być przeszukany automatycznie w celu rozpoznania tzw. rytmów będących podstawowymi elementami zapisu EEG. Użycie komputera ma szczególne znaczenie w przypadku zapisów dłuższych (np. 8-godzinnych).

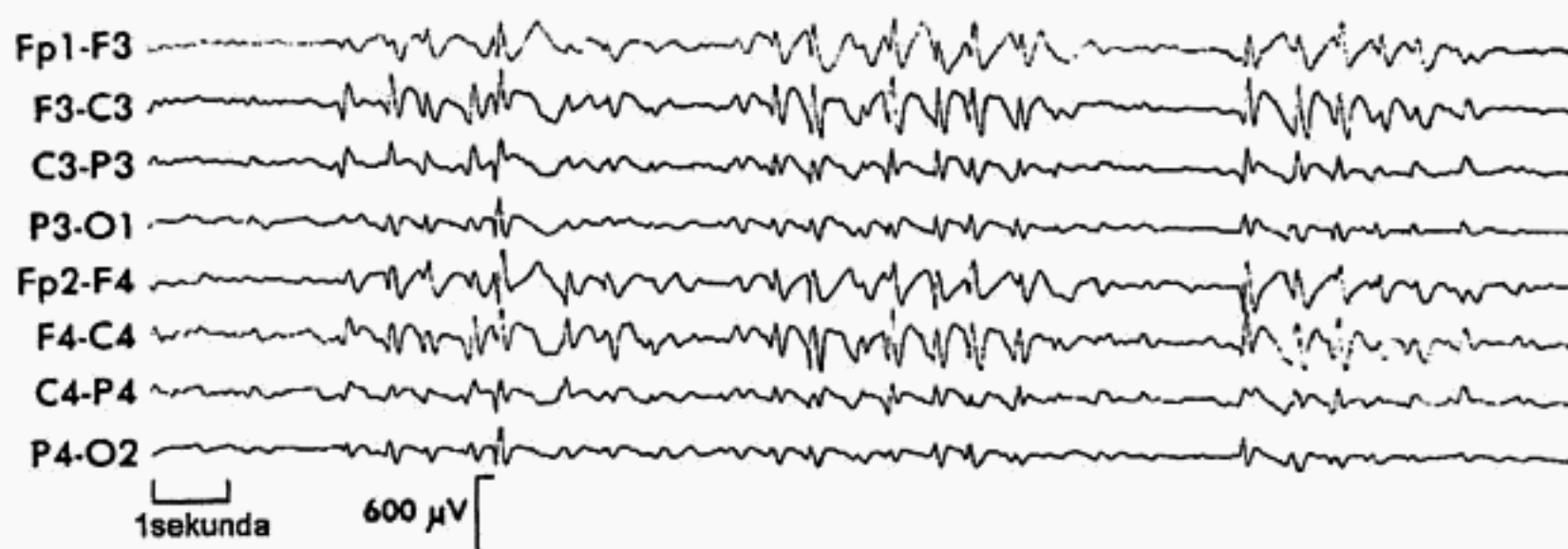
Identyfikacja rytmów jest zwykle związana z analizą widmową lub użyciem filtrów cyfrowych. Identyfikacja prowadzi do wskazania początków i końców odcinków zawierających dany rytm. Odcinki takie mogą zachodzić na siebie. Analiza widmowa (zwykle krótkookresowa transformacja Fouriera) jest prowadzona w ruchomym oknie o długości 5 s. Taka jest też dokładność określania początku i końca odcinka zawierającego poszukiwany rytm. Transformata dzielona jest na podpasma odpowiadające poszczególnym rytmom i obliczana jest zawartość energetyczna poszczególnych zakresów częstotliwości. Na podstawie proporcji energii chwilowych w zdefiniowanych zakresach częstotliwości stawiana jest diagnoza o obecności lub nieobecności rytmu.

Obserwacja podstawowych elementów zapisu elektroencefalograficznego może prowadzić do bardziej specjalistycznych poszukiwań diagnostycznych w kierunku wykrycia (weryfikacji podejrzeń) określonych zaburzeń funkcjonowania centralnego układu nerwowego. Jest to jedna z dziedzin diagnostyki, w której elektroencefalografia nie jest wypierana przez coraz doskonalsze metody diagnostyki obrazowej (CT czy NMR).

Przykładem takiego zaburzenia jest epilepsja (padaczka). W tym przypadku analiza EEG charakteryzuje się nie tylko wartością diagnostyczną czy prognostyczną, ale umożliwia także lokalizację ognisk padaczkowych, jeśli takie wystąpią.



Jeżeli podczas obserwacji chorego stwierdzono napady lub wyłączenia, to występowanie jednej z dwóch form (rys. 4.4): iglicy albo połączenia iglicy i fali, w określonych długotrwałych sekwencjach jest uważane za potwierdzenie padaczki w zapisie EEG. Istotna jest także topograficzna synchronizacja sekwencji iglic. Jednocześnie iglice oraz fale mogą występować pojedynczo u osób całkowicie zdrowych na pograniczu faz snu.



Rys. 4.4. Elektroencefalograficzny zapis iglic będących objawem padaczki

Iglice reprezentują wyzwolenie napadu padaczkowego, co następuje zwykle wtedy, gdy wyładowania w obrębie pojedynczego neuronu rozprzestrzenia się i obejmą dostatecznie dużą liczbę neuronów sąsiednich. Elektroencefalograficzny zapis napadu padaczki jest cennym źródłem informacji diagnostycznych, ale ponieważ moment wystąpienia napadu jest nieprzewidywalny, celowe jest stosowanie aktywacji sprzyjających nasileniu objawów. Zwykle dokonuje się tego podczas zapisu snu fizjologicznego albo na drodze farmakologicznej.

Reguły rozróżnienia iglic od artefaktów są stosunkowo łatwe do sformalizowania:

- iglice są niesymetryczne i zwykle następuje po nich wolna fala;
- iglice są zwykle dwu- lub tryfazowe na tle wolnozmiennnej aktywności podłoża;
- iglice występują jednocześnie w kilku odprowadzeniach, przy czym w jednym z nich amplituda jest najwyższa;
- iglice występują rytmicznie, niekiedy interwał podstawowy jest podwielokrotnością odstępu dwóch sąsiednich iglic.

Identyfikacja iglic może być zrealizowana:

- metodami korelacyjnymi, w których poprawne rozpoznanie opiera się na podobieństwie kształtów kolejnych iglic w danym kanale;
- w dziedzinie czasowo-częstotliwościowej poprzez stwierdzenie współwystąpień określonych komponentów czasowo-częstotliwościowych sygnału.

Dokładniejsza analiza kształtu iglicy daje pewne wskazówki diagnostyczne dotyczące przyczyn padaczki, co jednak wymaga doświadczenia klinicznego i nie stało się dotychczas przedmiotem rozpoznania automatycznego.

#### 4.2.2. Detekcja iglic w dziedzinie czasowo-częstotliwościowej

Definicja iglicy jako krótkotrwałego epizodu wysokoczęstotliwościowego synchronicznego w kilku kanałach i powtarzającego się w określonych odstępach czasu predestynuje wykorzystanie do ich detekcji metod czasowo-częstotliwościowych. W najprostszym przypadku detekcja polega na stwierdzeniu wystąpienia komponentów sygnału w określonym paśmie częstotliwościowym (dla iglicy: 24÷40 Hz) o energii znacznie przewyższających energię otoczenia. Stosunek tych energii jest miarą prawdopodobieństwa wystąpienia iglicy i przeciwnie do wartości bezwzględnej energii umożliwia wyłowienie iglicy z tła zakłóceń lub sygnałów fizjologicznych zarejestrowanych równocześnie. Kolejnym krokiem jest badanie korelacji podwyższonej energii w sąsiednich topologicznie kanałach zapisu EEG. Jednoczesne podwyższenie energii w kilku kanałach jest już przesłanką do stwierdzenia wystąpienia iglicy.

Metoda detekcji iglic jako izolowanych epizodów nie jest odporna na zakłócenia impulsowe pojawiające się we wszystkich odprowadzeniach EEG. Źródłem takich zakłóceń może być np. impuls elektromagnetyczny lub niestabilność napięcia zasilania. Uodpornienie metody detekcji na zakłócenia tego typu jest możliwe dzięki grupowej naturze iglic. Wykrycie pierwszej iglicy jest realizowane przez algorytm opisany poprzednio. Następnie jest określany interwał czasu, w którym spodziewane jest wystąpienie kolejnej iglicy (zwykle do 1,5 s). Jeżeli w tym interwale nie wykryto kolejnego epizodu synchronicznego podwyższenia energii komponentów wysokoczęstotliwościowych – informacja o pozytywnej detekcji jest anulowana. W przeciwnym przypadku zarówno pierwszy, jak i drugi epizod jest uznawany za iglicę, a następnie modyfikowane są warunki progowe detekcji:

- próg stosunku energii iglicy do energii otoczenia w zadanym paśmie częstotliwości jest ustawiany na 40% wartości średniej obu wykrytych iglic;
- interwał czasu, po którym spodziewane jest wystąpienie kolejnej iglicy, jest równy odstępowi iglic już znalezionych – wyznaczony przedział czasu jest dłuższy dla dłuższych interwałów i zwykle zawiera się w granicach 150÷200 ms.

Wykrycie kolejnych iglic przebiega następująco:

- każde synchroniczne podwyższenie energii komponentów wysokoczęstotliwościowych jest notowane;
- jeżeli bieżąca wartość stosunku energii jest większa od ostatniej wartości progowej i epizod przypada w przedziale czasu spodziewanej iglicy, to jest uznawany za iglicę;

- jeżeli w przedziale czasu spodziewanej iglicy nie wykryto epizodu podwyższenia energii, a epizod taki przypada po przedziale czasu spodziewanej iglicy i wartość jego stosunku energii jest większa niż podwojona wartość ostatniej średniej, to jest on uznawany za iglicę;
- w przypadku pozytywnej detekcji iglicy wartości interwału oraz średniego i progowego stosunku energii są modyfikowane; jeżeli w skład grupy wchodzi więcej niż sześć iglic to tylko sześć ostatnich wartości stosunku energii podlega uśrednianiu.

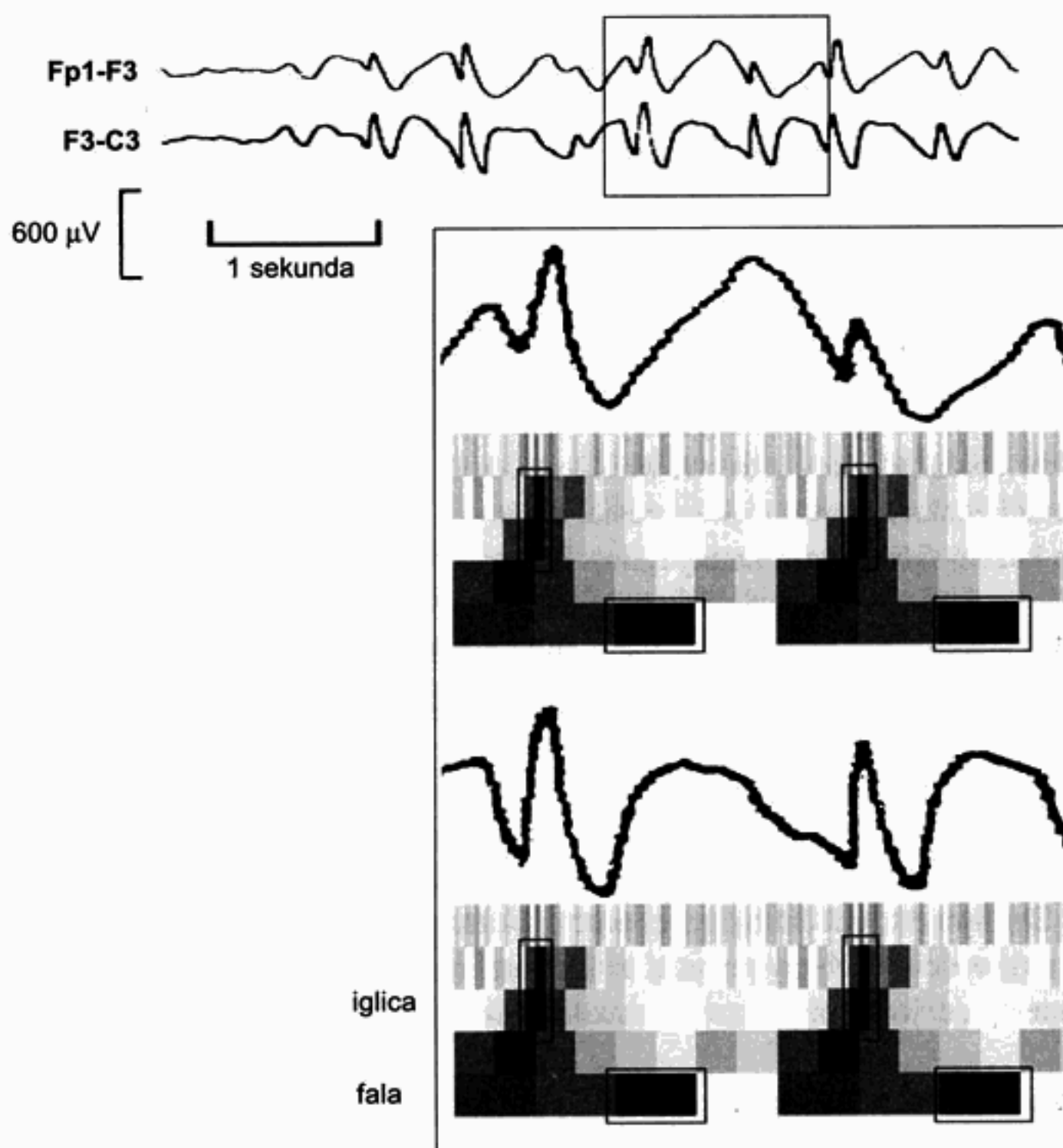
Odrębna detekcja fali jako epizodu EEG nie może być wykonana, gdyż zakres częstotliwościowy tworzących ją komponentów sygnału jest identyczny z rytmem alfa ( $8\div 13$  Hz). Fala może być za to rozpoznana w kontekście poprzedzającej iglicy. Podstawą rozpoznania jest w takim przypadku podwyższona względem otoczenia energia w paśmie częstotliwości  $8\div 13$  Hz oraz stałość odstępu czasowego od poprzedzającej iglicy. W przypadku iglicy z falą, fala jest spodziewana w okresie do 500 ms po iglicy. Pomimo że epizody iglicy z falą również występują synchronicznie w zbliżonych topologicznie kanałach i grupowo w czasie, może się okazać niemożliwe wyodrębnienie w zapisie każdej fali towarzyszącej iglicy. Dlatego dla odróżnienia grupy iglic z falą od grupy iglic wystarczy wykrycie fali za iglicą dla 40% iglic tworzących grupę.

Charakterystyczną cechą opisywanej metody jest zdefiniowanie wzorca czasowo-częstotliwościowego reprezentującego poszukiwane zjawisko, a następnie poszukiwanie lokalnej zgodności rzeczywistego zapisu z tym wzorcem. Opisany algorytm detekcji posługuje się wartościami bezwzględnych współczynników czasowo-częstotliwościowej reprezentacji sygnału. Podobnie jak w przypadku detekcji późnych potencjałów komorowych, dekompozycja jest dokonywana za pomocą analizy falkowej Morleta. Eliminuje to niedokładności precyzji czasowej lokalizacji komponentów wysokoczęstotliwościowych typowe dla krótkookresowego przekształcenia Fouriera, a także interferencje pasożytnicze niemożliwe do uniknięcia w transformacji Wignera–Ville’a.

W detekcji iglic i fal istotne znaczenie ma możliwość dowolnego wyboru współczynnika kontrakcji kolejnych wersji częstotliwościowych falek analizujących, dzięki czemu częstotliwości środkowe odpowiadających im filtrów mogą być ustawiane w sposób optymalny dla efektywności detekcji. W detekcji zjawisk padaczkowych z zapisu EEG nie wykorzystuje się transformacji odwrotnej, dlatego brak ortogonalności bazy dekompozycji jest bez znaczenia i nie stanowi wady w kontekście zadania detekcji. Ponieważ algorytm detekcji wykorzystuje wartości bezwzględne współczynników powierzchni czasowo-częstotliwościowej reprezentujące lokalną energię sygnału bez znaczenia pozostaje również fakt wykorzystywania falki, a zatem i reprezentacji o wartościach zespolonych.

Algorytm detekcji wykorzystuje cztery falki o częstotliwościach środkowych 35 Hz i 27 Hz dla detekcji iglicy oraz 11 Hz i 9 Hz dla detekcji fali (rys. 4.5). Precyzyjne

ustalenie współczynnika kontrakcji zależy oczywiście od zastosowanej częstotliwości próbkowania, ale za cenę rezygnacji z ortogonalności bazy dekompozycji i dopuszczenia powielenia informacji na przecięciu częstotliwościowych zakresów analizy falek poszczególnych pasm uzyskano pełną dowolność wartości tego współczynnika.



**Rys. 4.5.** Detekcja iglic w wieloodprowadzeniowym zapisie EEG w dziedzinie czasowo-częstotliwościowej

W pierwszej kolejności poszukiwane jest wystąpienie iglicy. Zakresy częstotliwościowe dwóch wyższych pasm są przeszukiwane przez trzy przyległe ruchome okna o jednakowej długości (120 ms), w których uśredniana jest wartość bezwzględna współczynników czasowo-częstotliwościowej reprezentacji sygnału. Długość okna jest tak dobrana, że przy dokładności przesuwania okien 30 ms wynikającej z częstotliwości środkowej falek reprezentacja iglicy zmieści się całkowicie w środkowym oknie, podczas gdy dwa sąsiednie okna zawierać będą reprezentację energii otoczenia. Na



kolejnych pozycjach okien badane jest spełnienie warunku maksimum energii w środkowym oknie, a następnie spełnienie przez stosunek energii w środkowym oknie do energii w sąsiednich oknach warunku progowego. Koniunkcja pozytywnych rezultatów progowania w topologicznie przyległych odprowadzeniach jest równoznaczna z wykryciem iglicy, a jej dokładna pozycja wyznaczana jest za pomocą paraboli najlepiej dopasowanej do wartości bezwzględnych współczynników reprezentacji czasowo-częstotliwościowej zawartych w środkowym oknie.

Po stwierdzeniu wystąpienia iglicy i określeniu jej pozycji sprawdzany jest warunek wystąpienia fali w zadanym przedziale czasowym. Dokonuje się tego przez analizę dwóch niższych pasm częstotliwości, począwszy od momentu wystąpienia iglicy. Wartości reprezentacji sygnału nie są w tym przypadku uśredniane, a wystąpienie maksimum lokalnego wśród próbek następujących po iglicy w okresie odpowiadającym 500 ms jest równoznaczne z detekcją fali.

### 4.3. Detekcja wrzecion snu w zapisie polisomnograficznym

Rozdział ten opisuje użycie techniki poszukiwania dopasowań (ang. *matching pursuit*) do analizy elektroencefalogramu zarejestrowanego podczas fizjologicznego snu całonocnego.

Podczas snu fizjologicznego wyodrębnić można powtarzającą się sekwencję faz: charakteryzującą się szybkimi ruchami gałek ocznych fazę REM (ang. *rapid eye movements*) oraz cztery fazy non-REM, rozróżniane na podstawie zapisu aktywności elektrycznej mózgu. Przedmiotem opisywanej tu analizy są tzw. wrzeciona snu, znane też pod nazwą fali sigma, znamienne dla drugiej fazy snu. Wrzeciona są obecne w zapisie EEG snu fizjologicznego osoby zdrowej (ok. 500÷3300 w czasie 8-godzinного snu). Ich brak jest charakterystyczny dla fazy REM, ok. 1% ogólnej liczby wrzecion przypada jednak na tę fazę snu za sprawą rozmycia granicy pomiędzy kolejnymi fazami.

Po zapadnięciu w sen dośrodkowe przewodzenie bodźców w kierunku od kory mózgowej do wzgórza zostaje zablokowane. Wrzeciona snu są generowane przez neurony siateczkowe wzgórza rytmicznie wytwarzające potencjały postsynaptyczne blokujące aktywność neuronów korowo-wzgórzowych. Wrzeciona obserwowane w zapisie odprowadzeń frontalnych mają zwykle niższą częstotliwość (ok. 12 Hz) niż wrzeciona zarejestrowane w okolicach potylicznych (ok. 14 Hz). Wrzeciona zarejestrowane w pobliżu linii środkowej (odprowadzenia: Fz, Cz, Pz) mają przeważnie wyższą amplitudę niż wrzeciona zarejestrowane w rejonach peryferyjnych. Zmiany amplitudy i częstości występowania wrzecion podczas upływu czasu snu są jeszcze obecnie przedmiotem badań neurologii.

Nawet ten uproszczony opis fizjologiczny daje obraz przydatności detekcji i analizy występowania wrzecion w zaburzeniach inicjacji i podtrzymania snu (insomnia)



oraz patologii snu polegających na częściowych lub całkowitych wybudzeniach nocnych. Zmiany morfologii wrzecion następują w wyniku procesów starzenia się oraz dysfunkcji w obrębie centralnego układu nerwowego. Także zmiany nowotworowe (guzy), epilepsja, depresja, wpływ leków oraz zaburzenia napięcia (hypotonia) mają swoją reprezentację w zmianach parametrów wrzecion snu. Ponadto, analiza wrzecion snu może być istotnym elementem monitorowania pacjentów po operacjach neurochirurgicznych.

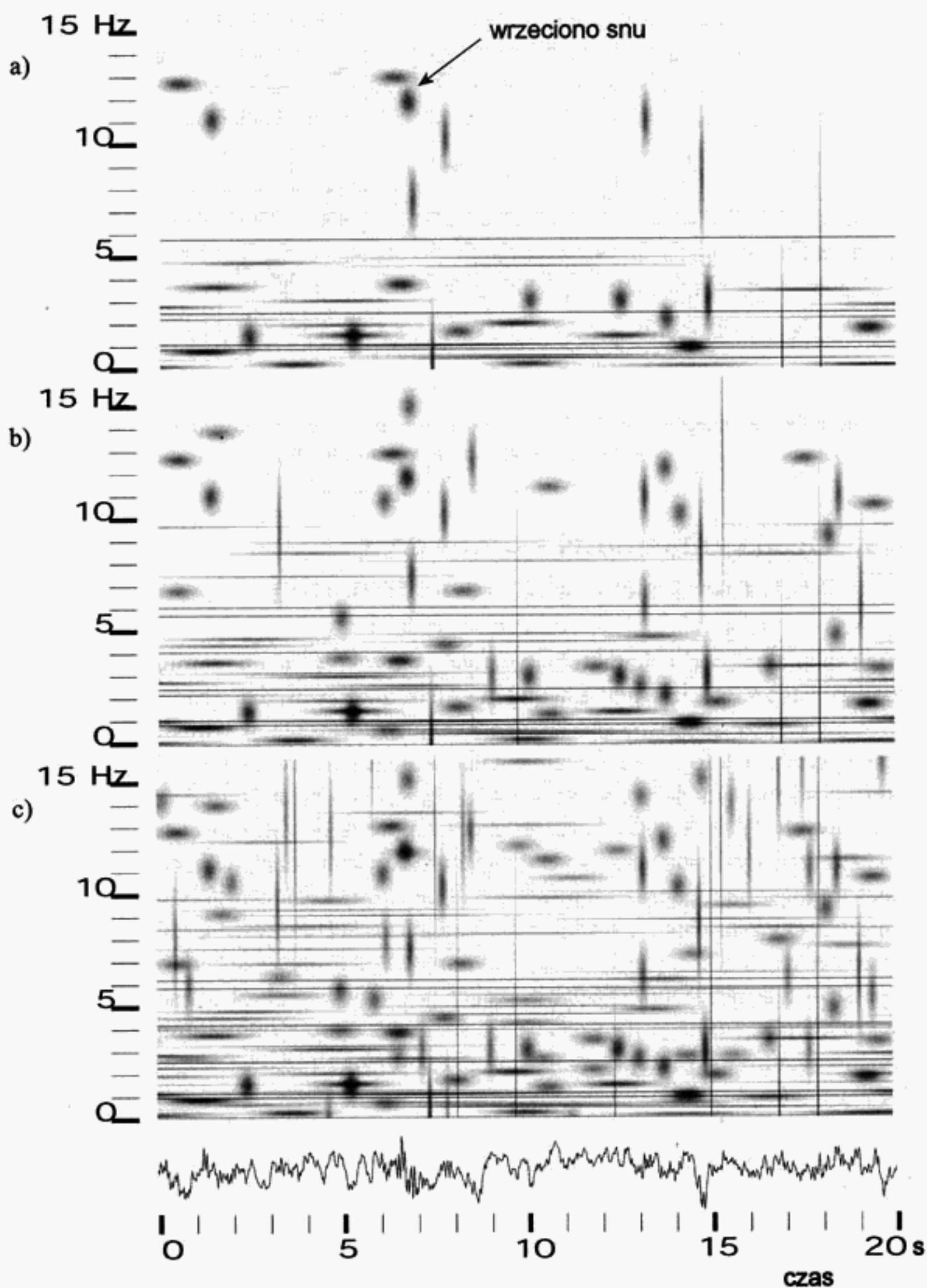
Wrzeciona są strukturami synchronicznych oscylacji o częstotliwości  $11\div 15$  Hz i czasie trwania  $0,5\div 6$  s (co najmniej  $6\div 7$  oscylacji). Są one dobrze zlokalizowane w czasie i częstotliwości oraz niekiedy niosą znaczną energię. Charakterystyczną ich cechą jest właśnie dobra lokalizacja, niekiedy do ich opisu wystarcza jeden atom reprezentacji oraz powtarzalne występowanie we względnie stałych interwałach czasu, co  $3\div 10$  s. Nazwa „wrzeciono” dobrze reprezentuje obwiednię oscylacji podobną do funkcji dzwonowej symetryczną względem połowy czasu trwania i względem linii izoelektrycznej.

Diagnostycznie istotnymi parametrami wrzecion analizowanymi z uwzględnieniem topologii odprowadzeń rejestrujących są:

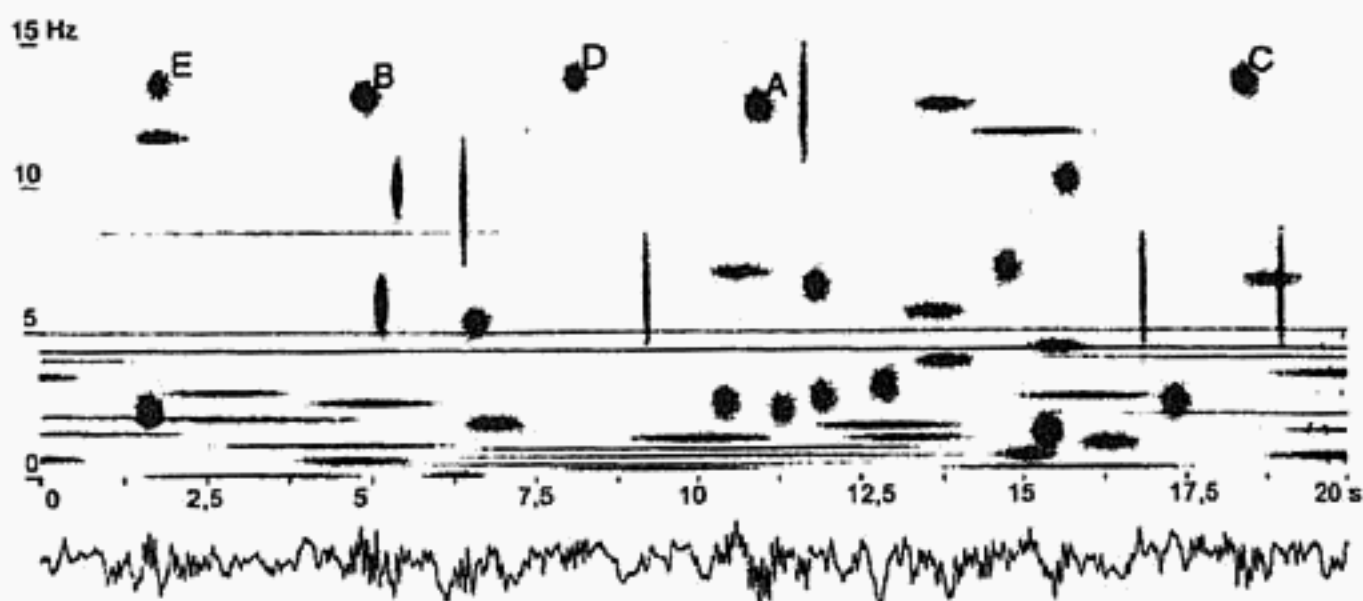
- częstotliwość oscylacji,
- czas trwania,
- amplituda międzyszczytowa,
- liczba wrzecion w jednostce czasu.

Detekcja wszystkich rytmicznych komponentów na płaszczyźnie czasowo-częstotliwościowej prowadziłaby do trudnego do interpretacji nadmiaru danych. Wykorzystanie dodatkowego kryterium dużej energii zawartej w reprezentacji wrzecion prowadzi do znacznie bardziej przejrzystego obrazu elektrycznej aktywności mózgu. Bez znaczenia pozostaje przy tym, że reprezentacja z użyciem ograniczonej liczby funkcji słownika nie wyczerpuje dokładnie całej energii sygnału (ok. 97%). Liczba niezbędnych funkcji analizujących zależy od ilości informacji zawartej w sygnale i od koherencji z wybranym słownikiem. Uzupełnianie reprezentacji o kolejne funkcje prowadzi w tej sytuacji tylko do wzrostu udziału reprezentacji składowych szumu (rys. 4.6). Ostatecznie, dla sygnału o długości 2048 próbek (częstotliwość próbkowania 100 Hz), ze słownika złożonego z miliona elementów o znacznej redundancji wybrano reprezentację złożoną ze 100 funkcji.

Wrzeciona snu są identyfikowane na podstawie dwóch kryteriów: częstotliwości z przedziału  $12\div 15$  Hz oraz rozpiętości czasowej  $500\div 650$  ms (rys. 4.7). Amplitudowy próg detekcji wrzecion na płaszczyźnie czasowo-częstotliwościowej jest zależny od fazy snu, przykładowo w fazie REM brak wrzecion o amplitudzie przekraczającej  $20\ \mu\text{V}$ , podczas gdy w pozostałych fazach snu spotykane są nawet w bezpośrednim sąsiedztwie wrzeciona o amplitudzie  $30\div 80\ \mu\text{V}$ .



**Rys. 4.6.** Reprezentacja zapisu EEG na płaszczyźnie czasowo-częstotliwościowej: a) użycie 50 funkcji bazowych zapewnia reprezentację 94,25% energii; b) użycie 100 funkcji – 97,7% energii; c) użycie 200 funkcji – 99,32 energii



Rys. 4.7. Identyfikacja wrzecion snu (A, B, C) oraz (D, E) na płaszczyźnie czasowo-częstotliwościowej metodą poszukiwania dopasowań

#### 4.4. Ocena zmienności rytmu serca za pomocą sieci falkowo-neuronowej

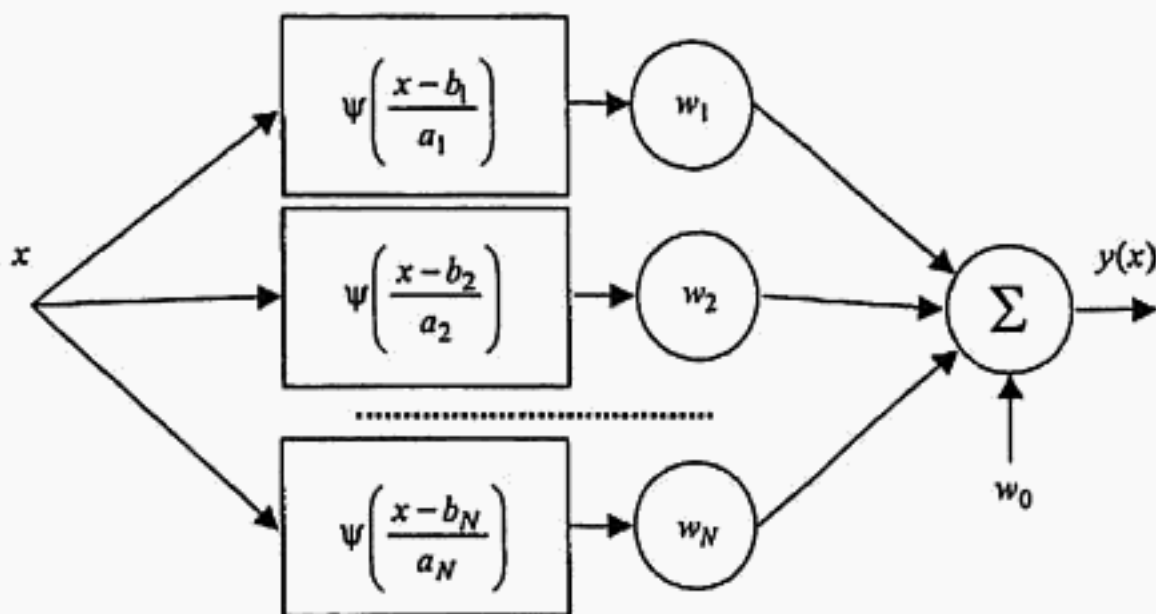
Wyodrębnianie cech charakterystycznych dla określonych patologii z sygnału elektrodiagnostycznego jest podstawowym celem analizy w dziedzinie czasowo-częstotliwościowej. Często stosowane jest porównanie reprezentacji sygnału bieżącego ze wzorcem sygnału patologii w dziedzinie czasowo-częstotliwościowej albo binaryzacja określonych współczynników reprezentacji z użyciem doświadczalnie wybranych wartości progowych. W przypadku bardziej skomplikowanych zależności te proste środki podejmowania decyzji przestają jednak wystarczać i nieodzowne jest użycie metod sztucznej inteligencji. Jedną z nich jest sieć falkowo-neuronowa WNN (ang. *wavelet neural networks*), która łączy w sobie cechy transformacji czasowo-częstotliwościowej i rozpoznawania cech.

Idea sieci falkowo-neuronowych została zaproponowana przez Zhanga i Benveniste'a w 1992 roku i polega na zastąpieniu funkcji aktywacji  $N$  neuronów jednowarstwowej sieci neuronowej typu perceptron przez funkcje falkowe wygenerowane z tej samej falki-matki za pomocą skalowania i przesuwania (4.1). Przekształcenie falkowe dodane na wejście struktury sieci neuronowej (rys. 4.8) pełni funkcję klasyfikatora mającego za zadanie wyodrębnienie tylko tej części informacji zawartej w sygnale, która jest najbardziej znamienna w aspekcie dokonywanej analizy.

$$y(x) = \sum_{i=1}^N w_i \psi\left(\frac{x - b_i}{a_i}\right) + w_0 \quad (4.4)$$

gdzie  $a_i$  oraz  $b_i$  oznaczają, odpowiednio, skalowanie i przesunięcie, natomiast  $w_i$  jest  $i$ -tym współczynnikiem wagowym.

Podobne rozwiązanie można zastosować w sieci wielowarstwowej, której warstwa wejściowa będzie stanowiła bazę dekompozycji. Baza dekompozycji może być ortogonalna, o ile zostanie użyta ortogonalna rodzina falkowa. Baza ortogonalna może być zupełna, jeżeli liczba neuronów będzie odpowiadała całkowitej liczbie atomów powierzchni czasowo-częstotliwościowej. Zwraca uwagę fakt, że w przeciwieństwie do klasycznej sieci neuronowej, poszczególne neurony warstwy wejściowej (wavelony) mają różne funkcje aktywacji, co oznacza wstępną „specjalizację” w wykrywaniu różnych cech sygnału. Z tego powodu sieci falkowo-neuronowe są przydatne przede wszystkim do wyszukiwania ściśle zdefiniowanych wzorców. Niestety, ceną za tę specjalizację jest znaczna wrażliwość procesu rozpoznania na warunki początkowe.



Rys. 4.8. Struktura sieci falkowej (rolę funkcji aktywacji pełnią funkcje falkowe)

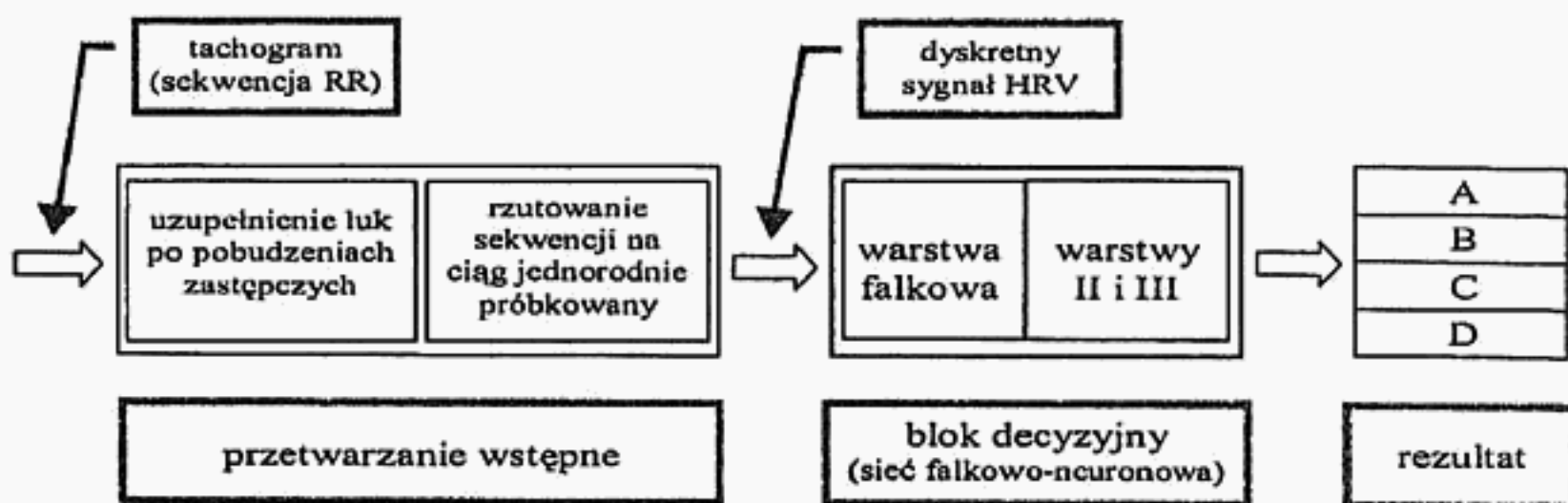
Znane są zastosowania sieci falkowo-neuronowych w analizie akustycznych potencjałów wywołanych, natomiast algorytm opisany w niniejszym podrozdziale wykorzystuje taką sieć do klasyfikacji sygnału zmienności rytmu serca HRV (ang. *heart rate variability*) pacjentów z chorobą naczyniową. Celem algorytmu jest określenie poziomu choroby naczyniowej na podstawie zapisu elektrokardiograficznego – taniej i nieinwazyjnej metody elektrodiagnostycznej.

Przed użyciem w charakterze sygnału wejściowego do systemu WNN ciąg interwałów międzyuderzeniowych RR zostaje wstępnie przetworzony (rys. 4.9). Ponieważ zmienność rytmu serca jest oceniana za pośrednictwem węzła zatokowego i pobudzenia zastępcze nie mogą być brane pod uwagę, na wstępie dokonywane jest uzupełnienie luk po zespołach QRS innych niż nadkomorowe. Kolejnym problemem jest niejednorodność interwałów próbkowania tachogramu, jako że wartość interwału RR określana jest tylko w momencie wystąpienia kolejnego uderzenia serca. Przed zastosowaniem transformacji falkowej konieczne jest więc przekształcenie go w sygnał, w którym określenie częstości akcji serca będzie możliwe w stałych interwałach czasowych. Spośród wielu możliwości takiego przekształcenia największe uzasadnienie na gruncie fizjologii pobudzania i hamowania czynności węzła zatokowego serca ma metoda



aproksymacji funkcjami sklejanymi trzeciego stopnia. Zastosowanie tej aproksymacji pozwala na dowolne umiejscowienie docelowych próbek sygnału. W opisywanym algorytmie zastosowano stały interwał próbkowania równy 100 ms.

Wybór falki-matki w opisywanym zastosowaniu nie był ograniczony warunkiem dokładnej rekonstrukcji sygnału. Zastosowano falę Morleta, dzięki czemu istnieje możliwość wygodnego doboru współczynnika kontrakcji poszczególnych skal, a ich częstotliwości środkowe są optymalnie dobrane do charakterystyk poszukiwanych cech. Warstwa wejściowa sieci neuronowej oblicza podzbiór współczynników czasowo-częstotliwościowej reprezentacji sygnału wybrany eksperymentalnie. Wstępna specjalizacja neuronów warstwy wejściowej została w ten sposób wykorzystana w celu optymalizacji procesu rozpoznania w sieci o jak najmniejszych rozmiarach. Analiza statystyczna poszczególnych współczynników reprezentacji czasowo-częstotliwościowej w próbkach pacjentów reprezentatywnych dla poszukiwanej choroby i grupie kontrolnej jest źródłem informacji o położeniu najbardziej charakterystycznych cech chorobowych w dziedzinie współczynników falkowych. Preferowane są te współczynniki, dla których wariancja jest maksymalnie skorelowana z poziomem choroby naczyniowej.



Rys. 4.9. Struktura blokowa algorytmu wykorzystującego sieci falkowo-neuronowe

Zastosowana sieć neuronowa ma strukturę dwuwarstwową. W pierwszej warstwie zastosowano tangsoidalną, a w drugiej liniową funkcję aktywacji. Analogicznie do klasycznej sieci neuronowej, konieczne jest przeprowadzenie procesu uczenia polegającego na iteracyjnych pokazach zadań o znanych rozwiązaniach i modyfikacji współczynników wagowych sygnałem błędu rozpoznania. Po ustaleniu wartości współczynników wagowych klasyfikator został podany testowaniu przy użyciu nieznanych dotąd zadań rozpoznawania. Opisywany klasyfikator charakteryzuje się czułością równą 78% i specyficznością równą 83%.



## **ROZDZIAŁ 5**

# **Filtracja sygnału w dziedzinie czasowo-częstotliwościowej o parametrach dobieranych adaptacyjnie**

Cechą wspólną metod opisanych w tym rozdziale jest przekształcanie czasowo-częstotliwościowej reprezentacji sygnału. W opisywanych zastosowaniach powierzchnia czasowo-częstotliwościowa jest alternatywną formą reprezentacji sygnału wybraną do manipulacji ze względu na szczególne własności rozłączności czasu i częstotliwości (skali) reprezentowanych na prostopadłych osiach. Sygnał zmodyfikowany w dziedzinie czasowo-częstotliwościowej jest następnie transformowany do dziedziny czasu. Modyfikacje sygnału w dziedzinie czasowo-częstotliwościowej implikuje zatem stosowanie bezstratnej transformacji falkowej posiadającej własność dokładnej rekonstrukcji; w przeciwnym przypadku zmiany zawartości informacyjnej sygnału nie będą kontrolowane przez operacje wykonywane w dziedzinie czasowo-częstotliwościowej.

### **5.1. Eliminacja zakłóceń sygnału okoruchowego**

#### **5.1.1. Opis sygnału okoruchowego i jego zastosowań diagnostycznych**

Jednym ze sposobów wykorzystania sygnału okoruchowego interesującym z punktu widzenia lekarza klinicysty jest badanie stanu i określanie przyczyn zaburzeń funkcjonowania układu równowagi przez wywoływanie oczopląsów i porównywanie ich parametrów. Według obecnego stanu wiedzy lekarskiej, stymulacja fizjologiczna lub sztuczna choćby jednego z trzech podstawowych ośrodków równowagi doprowadza do złożonych reakcji organizmu, wyrażających się zawrotami głowy, oczopląsem, zaburzeniami równowagi statycznej i dynamicznej oraz objawami wegetatywnymi. Proporcje występowania wymienionych zjawisk są uzależnione przede wszystkim od sposobu stymulacji oraz fizjologicznej sprawności narządu przedsionkowego, narządu wzroku i czucia proprioceptywnego. Najczęstszym objawem zakłóceń równowagi jest oczopląs, który – choć nie jest jedyną odpowiedzią na zadany bodziec – zawiera wiele cennych informacji diagnostycznych pozwalających na obiektywną ocenę reakcji układu równowagi, a przez to – jego stanu klinicznego. Przez oczopląs rozumiane są poziome, pionowe lub elipsoidalne ruchy gałek ocznych. Podstawową cechą klasycznego

oczopłasu jest występowanie w tym zjawisku faz: szybkiej i wolnej, rozróżnialnych na podstawie prędkości kątowej ruchu gałki ocznej. Przedmiotem badań jest kilkanaście rodzajów oczopłásów rozróżnianych w oparciu o przyczyny powstawania.

W praktyce klinicznej do diagnostyki narządu równowagi wykorzystuje się trzy następujące:

- 1) oczopłás przedsionkowy poobrotowy,
- 2) oczopłás przedsionkowy postkaloryczny,
- 3) oczopłás optokinetyczny.

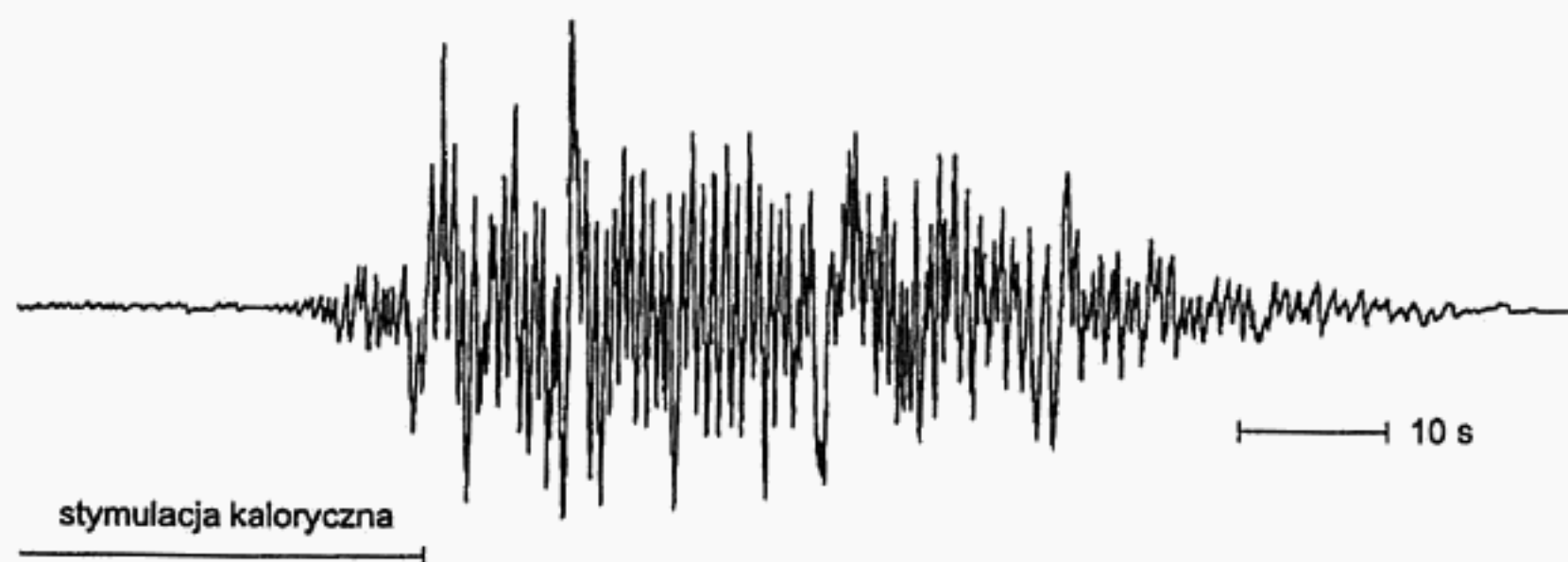
Wywoływanie oczopłasu poobrotowego jest najbardziej zbliżone do fizjologicznych warunków pracy układu równowagi. Na skutek występujących w ruchu obrotowym przyspieszeń następuje przemieszczanie się endolimfy zawartej w trzech wzajemnie prostopadłych kanałach półkolistych błędnika, związane z jej bezwładnością (tzw. limfokineza). Prędkość ruchu endolimfy reprezentuje wartość doznanego przyspieszenia, a proporcje prędkości w trzech kanałach – jego kierunek. Wewnątrz kanałów półkolistych znajdują się rzęski nabłonka zmysłowego (*stereocilia*), których wychylenia – powodowane ruchem endolimfy – wywołują pobudzanie lub hamowanie aktywności bioelektrycznej kanałów, co wyraża się wzrostem lub spadkiem częstotliwości impulsacji włókien nerwu przedsionkowego. Informacja o doznanym przyspieszeniu jest przetwarzana w mózdzku na sygnały sterujące mięśniami poruszającymi gałką oczną (sygnały okulomotoryczne) w taki sposób, aby zapewnić niezmiennosc osi optycznej patrzenia. Zjawisko to umożliwia np. zatrzymanie wzroku na nieruchomym celu podczas obrotu głowy. Występuje ono nawet w przypadku braku postrzeganego celu (tj. przy zamkniętych oczach), co dowodzi, iż przyczyną ruchu gałek ocznych jest bodziec pochodzenia błędnikowego. Ruch „śledzenia” tworzy fazę wolną oczopłasu. Wobec ograniczonego zakresu ruchu gałki ocznej w oczodole (zwykle ok.  $90^{\circ}$ ), naprzemian z opisanymi wyżej ruchami „śledzenia” muszą występować tzw. ruchy kompensacyjne, przemieszczające gałkę oczną w kierunku przeciwnym. Ruchy te tworzą fazę szybką oczopłasu. Prędkość kątowa ruchu gałki ocznej zależy od doznawanego przyspieszenia wyłącznie w fazie wolnej (bodziec błędnikowy), natomiast w fazie szybkiej jest stała. Dla celów diagnostycznych stosuje się przyspieszenia chwilowe (nagła zmiana prędkości obrotowej) i sinusoidalne. Fotel obrotowy przeważnie umieszczany jest w ciemni w celu eliminacji efektu fiksacji wzroku, a zatem ograniczenia wpływu bodźców okoruchowych pochodzenia pozabłędnikowego. Wadą metody obrotowej jest poddawanie obu błędników takim samym przyspieszeniom, co wynika ze sztywności ich mechanicznego połączenia przez kości czaszki. Wobec powyższego metoda ta nie umożliwia precyzyjnej lokalizacji źródła zaburzeń równowagi.

Kaloryczna (cieplna) metoda wywoływania oczopłasu polega na wprawianiu w ruch endolimfy w kanale półkolistym poziomym (usytuowanym pionowo, gdy pacjent znajduje się w pozycji leżącej) poprzez pośrednie miejscowe ogrzanie lub ochłodzenie. Konwekcyjny ruch endolimfy powoduje – podobnie jak w przypadku użycia

przyspieszenia – reakcję systemu nerwowego i ruch gałek ocznych typu oczopląsu. Choć reakcja ta nie jest typowo fizjologiczna, jej powszechne wykorzystanie do diagnostyki układu równowagi wynika z możliwości pobudzenia tylko jednego kanału. Pobudzenia kalorycznego dokonuje się zazwyczaj poprzez wprowadzenie do przewodu słuchowego wody o temperaturze różniącej się od temperatury ciała. Powszechnie przyjętym w praktyce klinicznej testem układu równowagi człowieka jest tzw. próba cieplna Hallpike’a, uważana za jedno z najlepszych kryteriów oceny narządu równowagi. Polega ona na podrażnieniu układu równowagi określonymi bodźcami i rejestracji, a następnie analizie jego odpowiedzi. Próba ta pozwala określić jako źródło nieśprawności obwodową lub ośrodkową część przedsionka, co ma istotne znaczenie dla leczenia pacjenta.

Powszechnie do drażnienia błędnika stosuje się wodę o temperaturze  $30^{\circ}\text{C}$  i  $44^{\circ}\text{C}$ , zwaną dalej „zimną” i „ciepłą”. Tak dobrane wartości temperatur są wystarczająco odległe od średniej temperatury ciała badanego, która wynosi  $37^{\circ}\text{C}$ , aby na skutek limfokinezy spowodować oczopląs. Oczopląs ten traktowany jest jako mierzalna reakcja badanego systemu. Bezpośrednio po rozpoczęciu pobudzenia cieplnego pojawia się reakcja w postaci oczopląsu, której szczyt (maksimum amplitudy) przypada zazwyczaj ok. 60÷80 s później. W badaniach klinicznych stosowany jest zatem następujący protokół badania: 30 s – pobudzenie (0,25 l wody), 30 s – przerwa, oraz 30 s – rejestracja ruchu gałek ocznych. Taki protokół badania pozwala uzyskać za każdym razem najwłaściwszy do diagnozy zapis na szczycie reakcji.

Reakcja oczopląsowa na wlanie ciepłej wody do prawego ucha (lub zimnej wody do lewego ucha) jest następująca (rys. 5.1): wolny ruch gałki ocznej w prawo, a następnie szybki ruch w lewo, określane odpowiednio jako faza wolna i szybka oczopląsu. Reakcja oczopląsowa na wlanie zimnej wody do prawego ucha (lub ciepłej wody do lewego ucha) jest odwrotna. W przypadku fizjologicznie prawidłowym prędkość kąta ruchu gałki ocznej w fazie wolnej oczopląsu jest jednakowa we wszystkich czterech wymienionych przypadkach. Pojawienie się asymetrii reakcji stanowi podstawową przesłankę diagnostyczną.



Rys. 5.1. Reakcja oczopląsowa podczas próby kalorycznej



Według opisanego schematu przeprowadzane są u każdego pacjenta cztery tzw. próby kaloryczne, a następnie na podstawie analizy statystycznej zarejestrowanych sygnałów obliczane są dwa współczynniki: CP (ang. *canal paresis*) określający symetrię pobudzenia błędników oraz DP (ang. *directional preponderance*) określający przewagę kierunkową oczopląsu:

$$CP = \frac{(V_{zl} + V_{cl}) - (V_{zp} + V_{cp})}{V_{zl} + V_{cl} + V_{zp} + V_{cp}} \cdot 100\% \quad (5.1a)$$

$$DP = \frac{(V_{zl} + V_{cp}) - (V_{zp} + V_{cl})}{V_{zl} + V_{cl} + V_{zp} + V_{cp}} \cdot 100\% \quad (5.1b)$$

gdzie:

- $V$  – szybkość fazy wolnej w poszczególnych badaniach;
- $zl$  – zimna woda, lewe ucho;
- $zp$  – zimna woda, prawe ucho;
- $cl$  – ciepła woda, lewe ucho;
- $cp$  – ciepła woda, prawe ucho.

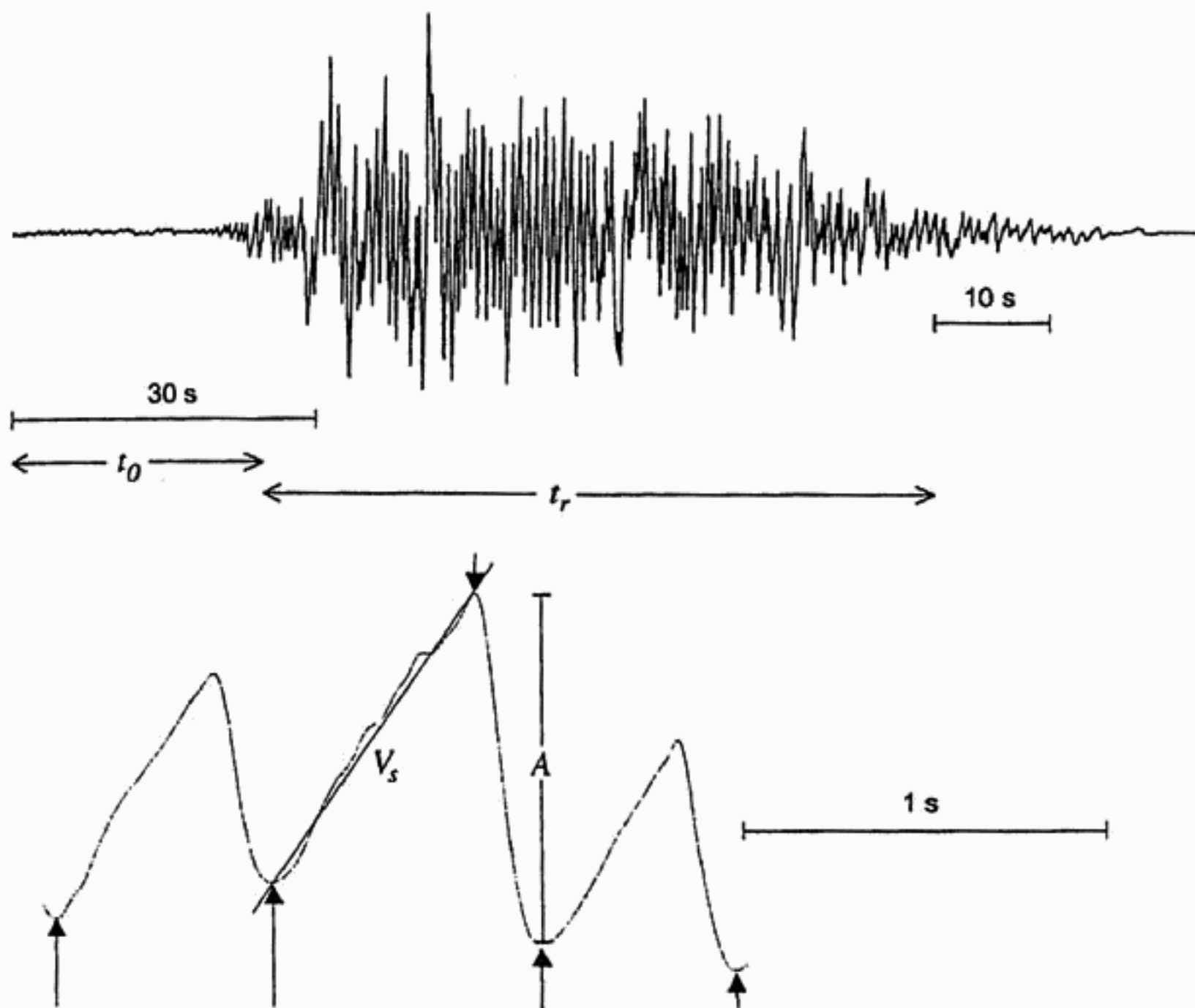
Podane wzory, zaproponowane przez Jongkeesa, są biologiczną podstawą ilościowej oceny elektronystagmogramów. Warto zauważyć, że współczynnik CP w sposób istotny zależy tylko od asymetrii reakcji oczopląsu na pobudzenia – z punktu widzenia kierunku pobudzeń (ucho prawe, ucho lewe), natomiast współczynnik DP – tylko od asymetrii reakcji na pobudzenia – z punktu widzenia kierunku fazy wolnej oczopląsu (w prawo, w lewo).

Z literatury przedmiotu wynika, że maksymalne, uważane jeszcze za fizjologicznie prawidłowe, wartości tych współczynników wynoszą odpowiednio:

- dla CP 22%,
- dla DP 28%.

Każde istotne przekroczenie tych granic wskazuje na obecność patologii i stanowi podstawową informację, interesującą prowadzącego badania lekarza.

Użycie wody do irygacji ucha zewnętrznego nastręcza wiele trudności technicznych. Trudny jest dokładny pomiar prędkości strumienia wody oraz utrzymanie jej temperatury, wzrasta również ryzyko porażenia elektrycznego. Podjęto więc próby zastąpienia wody powietrzem w charakterze medium transportującego energię cieplną. Ponieważ wymiana energetyczna pomiędzy powietrzem a tkanką jest znacznie mniej efektywna (na skutek niższego ciepła właściwego powietrza), konieczne jest w tym przypadku użycie znacznie większej objętości powietrza o temperaturach bardziej odległych od temperatury ciała pacjenta. Równoważnikiem 30-sekundowej irygacji z użyciem 0,25 l wody o temperaturze 30 (44)°C jest trwająca 60 s stymulacja powietrzem (8 litrów) o temperaturze 24 (50)°C.



**Rys. 5.2.** Parametry diagnostyczne typowe dla oczopląsów wywołanych pobudzeniem kalorycznym (opis w tekście)

Parametrami diagnostycznymi oczopląsów wywołanych pobudzeniem kalorycznym są (rys. 5.2):

- $A$  – amplituda (wartość średnia i odchylenie standardowe oraz wartość maksymalna),
- $V_s$  – prędkość kątowna fazy wolnej oczopląsu,
- $t_r$  – czas trwania reakcji oczopląsu,
- $t_0$  – czas utajenia reakcji,
- $k$  – liczba wychyleń gałek ocznych (częstotliwość podstawowa oczopląsu).

### 5.1.2. Pozyskiwanie sygnału okoruchowego

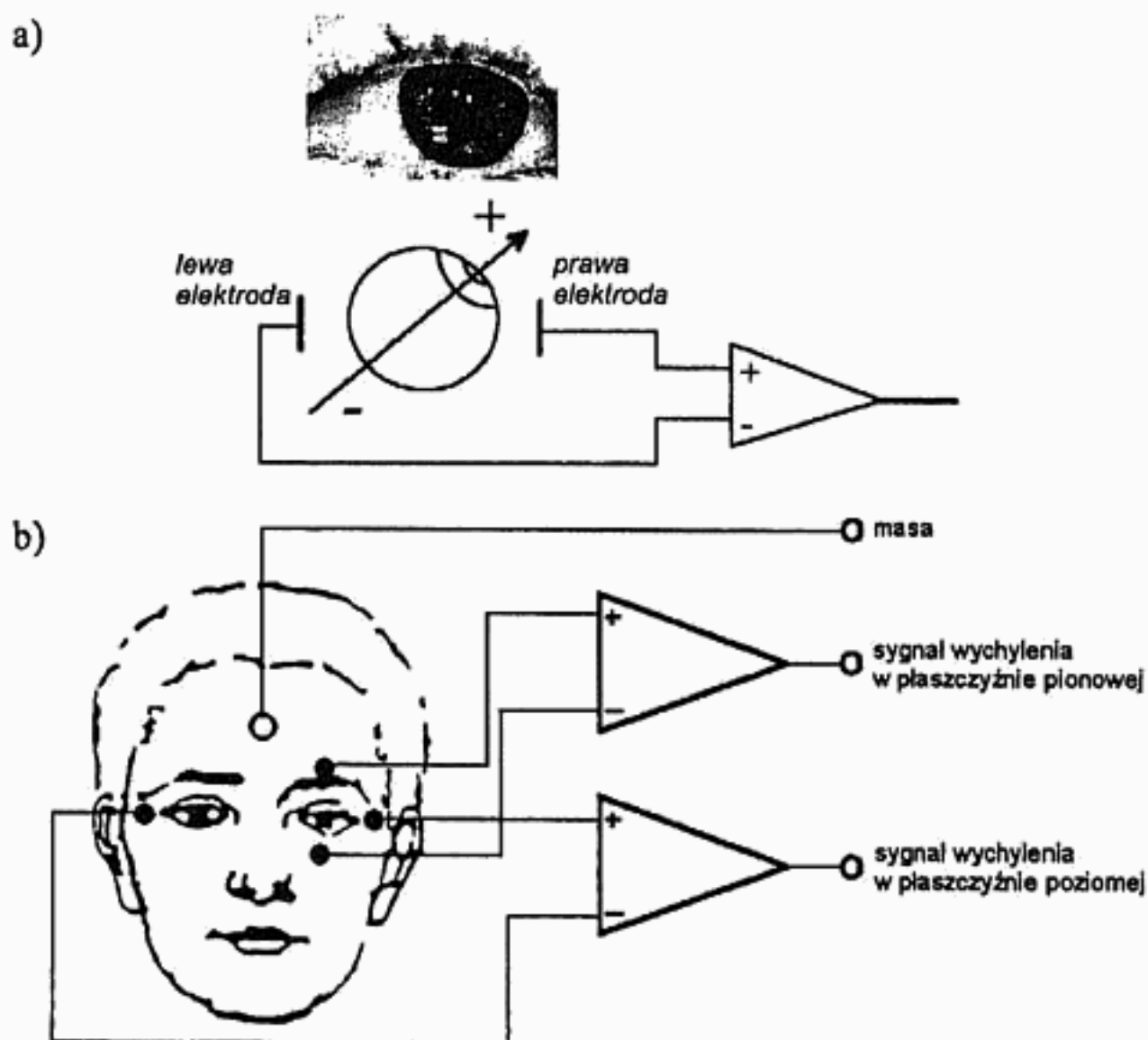
W celach klinicznej diagnostyki układu równowagi stosuje się przeważnie metodę elektryczną pozyskiwania sygnału okoruchowego. Sygnał pozyskany za pomocą tej metody nazywany jest **sygnałem elektronystagmograficznym** (w skrócie **ENG**).



Zaletami elektronystagmografii w tym konkretnym zastosowaniu są:

- nieinwazyjność,
- możliwość rejestracji ruchu gałek ocznych za zamkniętymi powiekami (nie ma konieczności stosowania ciemni),
- niski koszt aparatury,
- wystarczająca dokładność.

Swą zasadę fizyczną ta najwcześniej stosowana metoda rejestracji pozycji gałki ocznej opiera na doświadczeniu Du Bois Reymonda wykonanym w 1849 roku. Dowiedziono wówczas, że istnieje różnica potencjałów między siatkówką (*retina*) a rogówką (*cornea*), nadająca gałce ocznej charakter dipola elektrycznego. Różnica potencjałów wynosi ok.  $18 \pm 20$  mV i jest zależna (w zakresie 20% swej wartości) od takich czynników, jak oświetlenie, zmęczenie wzroku czy wilgotność powierzchni gałki ocznej. Poruszająca się w oczodole gałka oczna wywołuje obserwowalne na jego krawędziach zmiany pola elektrycznego (rys. 5.3). Wskutek bocznikującego charakteru otaczających tkanek i zakłóceń ze strony innych biopotencjałów praktyczna wartość napięcia sygnału okoruchowego mierzonego elektrodami naklejanymi na krawędziach oczodołu nie przekracza  $7 \text{ mV/}^\circ$ .



Rys. 5.3. Pobranie informacji okoruchowej metodą elektryczną: a) zasada fizyczna pomiaru; b) umiejscowienie elektrod i rozkład odprowadzeń

Spośród wad metody elektrycznej należy wymienić niewielką wartość opisanego wyżej napięcia oraz pośrednictwo naskórka przy jego przekazywaniu. Powodują one konieczność stosowania stopnia wstępnego urządzenia rejestrującego o znacznej impedancji wejściowej, wynikającą z dążenia do jak najmniejszego obciążenia źródła sygnału. Jest to przyczyną znacznej podatności sygnału okoruchowego na zakłócenia, co stanowi najistotniejszą wadę opisywanej metody. Amplituda zakłóceń różnej natury niejednokrotnie znacznie przekracza amplitudę sygnału, stwarzając trudności interpretacyjne nawet doświadczonemu personelowi medycznemu. W praktyce klinicznej redukcję zakłóceń uzyskuje się głównie poprzez znaczne ograniczenie pasma przenoszenia rejestratorów i stosowanie filtrów pasmowozaporowych (50 Hz, 35 Hz) o znacznym tłumieniu, co powoduje jednak zauważalne zubożenie sygnału. Odwzorowanie położenia oka jest także znacznie zakłócanie przez zjawiska zachodzące na styku elektrod ze skórą, będące źródłem sygnałów wolnozmiennych o znacznych amplitudach, sięgających nawet 20 mV. Powoduje to konieczność odcięcia również składowej stałej sygnału (poniżej 0,1 Hz) lub częstej kalibracji aparatury, co nie rekomenduje sygnału uzyskanego metodą elektryczną do użycia w systemach sterowania (z wyjątkiem sterowania przyrostowego).

Do najistotniejszych zalet akwizycji sygnału okoruchowego metodą elektryczną zaliczyć należy całkowitą nieinwazyjność oraz możliwość pobierania informacji okoruchowej przy zamkniętych powiekach (eliminacja zjawiska fiksacji wzroku). Inną zaletą opisywanej tu metody jest niewysoki koszt stosowanej aparatury przy wysokiej dokładności odwzorowania przemieszczeń gałki ocznej. Dzięki połączeniu tych zalet metoda elektryczna jest powszechnie stosowana w diagnostyce klinicznej narządu równowagi (tzw. próba kaloryczna Hallpike'a) opisaną w podrozdziale 5.1.1.

Stosowana wyłącznie na skalę laboratoryjną modyfikacja akwizycji sygnału okoruchowego metodą elektryczną, polegająca na zastosowaniu elektrod implantowanych, pozwala wyeliminować znaczną podatność sygnału na zakłócenia. Amplituda sygnału, możliwa do uzyskania z platynowych elektrod implantowanych, sięga 100 mV/°. Dodatkowo, dzięki wyeliminowaniu kontaktu elektroda – skóra, sygnał ten charakteryzuje się stałością amplitudy w czasie, przez co znacznie wzrasta wierność odwzorowania położenia gałki ocznej. Niestety, na skutek inwazyjności (elektrody są operacyjnie implantowane w kości skroniowej), zakres stosowalności tak zmodyfikowanej metody elektrycznej jest ograniczony do laboratoryjnych badań naukowych wykonywanych na zwierzętach doświadczalnych.

Fakt, że analiza oczopląsu podczas próby cieplnej Hallpike'a polega wyłącznie na obróbce geometrycznej (pomiar nachyleń linii będących miarą prędkości), a później na analizie statystycznej wprowadzonych danych, szczególnie predestynuje ją do przeprowadzenia za pomocą komputera oraz pozwala spodziewać się znacznych wynikających z tego korzyści. Mimo to, powszechnie spotyka się jeszcze manualną metodę oceny elektronystagmogramów poprzez mierzenie amplitud linijką i kątów nachylenia wykresów faz wolnych – kątomierzem, następnie uśrednianie i podstawianie do wzoru Jongkeesa.

### 5.1.3. Zakłócenia sygnału okoruchowego i metody ich ograniczania

Najpoważniejszą wadą metody elektrycznej jest jej znaczna podatność na zakłócenia, których źródła rozproszone są w całym praktycznie torze pozyskiwania sygnału. Jest to spowodowane przede wszystkim niską amplitudą pozyskiwanego sygnału, a także występowaniem w sąsiedztwie lub w samej gałce ocznej organów charakteryzujących się aktywnością elektryczną w podobnym zakresie napięć i częstotliwości.

Źródłami zakłóceń są:

- szумы układów elektronicznych toru wzmacnienia;
- iskrzenie styków połączeń komutowanych w obrębie aparatury;
- przydźwięk sieci wynikający z niedostatecznej filtracji tętnień zasilacza;
- przebiegi wolnozmiennne związane z temperaturowymi zmianami własności elementów elektronicznych toru wzmacnienia;
- zakłócenia impulsowe i przydźwięk sieci indukowane w doprowadzeniach elektrod i ciele pacjenta przez pole elektromagnetyczne z otoczenia;
- przebiegi wolnozmiennne związane ze zjawiskami elektrolitycznymi występującymi na styku elektrod z naskórkiem oraz ze zmienną w czasie impedancją naskórka;
- przebiegi biologiczne wynikające z aktywności mięśni okoruchowych, mięśni powiek, mięśni twarzy, a także elektroretinogramy, elektrokardiogramy oraz elektroencefalogramy nakładające się na sygnał ENG;
- wahania potencjału rogówkowego w zależności od natężenia padającego światła i innych czynników zewnętrznych.

Wpływ niektórych z wymienionych rodzajów zakłóceń może zostać ograniczony poprzez stosowanie niskoszumowej aparatury najwyższej jakości, ekranów elektromagnetycznych oraz przez odpowiednie przygotowanie naskórka w miejscu kontaktu elektrod. Są to obecnie czynności rutynowe stosowane przy rejestracji biopotencjałów. Dotyczy to jednak wyłącznie zakłóceń pochodzenia „technicznego”. Zakłócenia elektronystagmogramu, których przyczyną jest aktywność organizmu spoza zakresu zainteresowań diagnostyki układu równowagi, mogą być wprawdzie częściowo zminimalizowane poprzez staranny dobór warunków badania, jednakże sama natura tych zakłóceń nie pozostawia nadziei na ich całkowitą eliminację.

Konsekwencje występujących zakłóceń mogą być następujące:

- brak wpływu na parametry diagnostyczne (mały poziom zakłóceń),
- brak wpływu na parametry diagnostyczne (obliczanie parametrów diagnostycznych z wycinków sygnału uznanych za niezakłócone),
- obniżenie dokładności parametrów diagnostycznych związane z drobnymi pomyłkami interpretacji lub agresywną filtracją elektronystagmogramu,
- utrata znaczenia diagnostycznego obliczanych parametrów związana z grubymi pomyłkami interpretacji elektronystagmogramu.

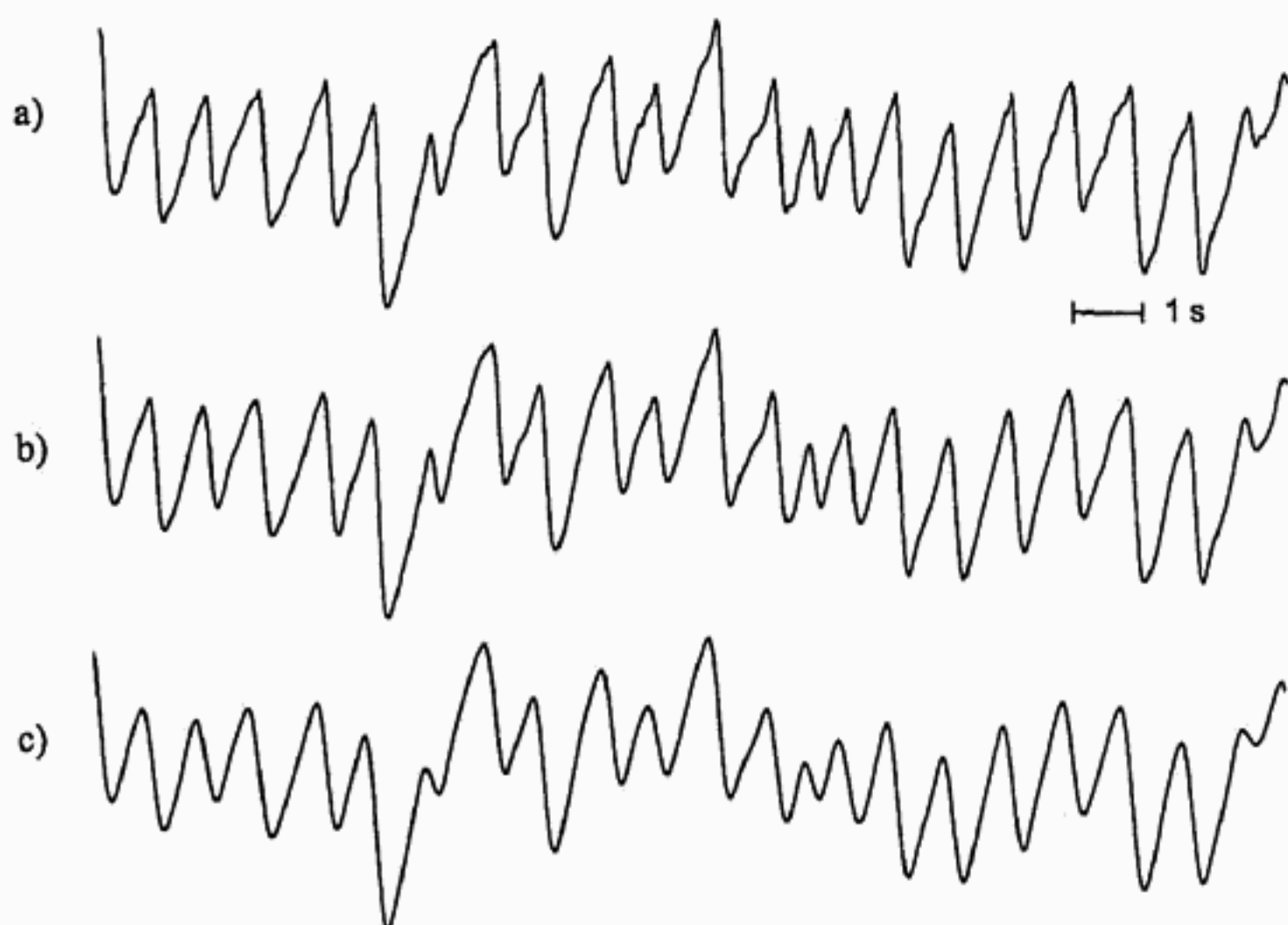


Operacje mające na celu zmniejszenie wpływu zakłóceń na analizowany sygnał można podzielić na dwie grupy:

- 1) modyfikacje pasma rejestrowanych częstotliwości (filtracja),
- 2) wybór najmniej zakłóconych partii zapisu (rozpoznawanie).

Filtracja częstotliwościowa rejestrowanego sygnału elektronystagmograficznego jest podstawową i powszechnie stosowaną metodą poprawy stosunku sygnału do zakłóceń. Wykorzystuje ona fakt znacznego odstępu częstotliwości podstawowej sygnału okoruchowego ( $0,5 \div 3$  Hz) zarówno od zakłóceń sieciowych (50 Hz), jak i od zakłóceń wolnozmiennych (poniżej 0,1 Hz). W zależności od potrzeb stosuje się filtry pasmowo-przepustowe o dolnej częstotliwości odcięcia rzędu 0,2 Hz i górnej – rzędu  $5 \div 25$  Hz oraz nachyleniu charakterystyk  $12 \div 18$  dB/okt. Stopień ingerencji filtracji w sygnał jest zazwyczaj dobierany eksperymentalnie dla danych warunków rejestracji (pracowni) z punktu widzenia minimalizacji grubych pomyłek interpretacji (rys. 5.4). Szczegółowa analiza częstotliwościowa sygnału elektronystagmograficznego prowadzi jednak do wniosku, że widmo jest skupione wokół częstotliwości podstawowej dla niewielkich dysproporcji prędkości ruchu gałki ocznej pomiędzy fazą wolną i szybką. W miarę wzrostu prędkości fazy szybkiej widmo sygnału zawiera coraz więcej składowych wysokoczęstotliwościowych o istotnej energii. Wykazano, że wierna reprezentacja wszystkich fizjologicznie spotykanych ruchów gałki ocznej (w tym sakkad) zawiera częstotliwości do ok. 70 Hz (przy częstotliwości podstawowej oczopląsu rzędu  $1 \div 2$  Hz). Dobór górnej częstotliwości granicznej filtracji pasmowoprzepustowej jest więc zwykle kompromisem pomiędzy częstością występowania grubych pomyłek interpretacji a dopuszczalnymi zniekształceniami sygnału, spowodowanymi pozbawieniem go składowych wysokoczęstotliwościowych. Zniekształcenia te są przyczyną niedokładności obliczeń parametrów diagnostycznych. Prowadzi to do zaniechania automatycznego wyznaczania niektórych parametrów lub do stosowania tabel korekcyjnych dostosowanych do częstotliwości odcięcia.

Interesującą alternatywę dla filtracji częstotliwościowej stanowi proces nazwany „rozpoznawaniem”. Wykorzystuje on fakt, że zakłócenia w przeważającej większości przypadków pojawiają się przypadkowo, a czas ich trwania jest ograniczony. Rozpoznawanie polega na wydzieleniu z nie filtrowanego sygnału okoruchowego fragmentów spełniających określone kryteria podobieństwa do wzorca. Proces ten wykorzystuje szczególne własności formy sygnału elektronystagmograficznego oraz statystyczny charakter większości parametrów diagnostycznych. Własności procesu rozpoznawania pozwalają uznać go za rodzaj filtracji sygnału w dziedzinie czasu. Większość parametrów diagnostycznych obliczanych jest w oparciu o nie zniekształcone fragmenty sygnału, a więc bez straty dokładności. Jest to okupione brakiem możliwości obliczania parametrów wymagających ciągłości sygnału. Dodatkowo, utrudnione jest obliczanie długości interwałów czasowych zawierających nieciągłości.



**Rys. 5.4.** Sygnał elektronystagmograficzny filtrowany dolnoprzepustowo przy częstotliwościach odcięcia: a) 25 Hz; b) 12 Hz; c) 5 Hz. Agresywna filtracja prowadzi do minimalizacji grubych pomyłek interpretacji, lecz jednocześnie jest przyczyną zaniku istotnych szczegółów sygnału, a w konsekwencji – utraty dokładności parametrów diagnostycznych

Analiza własności opisanych metod pozwala na sformułowanie następujących spostrzeżeń:

- sygnał składa się z występujących przemiennie faz: wolnej i szybkiej, w sygnale rzeczywistym występują dodatkowo zakłócenia;
- głównym źródłem grubych pomyłek interpretacji jest uznanie fazy szybkiej lub zakłócenia za fazę wolną albo uznanie fazy wolnej lub zakłócenia za fazę szybką;
- głównym źródłem niedokładności parametrów diagnostycznych jest nieprecyzyjne określanie punktów zwrotnych rozgraniczających fazy, co jest naturalną konsekwencją agresywnej filtracji sygnału;
- zakłócenia impulsowe i wysokoczęstotliwościowe na zboczu fazy wolnej powodują znacznie poważniejsze konsekwencje niż identyczne zakłócenia na zboczu fazy szybkiej;
- zakłócenia impulsowe i wysokoczęstotliwościowe w środkowej części zboczy faz powodują znacznie poważniejsze konsekwencje niż identyczne zakłócenia w pobliżu punktów zwrotnych;
- silnie zakłócone fragmenty sygnału mogą być odrzucane, zastępowane przebiegiem sztucznym itp., najwłaściwsze jednak wydaje się zastąpienie ich sygnałem filtrowanym w sposób zapewniający eliminację grubych pomyłek interpretacji.

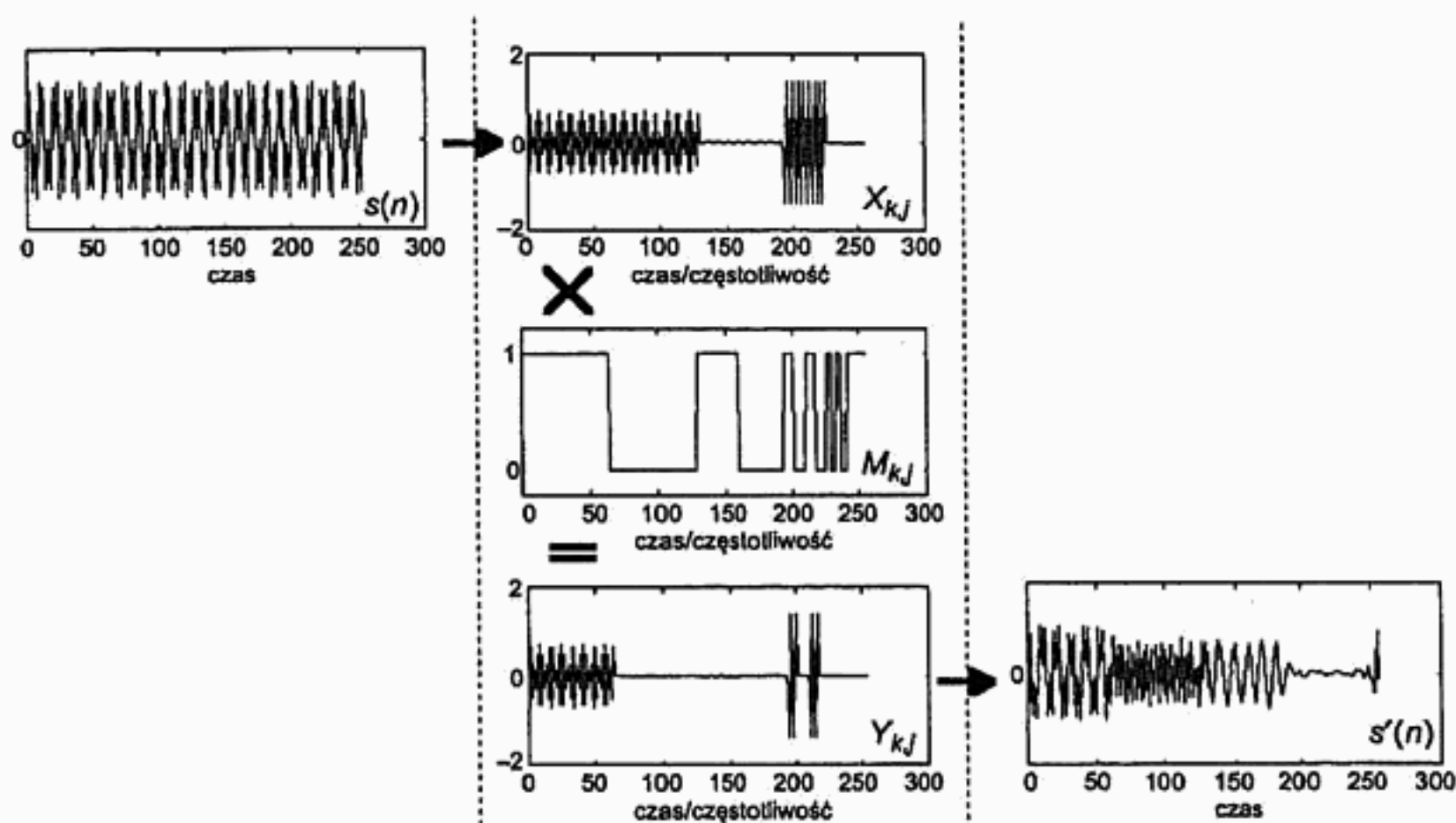


#### 5.1.4. Filtracja sygnału okoruchowego w dziedzinie czasowo-częstotliwościowej

Pomimo korzyści ze stosowania filtracji częstotliwościowej oraz rozpoznawania, żadna z tych metod nie może być uznana za wystarczająco precyzyjne narzędzie obróbki sygnału. Próba połączenia wyłącznie zalet obu opisanych metod prowadzi do wniosku, że najkorzystniejszy z punktu widzenia jakości parametrów diagnostycznych jest następujący schemat filtracji:

1. utworzenie na podstawie sygnału wejściowego kilku sygnałów filtrowanych filtrami dolnoprzepustowymi o różnej częstotliwości granicznej (jest to tzw. bank filtrów),
2. rekonstrukcja sygnału wyjściowego poprzez wybór najkorzystniejszego w danej chwili sygnału filtrowanego.

Filtracja sygnału elektronystagmograficznego w dziedzinie czasowo-częstotliwościowej o charakterystykach dobieranych adaptacyjnie (tj. na bieżąco dopasowywanych do sygnału) jest logicznym rozwinięciem powyższego schematu (rys. 5.5).



Rys. 5.5. Schemat filtracji sygnału w dziedzinie czasowo-częstotliwościowej

Zastosowanie ortogonalnej dekompozycji zapewnia bezstratną konwersję sygnału z dziedziny czasu do dziedziny czasowo-częstotliwościowej i vice versa. Taka metoda filtracji pozwala na precyzyjne wyznaczenie parametrów diagnostycznych z jednoczesnym uniknięciem grubych pomyłek interpretacji. Filtracja sygnału w dziedzinie cza-

sowo-częstotliwościowej o charakterystykach dobieranych adaptacyjnie pozwala na połączenie zalet, a zarazem wyeliminowanie wad, obu znanych wcześniej metod. Właśność ta została potwierdzona eksperymentalnie z użyciem rzeczywistych sygnałów okoruchowych oraz sztucznie wygenerowanych zapisów testowych.

Projektowanie adaptacyjnej filtracji sygnałów w dziedzinie czasowo-częstotliwościowej wymaga rozwiązania takich podstawowych zagadnień, jak:

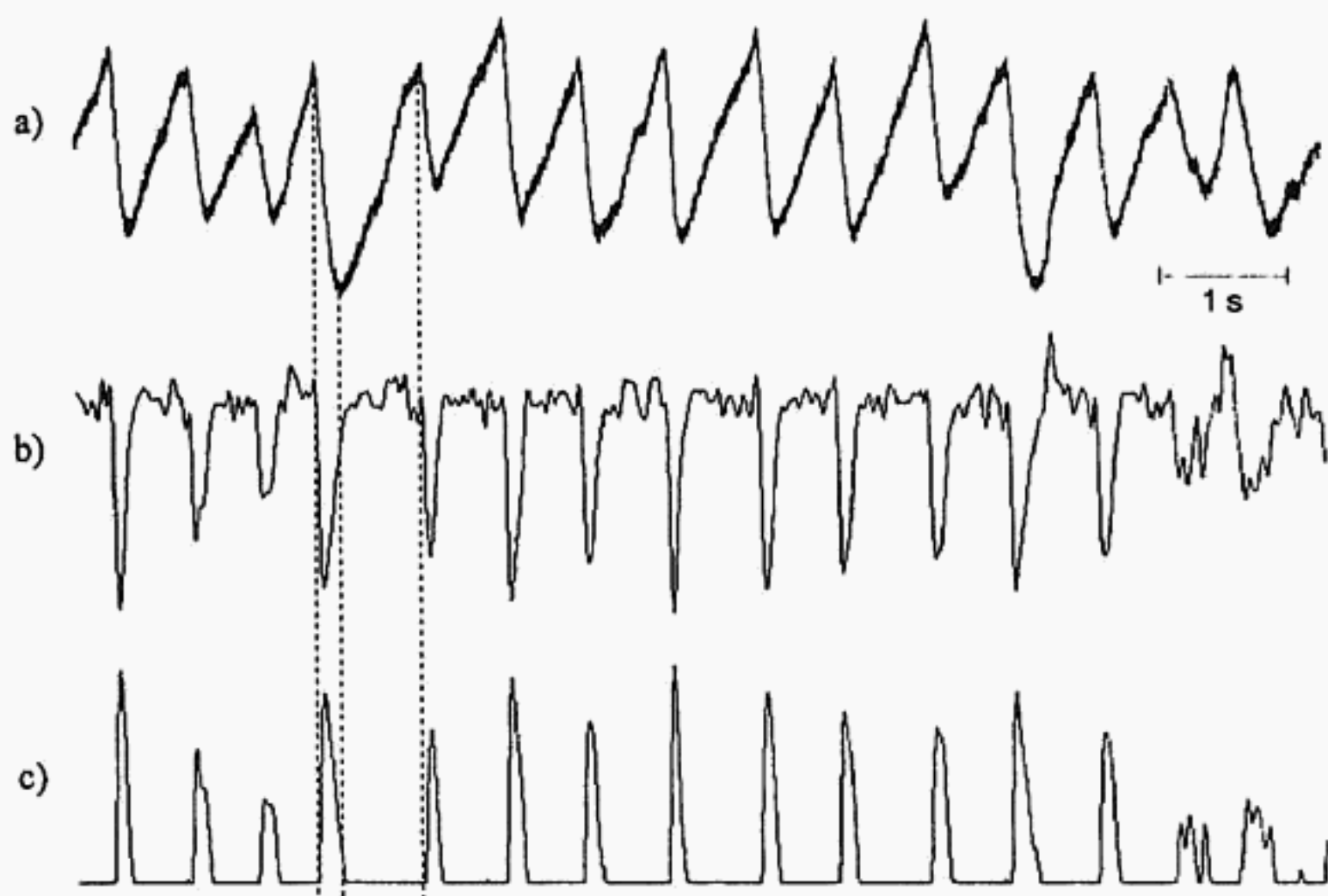
- wybór metody filtracji czasowo-częstotliwościowej sygnału (założeniem jest transformacja bezstratna wykorzystująca dekompozycję ortogonalną),
- wybór parametru (ewentualnie parametrów) sygnału sterujących procesem filtracji,
- zaprojektowanie algorytmu adaptacyjnego doboru charakterystyk filtracji.

Adaptacyjny dobór charakterystyk filtracji czasowo-częstotliwościowej sygnału ENG z wykorzystaniem proporcjonalnej funkcji detekcyjnej jest metodą automatycznej konstrukcji wektora filtru  $M_{k,j}$  (rys. 5.5) opartą na niżej wymienionych własnościach sygnału i spostrzeżeniach.

- Prędkość ruchu gałki ocznej, możliwa do wyliczenia na podstawie sygnału ENG  $s(n)$  jako  $v(n) = s(n) - s(n-1)$ , przyjmuje wartości przeciwnych znaków w fazach wolnej i szybkiej. Znak prędkości ruchu gałki ocznej może być podstawą przybliżonej separacji faz oczopląsu.
- Prędkość ruchu gałki ocznej obliczona w oparciu o rzeczywisty sygnał nie filtrowany nie może być uważana za parametr diagnostyczny z uwagi na obecność zakłóceń wysokoczęstotliwościowych i szumu, lecz z powodzeniem może być podstawą modyfikacji charakterystyki filtracji<sup>1)</sup>.
- Sygnał ENG powinien być filtrowany agresywniej na odcinkach reprezentujących fazę wolną oczopląsu niż na odcinkach reprezentujących fazę szybką.
- W obrębie zgrubnie określonych faz wolnych oczopląsu charakterystyka filtracji nie jest modyfikowana, natomiast w obrębie faz szybkich filtracja dopuszcza przechodzenie składowych wysokoczęstotliwościowych w stopniu proporcjonalnym do wartości prędkości. Wynika to z faktu, że większe wartości prędkości w fazie szybkiej świadczą o istnieniu znaczących składowych wysokoczęstotliwościowych, których przepuszczenie zapewnia utrzymanie zniekształceń sygnału ENG, a więc również niedokładności parametrów diagnostycznych, na minimalnym poziomie.

---

<sup>1)</sup> Modyfikacja charakterystyki filtracji musi uwzględniać ponadto rozdzielczość czasową transformacji czasowo-częstotliwościowej (wynikającą z zasady nieoznaczoności), więc z założenia obciążona jest niedokładnością.



Rys. 5.6. Etapy proporcjonalnego doboru charakterystyk filtracji sygnału ENG w dziedzinie czaso-częstotliwościowej: a) sygnał ENG; b) prędkość chwilowa ruchu gałki ocznej; c) funkcja detekcyjna fazy szybkiej oczopląsu

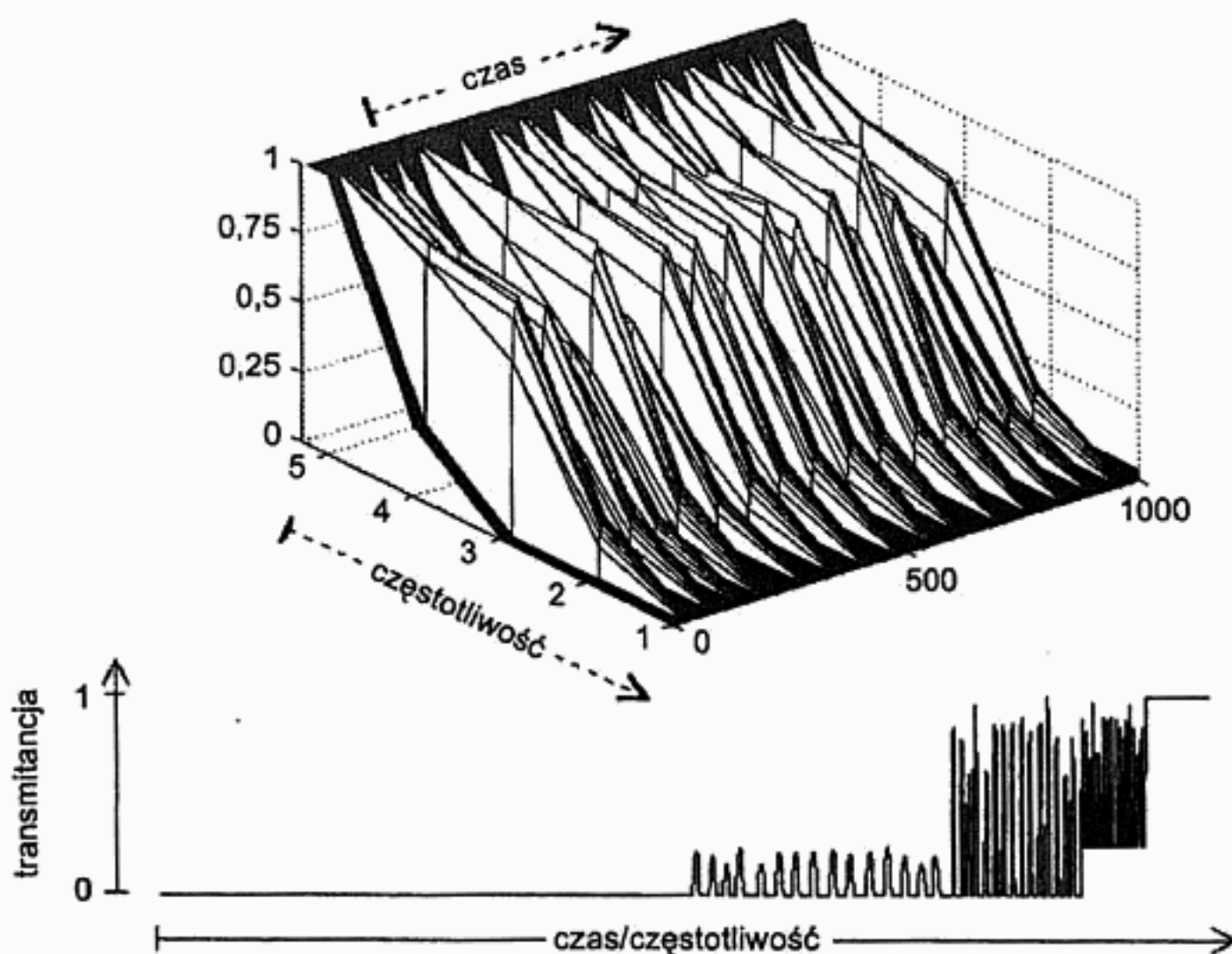
Na podstawie powyższych założeń zaprojektowany został algorytm realizujący proporcjonalny dobór charakterystyk filtracji w następujących etapach (rys. 5.6).

1. Na podstawie sygnału elektronystagmograficznego  $s(n)$  (rys. 5.6a) tworzona jest funkcja detekcyjna  $f_d(n)$  w oparciu o fakt przemienności kierunków prędkości w fazach wolnej i szybkiej. Na podstawie nie filtrowanego sygnału dokonuje się więc detekcji faz i obliczeń prędkości chwilowej ruchu gałki ocznej (rys. 5.6b).
2. Następnie, prędkości odpowiadające wykrytym fazom wolnym są anulowane, a prędkości odpowiadające fazom szybkim – normalizowane do przedziału  $\langle 0, 1 \rangle$ . Tak utworzona funkcja detekcyjna może przyjmować dowolne wartości z przedziału  $\langle 0, 1 \rangle$  (rys. 5.6c), stąd możliwa jest modyfikacja charakterystyk filtracji w sposób proporcjonalny do prędkości kątowej ruchu gałki ocznej w fazie szybkiej.
3. Modyfikacja przebiega niezależnie dla każdej z modyfikowanych oktav, jednak tak, że w obrębie fazy wolnej sygnał filtrowany jest silniej niż w obrębie fazy szybkiej. Przez „obręb fazy szybkiej” należy w tym przypadku rozumieć takie czasowe podzakresy funkcji detekcyjnej, w których przyjmuje ona wartości niezerowe. Podzakresy czasowe są oczywiście dodatkowo zmodyfikowane przez rozdzielczość czasową transformacji w danym zakresie częstotliwościowym (oktawie). Każdej z oktav  $j$  przyporządkowane są ponadto dwa stałe współczynniki

ki,  $a_j$  i  $b_j$ , o wartościach z przedziału  $\langle 0, 1 \rangle$ , będące proporcjonalnymi modyfikatorami funkcji detekcyjnej  $f_d(n)$  w tej oktawie podczas konstruowania wektora filtra  $M_{k,j}$  (rys. 5.5) zgodnie z zależnością

$$M_{k,j} = a_j \cdot f_d(n) + b_j \quad (5.2)$$

Parametry filtracji zostały tak dobrane, że składowe najwyższej oktawy są eliminowane niezależnie od położenia, a składowe oktaw 2, 3 i 4 są przenoszone tylko w obrębie fazy szybkiej oczopląsu ze współczynnikami odpowiednio 0,25, 1,0 i 0,75. Składowe niższych oktaw, a także składowa oktawy 4 ze współczynnikiem 0,25, są przenoszone niezależnie od położenia względem faz oczopląsu. Wektor współczynników czasowo-częstotliwościowych filtra oraz jego interpretacja na płaszczyźnie są przedstawione na rysunku 5.7

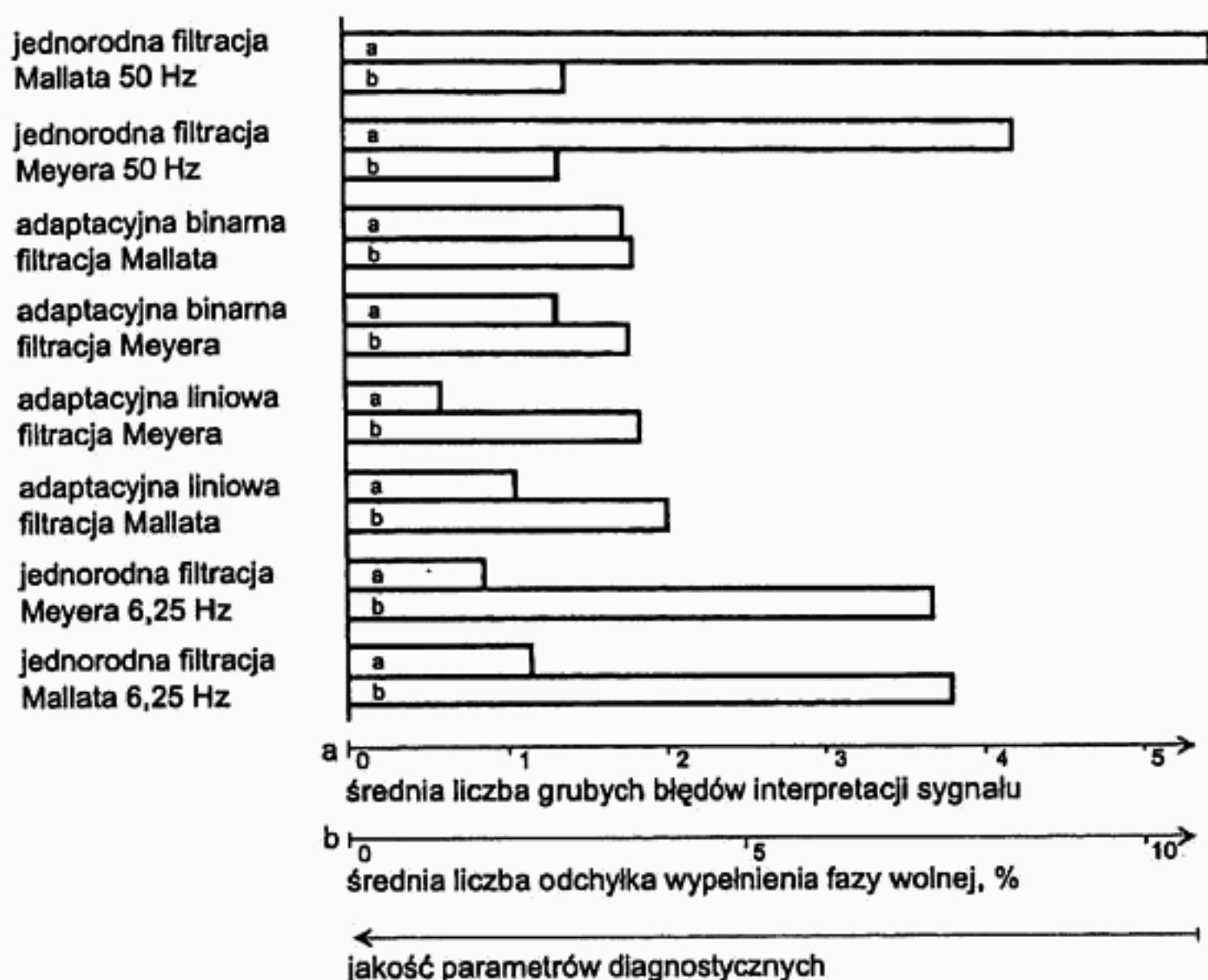


**Rys. 5.7.** Współczynniki czasowo-częstotliwościowe filtra dobranego adaptacyjnie do fragmentu sygnału oczopląsu: a) interpretacja na płaszczyźnie czasowo-częstotliwościowej; b) postać liniowa wektora filtra

Analiza rezultatów serii przeprowadzonych eksperymentów numerycznych i interpretacja uzyskanych wyników pozwoliły na określenie stopnia poprawy jakości parametrów diagnostycznych elektronystagmogramu filtrowanego w dziedzinie czasowo-częstotliwościowej w stosunku do tradycyjnych metod filtracji. Oceny poprawy jakości



parametrów diagnostycznych dokonano na drodze szczegółowej analizy tych parametrów (uzyskiwanych po filtracji sygnału) oraz przez porównywanie parametrów otrzymanych z użyciem testowanych metod filtracji z parametrami otrzymanymi w wyniku analizy sygnałów filtrowanych częstotliwościowo i sygnałów nie filtrowanych. Rysunek 5.8 jest graficzną prezentacją poprawy jakości obliczanych parametrów diagnostycznych dla rzeczywistych sygnałów oczopląsów. Zastosowanie adaptacyjnie dobieranych charakterystyk filtracji w dziedzinie czasowo-częstotliwościowej oraz dekompozycji z użyciem rodziny falek Meyera pozwoliło na uzyskanie jednocześnie najmniejszej liczby grubych pomyłek interpretacji i najmniejszej wartości odchyłki udziału fazy wolnej w sygnale w stosunku do sygnału oryginalnego.



Rys. 5.8. Ocena jakości obliczanych parametrów diagnostycznych dla rzeczywistych sygnałów oczopląsów

Najistotniejszym problemem stanowiącym w dalszym ciągu temat prac naukowych jest optymalizacja metody adaptacyjnego doboru charakterystyk filtracji czasowo-częstotliwościowej.

Filtracja sygnału elektronystagmograficznego w dziedzinie czasowo-częstotliwościowej o charakterystykach dobieranych adaptacyjnie jest metodą nową i obiecującą, ale zarazem taką, której możliwości nie zostały jeszcze w pełni poznane. Przeprowa-



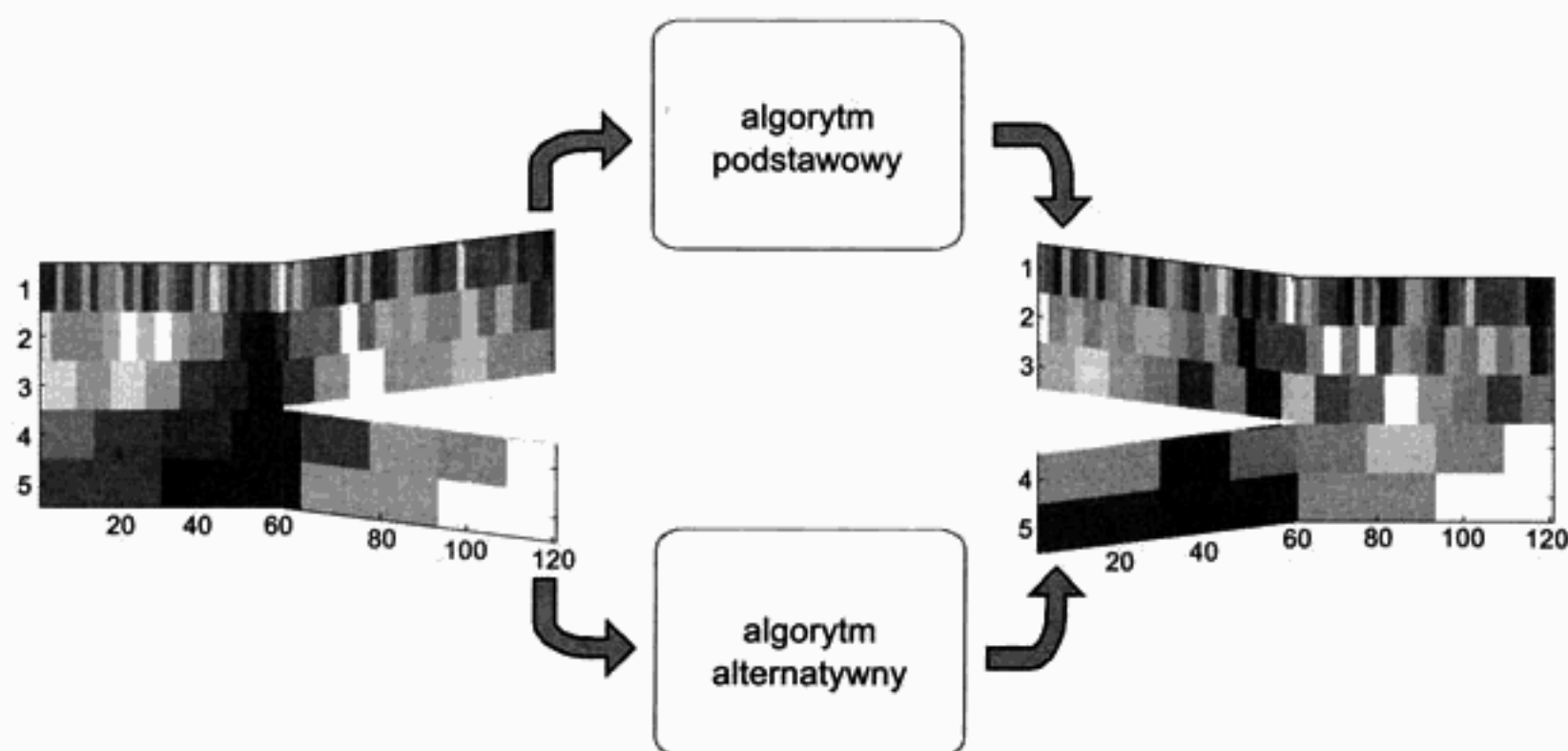
dzone eksperymenty pozwoliły rzucić nowe światło na omówione zagadnienie i otworzyły nowy obszar zastosowań nowoczesnych technik cyfrowego przetwarzania sygnałów w dziedzinach biomedycznych.

## 5.2. Ortogonalny podział pasma

Pasma sygnału zawarte pomiędzy składową stałą a częstotliwością maksymalną jest zwykle przetwarzane w sposób jednorodny, co oznacza, że każda próbka równoprawnie podlega przetwarzaniu z użyciem tego samego algorytmu. Nierzadko istnieje jednak potrzeba odmiennego przetworzenia pewnego zakresu częstotliwościowego lub zmiany metody przetwarzania synchronicznie z określonym zdarzeniem zewnętrznym. W przypadkach takich zastosowanie może znaleźć metoda ortogonalnego podziału pasma.

### 5.2.1. Wyodrębnianie podpasm częstotliwościowych

Zagadnienie wyodrębniania podpasm częstotliwościowych jest rozszerzeniem koncepcji częstotliwościowej filtracji sygnału. Zasadniczym elementem jest w obu przypadkach zwrotnica sygnału kierująca niektóre jego komponenty do wejścia podstawowego algorytmu przetwarzania, a inne do wejścia algorytmu alternatywnego (rys. 5.9). W swej najprostszej wersji filtracja zakłada dyskryminację określonych składowych sygnału, co odpowiada nieobecności algorytmu alternatywnego. Ortogonalny podział pasma posługuje się założeniem przeciwnym, a więc zapewnia wzajemną jednoznaczność oryginalnego ciągu próbek i ciągów próbek będących rezultatem dekompozycji sygnału.



Rys. 5.9. Schemat odmiennego przetwarzania sygnału w podpasmach częstotliwościowych

W swej podstawowej wersji ortogonalny podział pasma jest realizowany przez zestaw filtrów półpasmowych  $H(\omega)$  oraz  $G(\omega)$  połączonych zależnością wynikającą z warunku ortogonalności

$$G(\omega) = e^{-i\omega} \overline{H(\omega + \pi)} \quad (5.3)$$

gdzie kreska oznacza sprzężenie zespolone.

Rzeczywiste zastosowania stawiają jednak podziałowi pasma dalej posunięte wymagania:

- podział powinien odbywać się w dowolnej częstotliwości, niekoniecznie w połowie pasma;
- podział powinien umożliwiać wyodrębnienie podpasm nieciągłych, złożonych z kilku rozłącznych zakresów częstotliwości.

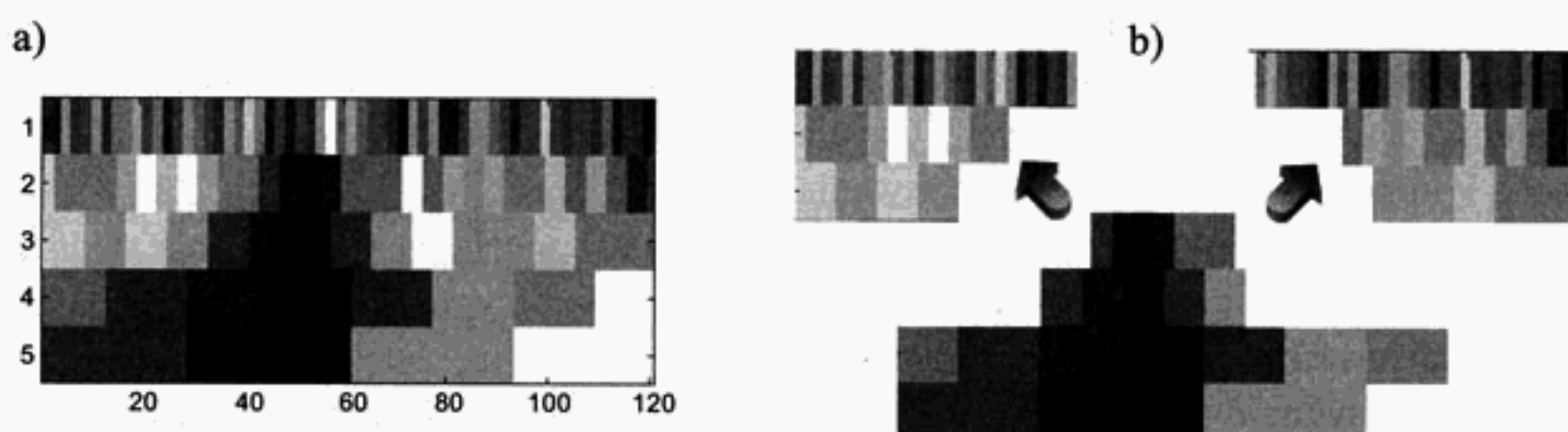
Pierwszą realizacją ortogonalnego podziału pasma spełniającą dodatkowe warunki jest modyfikacja widma sygnału w dziedzinie transformacji Fouriera. Podział pasma może następować przy dowolnej częstotliwości ograniczonej jedynie ziarnistością dyskretyzacji widma. Możliwe jest także grupowanie rozłącznych podpasm częstotliwości. Wreszcie, transformacja używa ortogonalnej bazy dekompozycji i zapewnia dokładną rekonstrukcję sygnału, co gwarantuje wzajemną jednoznaczność sygnału oryginalnego i sygnałów wynikowych. Użycie transformacji Fouriera pociąga jednak za sobą konieczność ostrożnego „wykrawania” podpasm częstotliwościowych. Za sprawą nieskończonych okresowych funkcji bazowych każda nieciągła modyfikacja widma spowoduje pojawienie się oscylacji w czasowej postaci sygnału, które, pomimo że skompensowane w komplementarnym sygnale wynikowym, zniekształcają przebieg każdego z sygnałów składowych. Dlatego podział pasma w dziedzinie Fouriera jest stosowany sporadycznie (np. jako jedna z metod projektowania filtrów), a używając funkcji o łagodnym przebiegu w celu stłumienia oscylacji, nie wykorzystuje precyzji, którą pozornie oferuje interwał dyskretyzacji widma.

Alternatywna metoda ortogonalnego podziału pasma, także spełniająca warunki spotykane w praktycznych zastosowaniach, polega na podziale czasowo-częstotliwościowej reprezentacji sygnału. Oczywiście dla zachowania spójności treści informacyjnej reprezentacja musi być uzyskana z użyciem ortogonalnej transformacji falkowej. Ortogonalny podział pasma w dziedzinie czasowo-częstotliwościowej oferuje dodatkowo możliwość zmiany przyporządkowania określonych komponentów do kanału przetwarzania w zależności od czynników zewnętrznych zdefiniowanych odrębną funkcją czasu synchroniczną z sygnałem. Przypadek, kiedy taka funkcja jest obliczana na podstawie sygnału, opisano w podrozdziale 5.1, choć może ona także reprezentować niezależny sygnał rejestrowany synchronicznie.

Ortogonalny podział pasma zrealizowany za pomocą transformacji falkowej charakteryzuje się precyzją zależną od zakresu częstotliwościowego, co jest skutkiem za-

sady nieoznaczoności. Przeciwnie do transformacji Fouriera użyto tutaj funkcji o nośniku zwartym i choć nie wyklucza to istnienia oscylacji na krańcach pasm częstotliwościowych (tzw. efekt brzegowy), pozostają one pod pełną kontrolą. Ponieważ ze względu na tłumienie oscylacji pożądane jest użycie falek o nośnikach krótszych niż szerokość wykrawanego pasma, praktyczna precyzja dokonanego podziału jest niższa niż wynikałoby to z zasady nieoznaczoności. Stosownie do potrzeb można wykorzystywać faleki o nośnikach krótszych, mających łagodniejsze zbocza charakterystyki widmowej, ale umożliwiającących szybsze przełączanie, lub dłuższych, gdy precyzja separacji częstotliwościowej jest zadaniem priorytetowym, a przełączanie – drugoplanowym lub wręcz nie wykorzystywanym.

Ortogonalny podział pasma w dziedzinie czasowo-częstotliwościowej polega na podziale płaszczyzny czasowo-częstotliwościowej na podpłaszczyzny należące do co najmniej dwóch grup (rys. 5.10). W najprostszym przypadku są to dwie podpłaszczyzny. Użycie większej liczby grup, jakkolwiek metodologicznie dopuszczalne, może być zawsze sprowadzone do podziału binarnego i dlatego nie będzie odrębnie opisywane. Bezstratność informacji podczas podziału płaszczyzny czasowo-częstotliwościowej jest zapewniona przez uczestnictwo każdego atomu w dokładnie jednej podpłaszczyźnie.



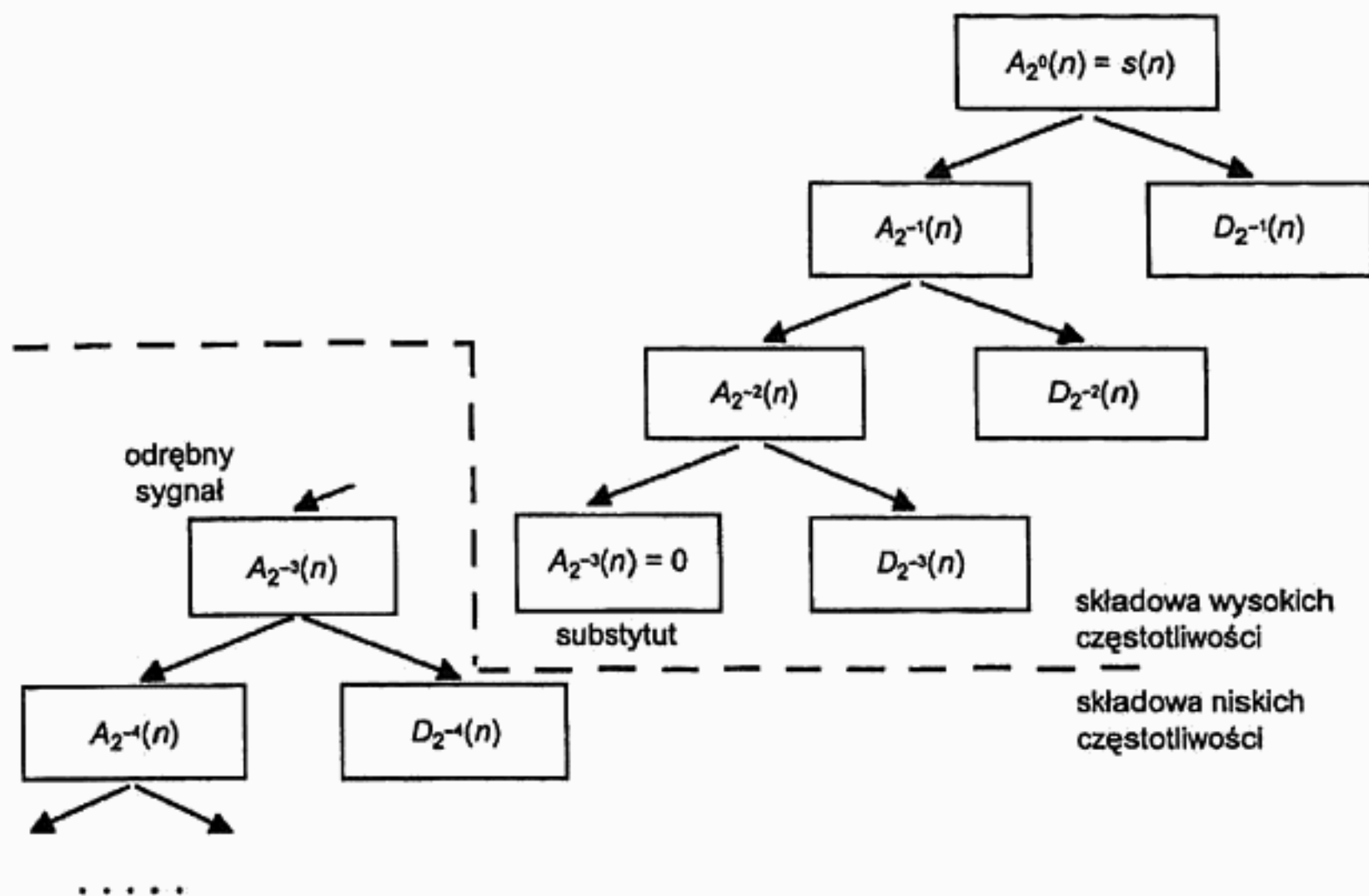
**Rys. 5.10.** Podział płaszczyzny czasowo-częstotliwościowej: a) czasowo-częstotliwościowa reprezentacja ewolucji serca; b) jej ortogonalny podział na trzy składowe

W przypadku gdy podział pasma jest dokonywany niezależnie od czasu (ang. *time-invariant split*) operacja podziału może być dokonana przez kształtowanie drzewa dekompozycji.

Operacja taka może przebiegać w dwóch wariantach, jako:

- 1) podział drzewa piramidowej dekompozycji falkowej – precyzja definicji częstotliwości podziału jest ograniczona do podwielokrotności częstotliwości próbkowania sygnału;
- 2) podział drzewa dekompozycji *Wavelet Packets* – precyzja definicji częstotliwości podziału jest ograniczona długością sygnału i zasadą nieoznaczoności – w tym przypadku nawet dla wysokich częstotliwości możliwe jest wyodrębnienie wąskiego pasma.

W obydwu przypadkach nie jest wymagana pełna dekompozycja sygnału, zwykle transformacja jest więc zatrzymywana po osiągnięciu poziomu wystarczającego dla podziału pasma. Podział drzewa dekompozycji jest przedstawiony na rysunku 5.11.

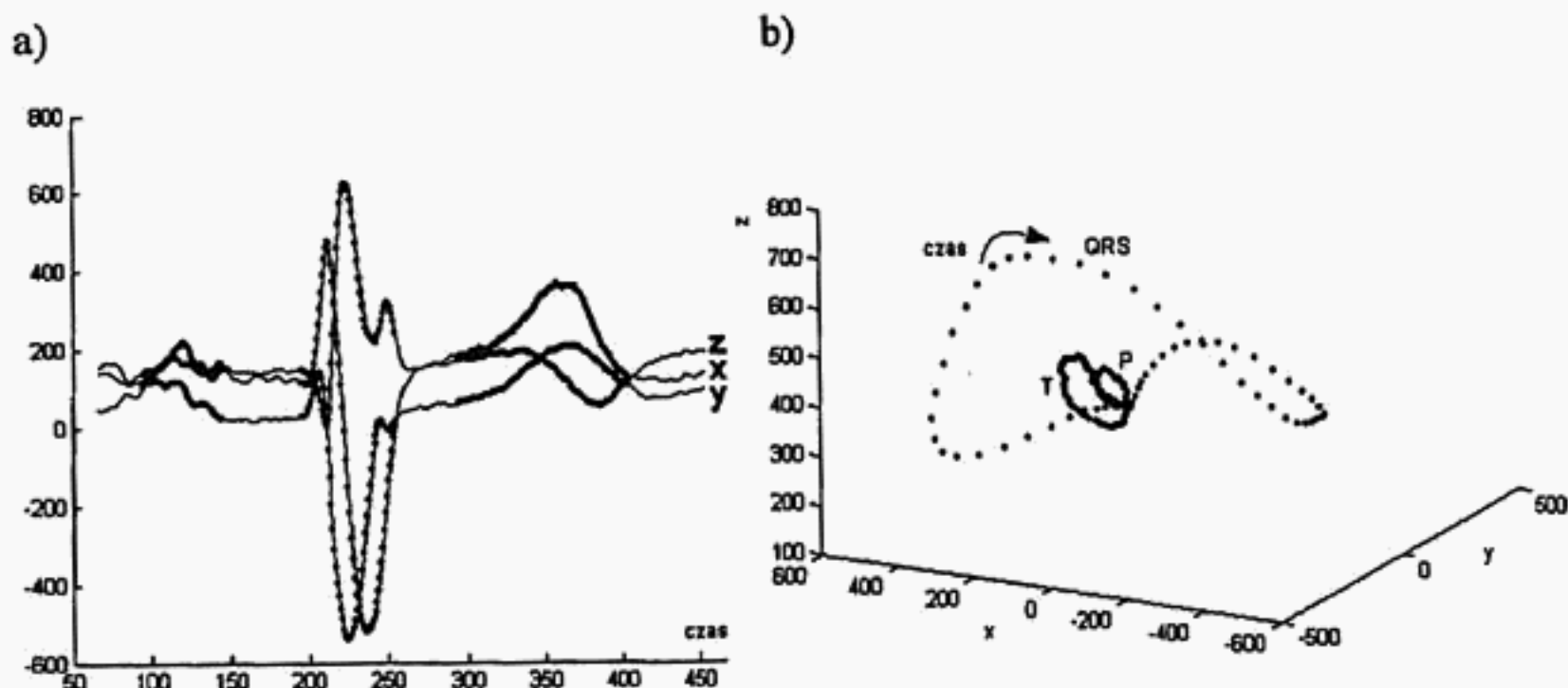


Rys. 5.11. Ortogonalny podział pasma zrealizowany przez podział drzewa dekompozycji

### 5.2.2. Zastosowanie podziału pasma w sekwencyjnej kompresji wektokardiogramów

Trójwymiarowa rejestracja pola elektrycznego pracującego serca WKG (wektokardiografia) umożliwia wizualizację przebiegu pola elektrycznego pracującego serca w przestrzeni dzięki specjalnemu układowi odprowadzeń (pseudoortogonalne odprowadzenia Franka). Wektokardiografia w swym podstawowym zakresie pozwala obserwować zjawiska dotąd trudne do diagnozowania (np. repolaryzacja przedsionków). Niestety, do rzutowania pętli zakreślonej przez wektor wypadkowy pola elektrycznego na płaszczyznę: czołową ( $X-Y$ ), poprzeczną ( $X-Z$ ) i strzałkową ( $Y-Z$ ), wymagany był specjalistyczny sprzęt i dopiero wykorzystanie technologii typowych dla grafiki komputerowej spowodował ożywiony rozwój wektokardiografii. Obecnie, właśnie dzięki cyfrowej akwizycji i automatycznej obróbce sygnałów, wektokardiografia przeżywa swą drugą młodość, a jej coraz powszechniejsze wdrażanie idzie w parze z napływającymi doniesieniami o nowych zastosowaniach diagnostycznych.





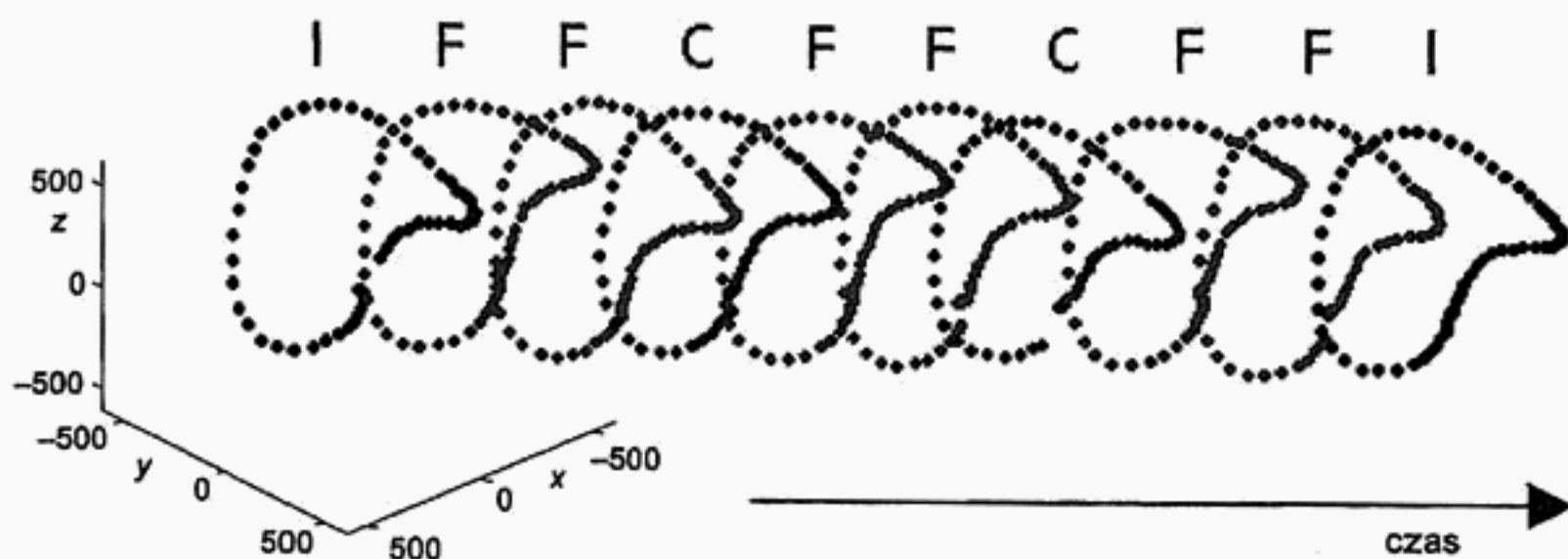
Rys. 5.12. Zapis wektokardiograficzny jednej ewolucji serca: a) zapis liniowy sygnałów ortogonalnych; b) zapis przestrzenny pętli wektokardiograficznych załamków P, QRS i T

Wektokardiogram jest rejestrowany w postaci trzech sygnałów  $X$ ,  $Y$  i  $Z$  odpowiadających ortogonalnej rejestracji dwubiegunowej pola elektrycznego (rys. 5.12). Sygnał liniowy każdego z kanałów może być przetwarzany identycznie do standardowego elektrokardiogramu (z wyjątkiem diagnostyki położenia osi serca, która w wektokardiografii jest wykonywana inaczej). Wektokardiografy spotykane są także w analizie sygnału próby wysiłkowej i jako rejestratory holterowskie wykonujące całodobowy zapis sygnału do pamięci cyfrowej. Szczególnie cennych informacji z diagnostycznego punktu widzenia dostarcza analiza zmian przestrzennego obrazu pętli załamka T (tzw. alternans załamka T).

Sekwencyjna kompresja wektokardiogramów polega na zastosowaniu metod właściwych do kompresji sekwencji równoodległych epizodów, gdzie spodziewane jest znaczne podobieństwo sąsiednich elementów. Metody takie są szeroko wykorzystywane w kodowaniu strumienia danych wizyjnych. W przypadku wektokardiogramu będącego zapisem przeważającej ilości ewolucji serca pochodzenia zatokowego, również spodziewane jest znaczne podobieństwo przebiegu pola elektrycznego epizodów sąsiednich. Wykorzystanie tej własności doprowadziło do opracowania algorytmu kompresji dedykowanego dla wektokardiografii, możliwego do wykorzystania również przez elektrokardiografię konwencjonalną, w którym pętle zataczone przez pole elektryczne pracującego serca podczas skurczu komór (tzw. zespoły QRS) tworzą sekwencję „obrazów”. Analogicznie traktowane są załamki P odpowiadające skurczowi przedsionków oraz załamki T reprezentujące proces repolaryzacji komór. Podobnie do algorytmów MPEG, wyróżnione elementy sekwencji ( $I$  – rys. 5.13) kodowane są bezpośrednio (w przypadku zaburzeń rytmu są to skrajne ewolucje odcinków rytmu regularnego), zaś pozostałe są kodowane z użyciem interpolacji jedno- lub dwukierunkowej (odpowiednio  $F$  oraz  $C$ ). Z wyjątkiem pobudzeń nadmiarowych, czasowa zmienność



aktywności serca jest bardzo powolna, toteż przewidywania przebiegu pętli interpolowanych na podstawie sąsiednich ewolucji są zazwyczaj tylko nieznacznie rozbieżne z ich rzeczywistym przebiegiem. Dzięki niewielkiej liczbie efektywnie wykorzystywanych symboli kodowanie błędów interpolacji pozwala na osiągnięcie wysokiej efektywności kompresji.

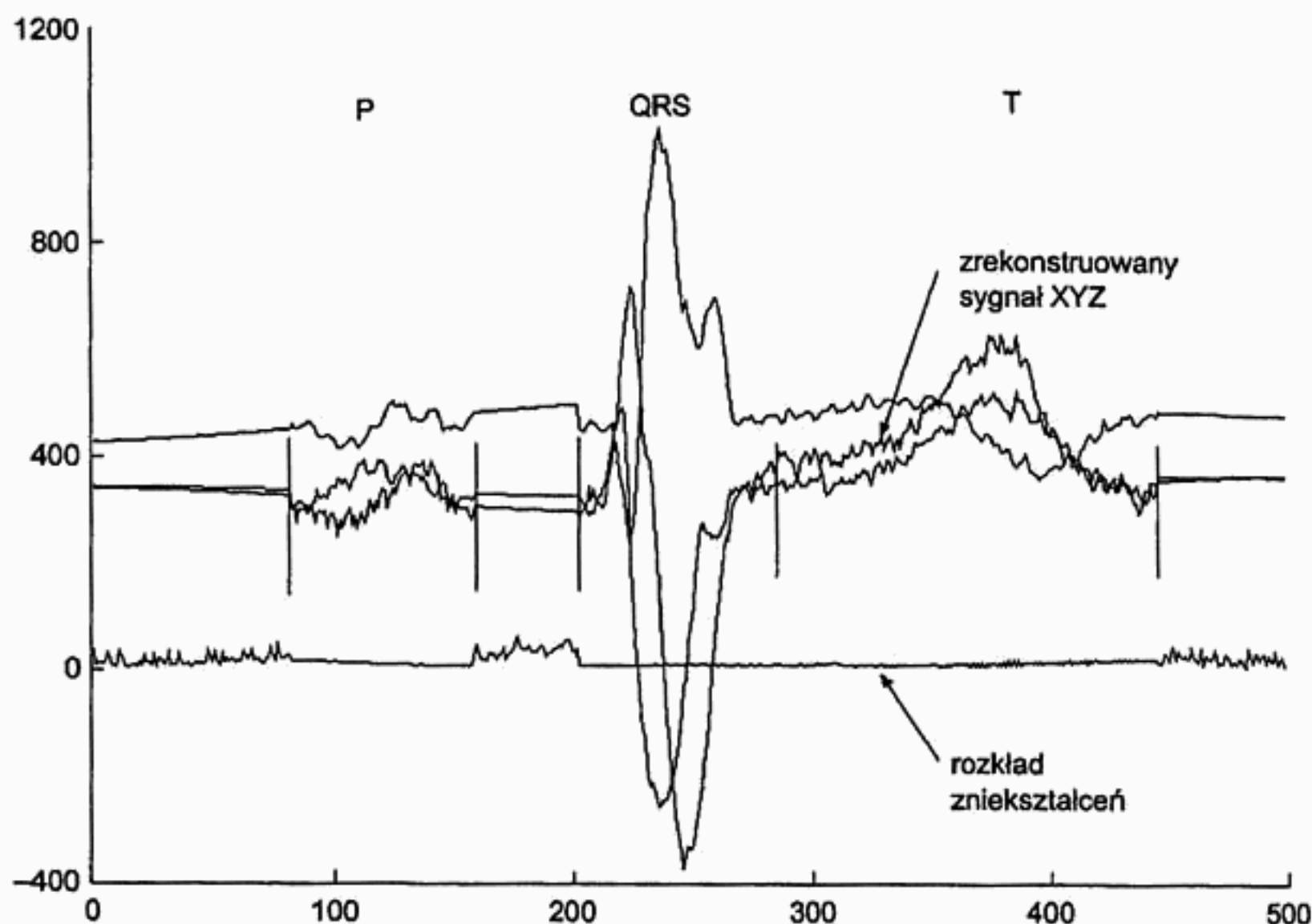


**Rys. 5.13.** Pętle wektokardiograficzne zespołów QRS sąsiednich ewolucji serca jako elementy sekwencji – metoda kodowania i predykcji sąsiadujących pętli (*I* – bezpośrednie kodowanie, *C* – zgrubna predykcja, *F* – dokładna predykcja)

W zastosowaniu do rzeczywistych zapisów WKG pozyskiwanych w różnych warunkach należy się liczyć z nałożeniem się elektrycznej aktywności serca i różnych zjawisk elektrycznych. Przykładem może być wolnozmiennie wahanie linii izoelektrycznej będące następstwem niestałości kontaktu elektrod pomiarowych, a więc nieistotne z diagnostycznego punktu widzenia. Degraduje ono trafność przewidywań interpolacyjnych, a więc i własności statystyczne błędów podlegających kodowaniu. Dodatkowo, spełnienie postulatu kompresji bezstratnej wymaga odtworzenia wahań linii izoelektrycznej w sygnale zrekonstruowanym po kompresji.

Skuteczne ograniczenie wpływu wolnozmiennych wahań linii izoelektrycznej na efektywność kompresji zostało zrealizowane za pomocą ortogonalnego podziału pasma (rys. 5.14). Jako częstotliwość podziału wybrano wartość 2 Hz, przy której uzyskano dobrą przewidywalność pętli, a jednocześnie stosunkowo niski strumień danych związany z sygnałem reprezentującym składowe wolnozmiennie (efektywna częstotliwość próbkowania 4 Hz). Podział pasma na dwie wzajemnie komplementarne części odbywa się niezależnie od czasu i został zrealizowany za pomocą obcięcia drzewa dekompozycji falkowej na poziomie 6 (częstotliwość próbkowania sygnału wynosiła 500 Hz). Składowe sygnału o częstotliwościach 0÷2 Hz zostały zakodowane odrębnie jako ciągły sygnał z użyciem algorytmu przyrostowego. Pozostałe składowe w przedziale częstotliwości 2÷250 Hz są kodowane jako trzy współbieżne sekwencje pętli (odrębnie dla zespołów QRS, załamków P i załamków T), w których co piętnasta

ewolucja jest kodowana bezpośrednio, a pozostałe – z użyciem interpolacji. W przypadku zaburzeń rytmu sekwencja kodowana z użyciem interpolacji jest skracana, a w przypadku niektórych arytmii (bigeminia, trigeminia) kodowanie interpolacyjne nie jest używane.



Rys. 5.14. Sygnał WKG zrekonstruowany z kompresji metodą sekwencyjną. Zniekształcenia pomiędzy załamkami są konsekwencją ograniczenia pasma sygnału do 2 Hz

## ROZDZIAŁ 6

# Kompresja sygnałów biomedycznych z wykorzystaniem reprezentacji czasowo-częstotliwościowych

Powszechna dostępność testów elektrodiagnostycznych, podobnie jak innych usług medycznych, stanowi o poziomie rozwoju cywilizacji. Niezależnie od gwałtownego rozwoju technologii przechowywania i przesyłania danych problem optymalnej reprezentacji sygnałów elektrofizjologicznych niezmiennie pozostaje aktualny zwłaszcza w takich dziedzinach, jak: badania długookresowe (holterowskie), teleakwizycja oraz zarządzanie elektroniczną dokumentacją pacjentów. Problem kompresji takich sygnałów wiąże się z dwoma zasadniczymi zagadnieniami: redukcją objętości pliku oraz wiernością reprezentacji zjawisk biofizycznych będących podstawą diagnozowania. Istniejące bezstratne metody kodowania danych charakteryzują się niewielką efektywnością w stosunku do sygnałów elektrodiagnostycznych, a na dodatek bezstratność, rozumiana jako identyczność oryginalnej i wynikowej cyfrowej reprezentacji sygnału, nie gwarantuje pozyskiwania tej reprezentacji w sposób obojętny dla informacji diagnostycznej.

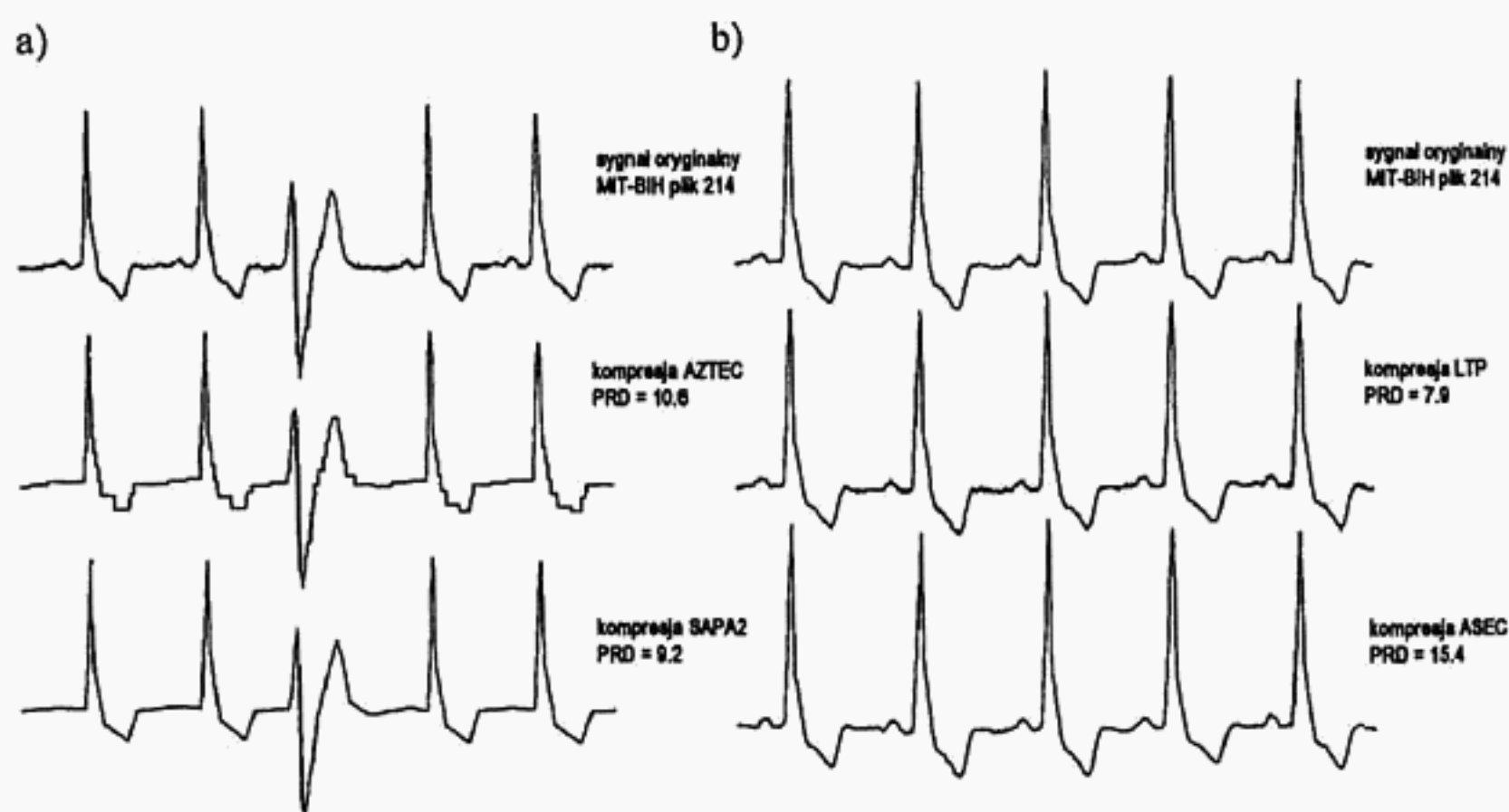
Przykładowo, elektrokardiogram jest pozyskiwany od pacjenta bez żadnych założeń wstępnych, z czym wiąże się obowiązek zapewnienia maksymalnej dokładności odwzorowania sygnału w jego reprezentacji cyfrowej przez cały czas, jako że rzadko występujące komponenty szerokopasmowe mogą pojawić się w dowolnej chwili. Zapis elektrycznej aktywności serca, podobnie jak i inne sygnały elektrofizjologiczne, charakteryzuje się niejednorodnością dystrybucji informacji diagnostycznych w czasie. W aspekcie medycznym poszczególne fragmenty zapisu postrzegane są jako bardziej istotne od pozostałych, natomiast w aspekcie inżynierskim lokalna szerokość pasma sygnału podlega znacznym wahaniom. Ponadto, szczęśliwie w przypadku elektrokardiogramu i niektórych innych sygnałów biomedycznych, moment wystąpienia komponentów szerokopasmowych nie jest całkowicie dowolny, wynika bowiem z ograniczeń narzuconych przez fizjologię. Może on być wyznaczony z zadowalającym poziomem ufności na drodze automatycznej analizy sygnału.

Tradycyjna klasyfikacja metod kompresji cyfrowej reprezentacji sygnału wyróżnia metody stratne i bezstratne na podstawie porównania ciągu próbek będącego rezultatem dekompresji z oryginalnym ciągiem próbek:

- w przypadku kompresji bezstratnej są to dwa identyczne ciągi próbek;
- w przypadku kompresji stratnej są to dwa różne ciągi próbek, a różnica między nimi jest powszechnie przyjętą miarą jakości kompresji.

Ścisłość tej definicji stwarza pozory gwarancji idealnej reprezentacji sygnału w przypadku stosowania metod bezstratnych, należy jednak pamiętać, że kompresja dotyczy dyskretnej reprezentacji sygnału, która jedynie w sposób przybliżony reprezentuje sygnał rzeczywisty. Tak więc straty informacji, nieuchronnie związane z kwantyzacją i dyskretyzacją sygnałów elektrodiagnostycznych, występują już na etapie akwizycji. Można wykazać, że wraz ze wzrostem dokładności dyskretnej reprezentacji sygnału (częstotliwość próbkowania i rozdzielczość kwantyzacji) własności statystyczne otrzymanego ciągu próbek, będące podstawą bezstratnych metod kompresji, pogarszają się, zmniejszając jej efektywność. W przypadku granicznym, tj. dla sygnału analogowego reprezentowanego przez dowolne wartości rzeczywiste położone nieskończenie blisko, nie można w ogóle wskazać wartości bardziej prawdopodobnej. Przeciwnie, jeśli dyskretna reprezentacja sygnału została pobrana przy założeniu znacznego uproszczenia sygnału (duża ziarnistość kwantyzacji) – rozkład wartości próbek może być bardzo niejednorodny i wystąpienie niektórych wartości będzie znacznie bardziej prawdopodobne niż pozostałych. Jest to wprawdzie podstawą do znakomitych rezultatów kompresji, bezstratnej w myśl definicji przedstawionej poprzednio, ale zastosowanie diagnostyczne takiej reprezentacji powinno budzić niepokój.

Ocena rezultatów kompresji powinna obejmować również proces kwantyzacji sygnału analogowego, gdyż sposób pozyskania dyskretnej reprezentacji sygnału ma istotne znaczenie dla oszacowania przydatności diagnostycznej sygnału (rys. 6.1).



**Rys. 6.1.** Ocena zniekształceń sygnału za pomocą różnicy średniokwadratowej PRD: a) sygnały istotnie różne z medycznego punktu widzenia o niewielkiej wartości PRD; b) sygnały identyczne z medycznego punktu widzenia o znacznej wartości PRD



W tym kontekście powszechne stosowanie w charakterze estymatora różnicy dwóch sygnałów wartości, tzw. różnicy średniokwadratowej PRD (ang. *percent root-mean-square difference*), ustępuje miejsca maksymalnej różnicy wyrażonej w mikrowoltach ( $\mu V$ ).

$$PRD = \left\{ \frac{\sum_{i=1}^n [x_1(i) - x_2(i)]^2}{\sum_{i=1}^n [x_1(i)]^2} \right\} \cdot 100\% \quad (6.1)$$

Podobna sytuacja ma miejsce przy ocenie efektywności kompresji, która w znacznej mierze zależy od kwantyzacji i częstotliwości próbkowania. Stwierdzenie, czy trzykrotna redukcja objętości pliku danych jest lepsza niż dwukrotna, nie jest jednoznacznie możliwe, jeżeli parametry próbkowania sygnałów, na których testowane były konkurencyjne algorytmy, są różne. Zamiast więc podawać stopień redukcji ilości informacji osiągany przy użyciu konkretnej metody przeważnie podaje się natężenie wyjściowego strumienia informacji (ang. *bitrate*) w kb/s. Wartość ta ma dodatkowo bardzo konkretne znaczenie w zastosowaniach transmisji sygnałów elektrodiagnostycznych, określa bowiem bezpośrednio wymaganą przepustowość (pasmo) kanału transmisyjnego.

Wyrazem braku zaufania środowisk medycznych do dotychczasowych osiągnięć w dziedzinie kompresji sygnałów elektrodiagnostycznych jest prawny zakaz używania stratnych metod kompresji do sygnałów będących podstawą diagnozowania w wielu krajach, w tym USA, Unii Europejskiej i w Polsce.

## 6.1. Stratne i bezstratne kodowanie reprezentacji czasowo-częstotliwościowej

Kompresja sygnału w dziedzinie czasu wykorzystuje szereg znanych technik stratnych takich, jak algorytm punktu zwrotnego (ang. *turning point*), AZTEC, CORTES, „fan” i inne) oraz bezstratnych (kodowanie Huffmana i przyrostowe). Warto zauważyć, że np. algorytm punktu zwrotnego był w założeniu przeznaczony do zastosowań w elektrokardiografii. Techniki bezstratne są oparte na wykorzystaniu niejednorodności rozkładu wartości próbek (kodowanie Huffmana) lub spodziewanej korelacji kolejnych wartości takich, jak kodowanie przyrostowe. W dziedzinie czasowo-częstotliwościowej sygnał jest reprezentowany przez taką samą liczbę współczynników jak w dziedzinie czasu (przy założeniu dekompozycji z użyciem ortogonalnej dyskretniej transformacji falkowej), jednak własności statystyczne ich wartości są zdecydowanie



odmienne. Już pobieżna obserwacja płaszczyzny czasowo-częstotliwościowej pozwala stwierdzić, że dzięki diadycznej strukturze drzewa dekompozycji przeważająca ilość współczynników reprezentuje górne pasma sygnału, gdy tymczasem energia większości sygnałów elektrodiagnostycznych skoncentrowana jest w dolnych zakresach częstotliwościowych. Zależność tę można wykorzystać do bezstratnego kodowania czasowo-częstotliwościowej reprezentacji sygnału metodą Huffmana. Wykorzystanie kodowania przyrostowego nie jest już takie oczywiste, bowiem korelacja kolejnych próbek w poszczególnych pasmach częstotliwościowych została zminimalizowana przez użycie ortogonalnych falek analizujących podczas dekompozycji sygnału.

### **6.1.1. Eliminacja progowa współczynników czasowo-częstotliwościowych**

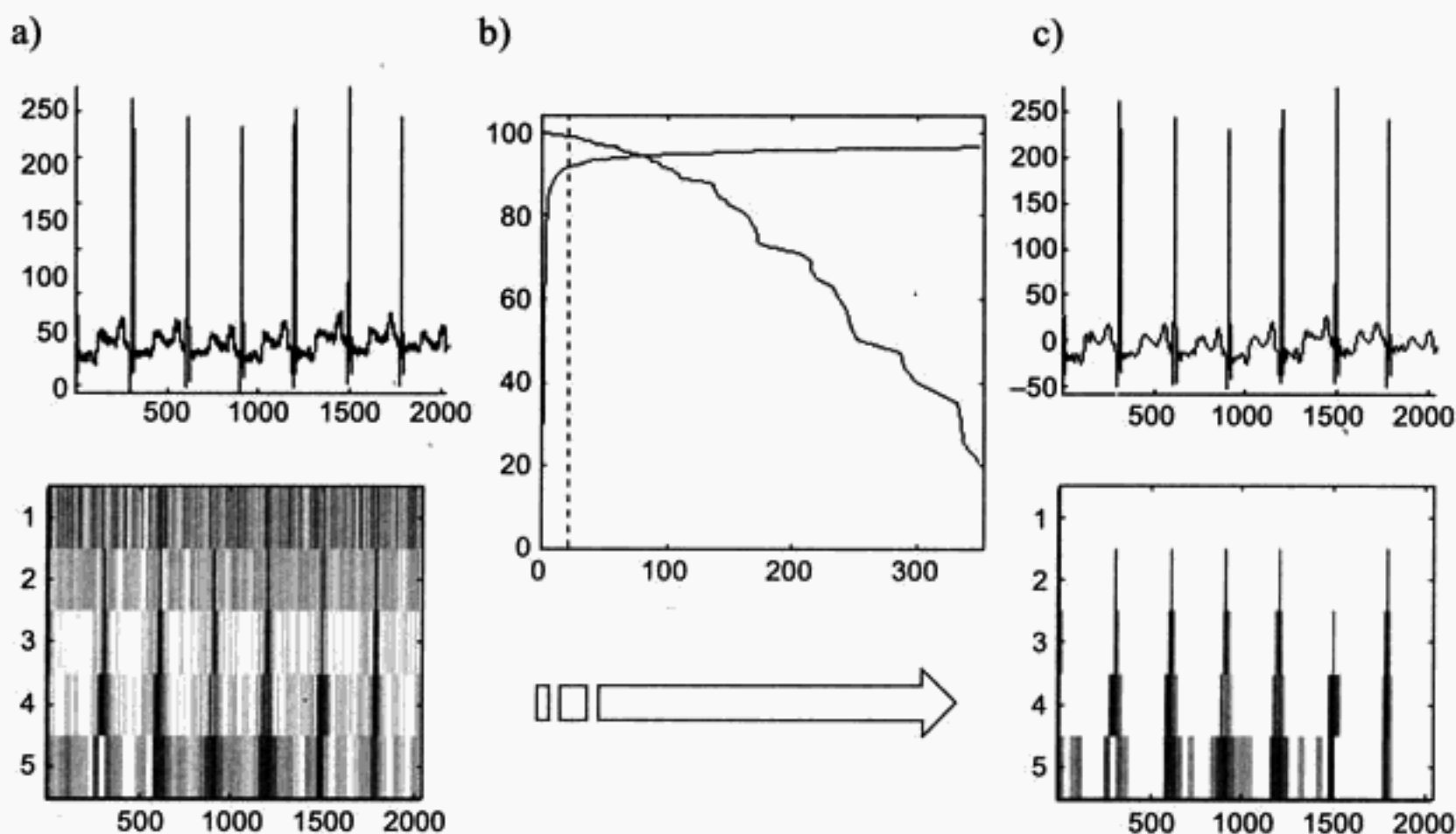
Dopuszczając możliwość niewielkiej odchyłki treści informacyjnej sygnału wynikowego (kompresja stratna), można dokonać eliminacji współczynników czasowo-częstotliwościowych nieznacznie różnych od zera. Korzystając z faktu, że używana do oceny jakości kompresji stratnej wielkość PRD reprezentuje energię, można założyć współczynnik zniekształceń na określonym poziomie globalnej energii sygnału (np.  $1 \div 5\%$ ). Następnie począwszy od współczynników powierzchni czasowo-częstotliwościowej o najmniejszych wartościach bezwzględnych wystarczy zastępować te wartości zerami, jednocześnie kumulując wartości wyeliminowane w zmiennej pomocniczej. Przekroczenie przez tę zmienną założonej wcześniej wartości oznacza koniec procesu eliminacji. W zależności od sygnału udaje się tą drogą wyeliminować  $50 \div 85\%$  współczynników przy zniekształceniach nie przekraczających 5%.

Metoda ta w prosty sposób uświadamia kompromis pomiędzy efektywnością kompresji a zniekształceniami sygnału, który zawsze towarzyszy stratnym metodom kompresji.

Metoda ta może również posłużyć do wykazania wpływu kształtu falki-matki na rezultaty kompresji. Im bardziej kształt ten przypomina fragmenty sygnału typowe np. dla elektrokardiogramu, tym mniejsze jest rozproszenie energii reprezentacji na współczynniki powierzchni czasowo-częstotliwościowej. Uzyskujemy zatem zaledwie kilka współczynników o znacznych wartościach, które w sposób niemal kompletny reprezentują sygnał, podczas gdy znaczna większość pozostałych ma wartości nieznaczące i może zostać pominięta bez ryzyka znacznych zniekształceń sygnału (rys. 6.2). W przypadku granicznym, w praktyce spotykanym jednak tylko przy użyciu sygnałów testowych, cała zawartość sygnału może być reprezentowana przez jeden współczynnik falkowy (wykluczając jednak trywialny przypadek sygnału stałego o zerowej energii).

Z opisanych wyżej powodów w zastosowaniach praktycznych rzadko wykorzystywane są falki niższych rzędów, gdyż ich nieregularny przebieg w dziedzinie czasu „nie

pasuje” do sygnałów elektrodiagnostycznych, stąd konieczność ich reprezentacji za pomocą wielu współczynników o istotnych wartościach. Przeciwnie, falki wyższych rzędów dzięki gładkiemu przebiegowi w dziedzinie czasu lepiej korelują z sygnałami pochodzenia biologicznego będącymi reprezentacją naturalnych procesów. Dlatego pełna reprezentacja tych sygnałów wymaga w takim przypadku mniejszej liczby współczynników, stwarzając nadzieję na wyższą efektywność kompresji.



**Rys. 6.2.** Eliminacja progowa współczynników czasowo-częstotliwościowych (uzyskano współczynnik kompresji 10,9 przy zniekształceniach  $PRD = 1\%$ ): a) sygnał oryginalny; b) współzależność zachowanej energii sygnału i ilości zerowanych współczynników reprezentacji, w %; c) sygnał zrekonstruowany po kompresji

Pewnym zastrzeżeniem jest sposób interpretacji współczynnika zniekształceń, który zakłada jednakową istotność dowolnej porcji sygnału (rys. 6.1). W przypadku sygnałów elektrodiagnostycznych zwłaszcza np. elektrokardiogramu, za sprawą niejednorodnej gęstości informacji diagnostycznych, określony poziom zniekształceń akceptowany w określonym interwale czasowym nie może być akceptowany w innej porcji sygnału.

### 6.1.2. Kodowanie współczynników czasowo-częstotliwościowych

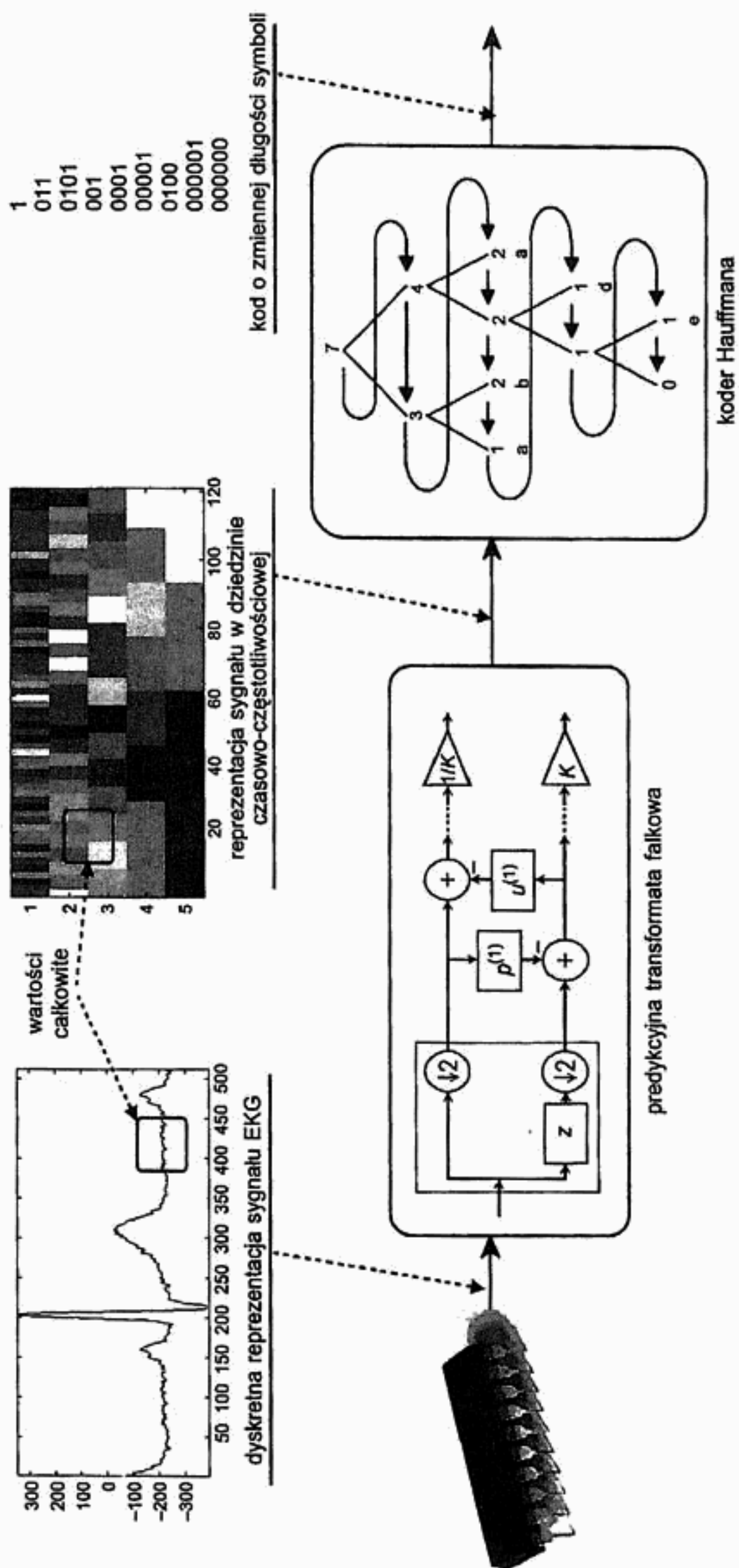
Kodowanie próbek metodą Huffmana polega na przyporządkowaniu unikalnego symbolu dla każdej wartości ze skończonego zbioru wartości (a z takim mamy do czynienia w przypadku dyskretnej reprezentacji sygnału). Przyporządkowanie to dokony-

wane jest z uwzględnieniem prawdopodobieństwa wystąpienia danej wartości, musi być zatem poprzedzone analizą statystyczną ciągu wejściowego. Symbole służące do zakodowania wartości próbek mają różną długość, stąd dla rozgraniczenia poszczególnych symboli konieczne jest stosowanie kodu przedrostkowego. Oznacza to, że żadne słowo kodu nie jest przedrostkiem żadnego innego słowa kodu. Gwarantuje to możliwość jednoznacznego zdekodowania zakodowanych wartości, czyli przyporządkowania symbolom kodu oryginalnych wartości próbek sygnału.

Kodowanie Huffmana może być wykorzystane do kompresji danych, w tym także dyskretnej reprezentacji sygnałów. W tym zastosowaniu jako zasadę kodowania przyjęto przyporządkowanie częściej występującym symbolom krótszych kodów, pozostawiając dłuższe dla wartości występujących sporadycznie. Wprawdzie liczba próbek reprezentacji kodowej jest identyczna, ale dzięki zastosowaniu symboli o różnej długości sumaryczna liczba bitów reprezentacji kodowej jest mniejsza niż w przypadku oryginalnej reprezentacji sygnału. W reprezentacji kodowej przeważają bowiem symbole krótsze.

Kodowanie Huffmana bywa często stosowane do bezstratnej kompresji sygnałów w dziedzinie czasu. Jednakże wspomniana już wcześniej zależność efektywności kompresji od rozkładu wartości ciągu próbek oraz atrakcyjne własności statystyczne czasowo-częstotliwościowej reprezentacji sygnału otwierają nowe możliwości bezstratnej kompresji sygnałów zrealizowanej według schematu przedstawionego na rysunku 6.3.

Podkreślając, że rozważamy kompresję bezstratną, więcej uwagi należy poświęcić połączeniu transformacji falkowej i kodera Huffmana. Oczywiście jest, że od zastosowanej transformacji oczekiwać należy spełnienia warunku dokładnej rekonstrukcji. Niestety, tradycyjne transformacje posiadające tę własność wykorzystujące np. banki filtrów wykorzystują falki rzeczywiste (niekiedy nawet zespolone), a więc reprezentacja czasowo-częstotliwościowa sygnału jest również rzeczywista. Oznacza to, że zbiór wartości tej reprezentacji może zawierać dowolnie wiele elementów, co przeczy podstawowemu założeniu metody Huffmana opartej na rozkładzie w skończonym zbiorze wartości. Ponieważ kwantyzacja współczynników powierzchni czasowo-częstotliwościowej (np. poprzez zaokrąglenie do najbliższej wartości całkowitej) prowadzi do utraty informacji, w tym szczególnym zastosowaniu jej użycie jest wykluczone. Pozostaje jedynie wykorzystanie transformacji czasowo-częstotliwościowej, która ponad spełnienie warunku dokładnej rekonstrukcji wytwarza reprezentację sygnału w postaci całkowitoliczbowej. Warunek ten spełnia liftingowa transformacja falkowa, będąca w istocie alternatywnym w stosunku do filtracji algorytmem dekompozycji sygnału. Chociaż rolę falki analizującej i funkcji skalującej przejmują iteracyjne stosowanie predykcji i uaktualnień, podobnie jak w klasycznej transformacji falkowej, istnieją metody kształtowania tych dwóch procedur w celu otrzymania optymalnej, w sensie koncentracji energii, reprezentacji czasowo-częstotliwościowej.



Rys. 6.3. Schemat bezstratnego kodowania współczynników czasowo-częstotliwościowej reprezentacji sygnału



Projektowanie filtrów dla liftingowej transformacji falkowej w kontekście szczególnych zastosowań jest nową i intensywnie obecnie rozwijaną dziedziną przetwarzania sygnałów i, niestety, wykracza poza ramy niniejszego opracowania.

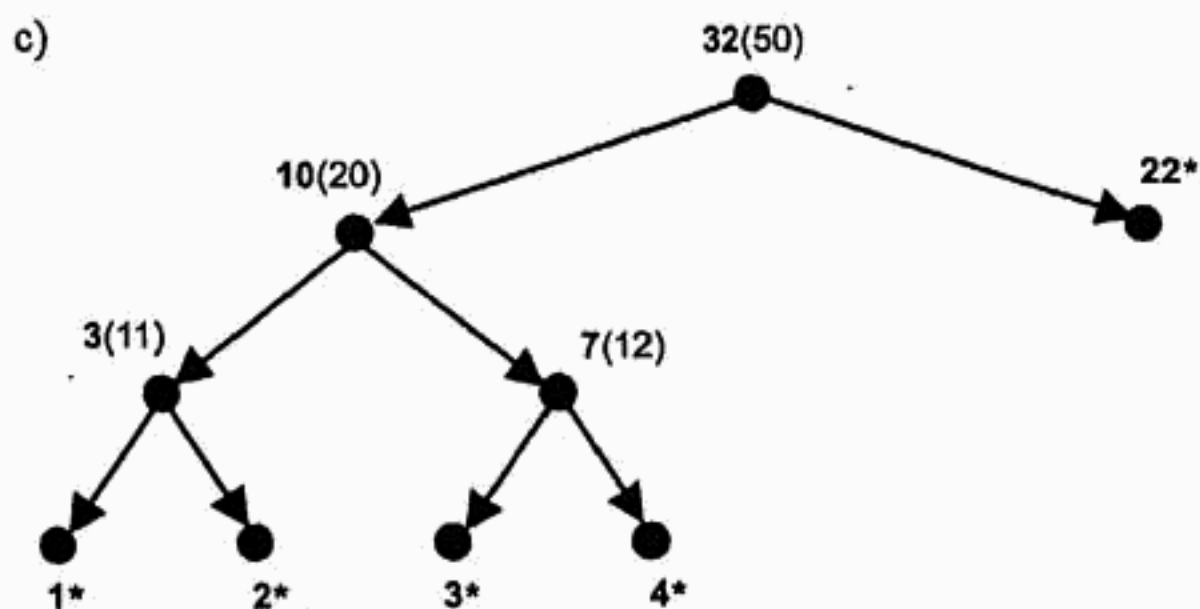
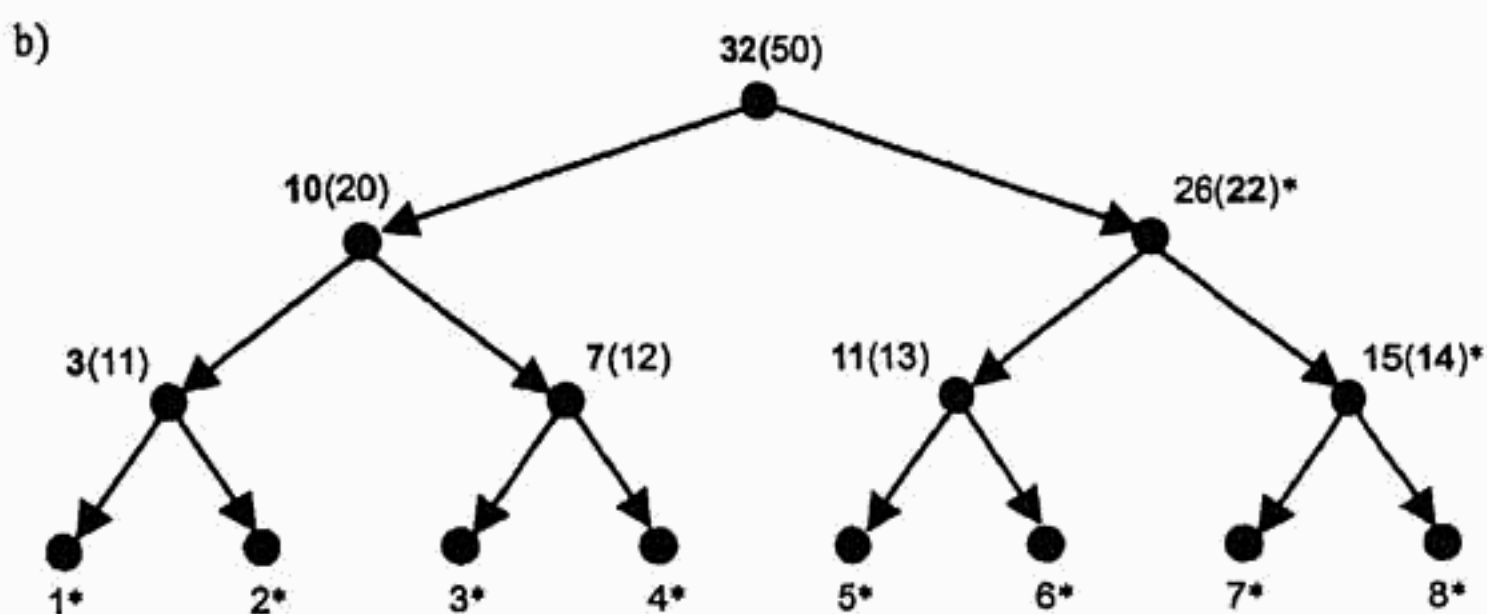
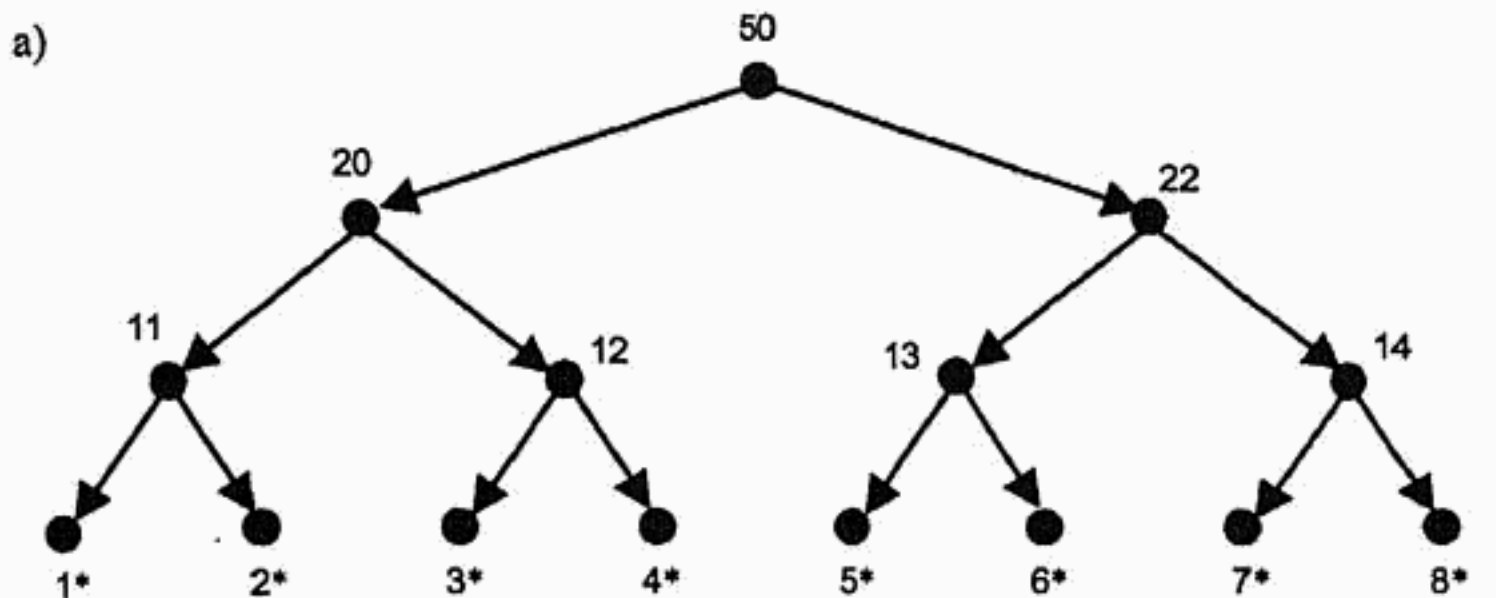
Dotychczasowe rozważania prowadzą do wniosku, że w celu zapewnienia dokładnej rekonstrukcji kodowanie współczynników powierzchni czasowo-częstotliwościowej musi być poprzedzone całkowitoliczbową transformacją sygnału. Efektywność kompresji wykorzystującej specjalnie zaprojektowane filtry pozwala oczekiwać 3,2-krotnej redukcji objętości pliku wynikowego (baza MIT-BIH, częstotliwość próbkowania 360 Hz, kwantyzacja 12 bitów). Wartość współczynnika kompresji jest znaczna jak na metodę bezstratną i przewyższa efektywność wielu wcześniejszych metod kompresji (w tym wszystkich algorytmów ogólnego przeznaczenia) testowanych na tym samym zestawie zapisów elektrokardiograficznych.

## 6.2. Kompresja optymalnego drzewa dekompozycji *Wavelet Packets*

Kompletna dekompozycja sygnału za pomocą pakietów falkowych (ang. *Wavelet Packets*) prowadzi do uzyskania współczynników reprezentujących sygnał w dyskretniej dziedzinie częstotliwości. Jednakże przeciwnie do klasycznej dekompozycji falkowej wykonywanej zawsze w oparciu o tę samą strukturę drzewa binarnego, wybór pomiędzy oryginalną a zdekomponowaną reprezentacją sygnału jest dokonywany dla każdego węzła drzewa niezależnie. Rezultatem jest adaptacja struktury dekompozycji do analizowanego sygnału, a struktura wynikowa nosi nazwę optymalnego drzewa dekompozycji. Oczywiście optymalizacja wymaga sformułowania kryterium decyzyjnego określającego, która z wersji reprezentacji jest preferowana. Kryterium tym zazwyczaj jest entropia.

Aby otrzymać optymalne drzewo dekompozycji, dokonywana jest pełna dekompozycja porcji sygnału z użyciem pakietów falkowych, a dla każdego z węzłów dekompozycji (włącznie z ostatnimi) obliczana jest wartość entropii. Celem optymalizacji drzewa dekompozycji jest zmniejszenie sumy entropii najniższych węzłów. Jeżeli suma entropii dwóch węzłów na poziomie niższym (dzieci) jest większa niż entropia połączonego z nimi węzła na poziomie wyższym (ojca), to z punktu widzenia minimalizacji entropii korzystniejsze jest pozostawienie reprezentacji w tym węźle w postaci nie zdekomponowanej. W przeciwnym razie węzłowi na poziomie wyższym przyporządkowana jest nowa (niższa) wartość entropii równa sumie entropii połączonych z nim węzłów na poziomie niższym, a dekompozycja zostaje zachowana. Iteracyjne powtarzanie tego postępowania dla wszystkich węzłów drzewa doprowadzi do usunięcia niektórych z nich, a pozostała struktura będzie zaadaptowana do reprezentowanej porcji sygnału (rys. 6.4).





Rys. 6.4. Optimalizacja drzewa dekompozycji *Wavelet Packets*: a) pełne drzewo dekompozycji (liczba oznacza entropię węzła wyznaczoną z reprezentacji); b) wyliczenie sumy entropii węzłów potomnych (przed nawiasem gwiazdką oznaczono węzły, w których suma entropii węzłów potomnych przewyższa entropię węzła); c) optymalne drzewo dekompozycji

### 6.2.1. Zastosowanie pakietów falkowych do bezstratnej kompresji elektrokardiogramu

Ponieważ poszukiwania optymalnego drzewa dekompozycji pozwalają na użycie dowolnego kryterium optymalizacji, w przypadku pakietów falkowych zastosowanych do bezstratnej kompresji elektrokardiogramu celem jest minimalizacja liczby bitów niezbędnych do zakodowania informacji. Operacje dla pojedynczego węzła dekompozycji na poziomie  $j$  można więc przedstawić następująco:

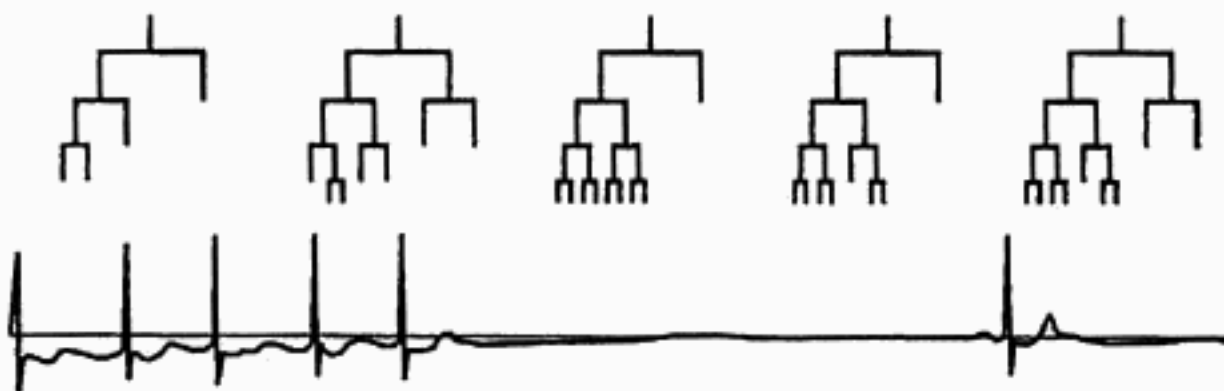
$$A_{j+1,n} = \begin{cases} B_{j+1,n} & E(x, B_{j+1,n}) < \frac{1}{2} E(x, A_{j,2n}) + \frac{1}{2} E(x, A_{j,2n+1}) \\ A_{j,2n} \oplus A_{j,2n+1} & \text{w pozostałych przypadkach} \end{cases} \quad (6.2)$$

gdzie:  $B_{j+1,n}$  jest reprezentacją w rozpatrywanym węźle uzyskaną przy pełnej dekompozycji, natomiast  $E$  oznacza entropię Shannona – wielkość stowarzyszona ze zbiorem niezależnych zdarzeń  $S = \{x_1, \dots, x_n\}$  i z rozkładem prawdopodobieństwa ich zajścia  $P = \{p_1, \dots, p_n\}$

$$E(p_1 \dots p_n) = E(S) = - \sum_{i=1}^n p_i \cdot \log(p_i) \quad (6.3)$$

Wykorzystywana w tym przypadku definicja entropii jest dolnym ograniczeniem średniej długości słowa kodu użytego do zakodowania komunikatów  $\{x_1, \dots, x_n\}$ .

Ponieważ nie wszystkie węzły drzewa dekompozycji zostały zachowane podczas optymalizacji, oprócz współczynników czasowo-częstotliwościowej reprezentacji sygnału, istnieje konieczność zapamiętania w skompresowanym pliku wynikowym także struktury drzewa według którego dokonano dekompozycji. Struktura ta może być inna dla każdej porcji sygnału (rys. 6.5).



Rys. 6.5. Adaptacja drzewa dekompozycji do zawartości sygnału EKG

Pogarsza to nieco efektywność kompresji opartej na optymalnym drzewie dekompozycji. Mimo to efektywność kompresji jest znaczna i przy wykorzystaniu odpowiednio zaprojektowanej transformacji całkowitoliczbowej połączonej z kodowaniem

reprezentacji czasowo-częstotliwościowej metodą Huffmana pozwala oczekiwać 3,5-krotnej redukcji objętości pliku wynikowego (baza MIT-BIH, częstotliwość próbkowania 360 Hz, kwantyzacja 12 bitów). Zważywszy że algorytm kompresji posiada własność dokładnej rekonstrukcji dyskretnej reprezentacji sygnału, jest to rezultat bardzo dobry.

### 6.2.2. Stratny algorytm kompresji elektrokardiogramu z użyciem pakietów falkowych

Współczynniki optymalnej czasowo-częstotliwościowej reprezentacji sygnału mogą zostać efektywnie zakodowane za pomocą stratnego algorytmu kompresji. Podobnie jak w przypadku stratnej kompresji wykorzystującej transformację falkową, założony zostaje pewien energetyczny margines dokładności odtwarzanego sygnału (np.  $1 \pm 5\%$  całkowitej energii sygnału). Następnie reprezentacja czasowo-częstotliwościowa jest upraszczana poprzez eliminację węzłów optymalnego drzewa dekompozycji, poczynwszy od węzłów o najmniejszej energii. Jednocześnie energia wyeliminowanych węzłów jest akumulowana w zmiennej pomocniczej. Osiągnięcie przez nią założonego poziomu zniekształceń (odsetka całkowitej energii sygnału) kończy proces upraszczania drzewa dekompozycji.

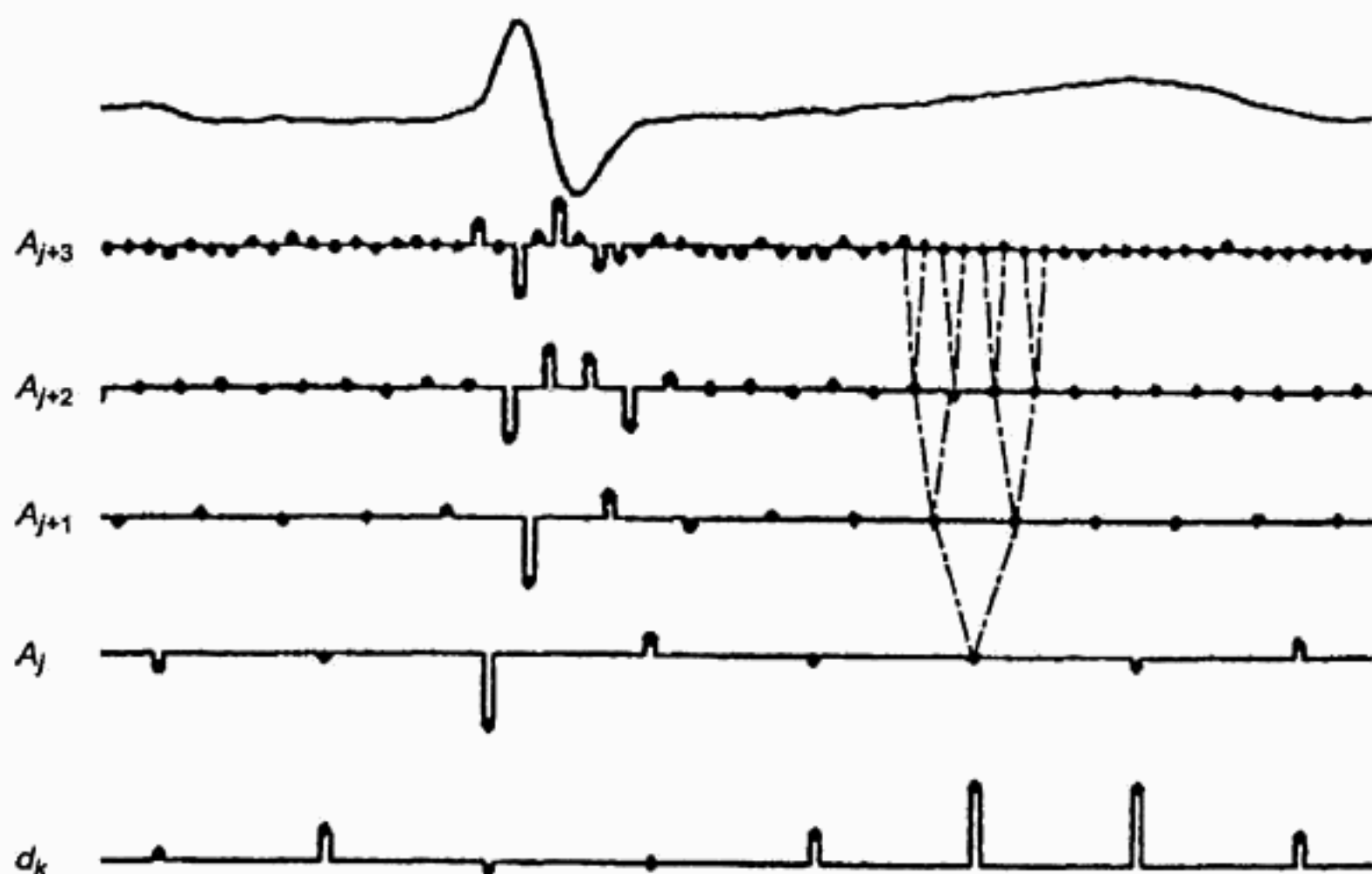
W porównaniu z upraszczaniem czasowo-częstotliwościowej reprezentacji sygnału otrzymanej przy użyciu klasycznej transformacji falkowej, upraszczanie optymalnego drzewa dekompozycji przebiega identycznie. Jednak w drugim przypadku kodowanie struktury drzewa dekompozycji wynika bezpośrednio z optymalizacji, tak więc zapamiętanie pozycji wyeliminowanych węzłów nie wymaga dodatkowego nakładu zasobów.

Alternatywne podejście do zagadnienia kompresji umożliwia algorytm EZW (ang. *embedded zerotree wavelet*) zaproponowany przez Shapiro. W algorytmie tym parametrem wejściowym jest wynikowe natężenie strumienia informacji lub współczynnik kompresji sygnału (redukcja objętości pliku wynikowego), natomiast kodowanie jest wykonywane tak, aby zapewnić minimalną wartość współczynnika zniekształceń. Podejście takie jest bardziej uzasadnione w realizacjach praktycznych, gdzie do dyspozycji pozostaje zwykle tor transmisyjny o z góry określonej przepustowości. Osiągnięcie minimalnej (dla danego natężenia strumienia informacji) wartości poziomu zniekształceń jest osiągane dzięki stosowaniu kodowania zagnieżdżonego. W kodowaniu takim uproszczone reprezentacje sygnału, odpowiadające niskiemu natężeniu strumienia informacji, są stopniowo obudowywane reprezentacjami szczegółów. W wyjściowym strumieniu danych najprostsze wersje reprezentacji stają się kolejno przedrostkami wersji bardziej szczegółowych. Po osiągnięciu docelowego natężenia strumienia informacji proces uzupełniania jest przerywany, dając gwarancję, że każdy z odrzuconych współczynników reprezentacji sygnału jest mniej istotny dla jego rekonstrukcji niż ostatni współczynnik uwzględniony w wersji skompresowanej.

Algorytm EZW polega na:

- priorytetowej transmisji położenia i znaku istotnych współczynników czasowo-częstotliwościowych,
- zwartym zakodowaniu nieistotnych współczynników z wykorzystaniem wzajemnego podobieństwa transformaty falkowej w poszczególnych skalach,
- kolejnym przybliżaniu amplitud istotnych współczynników czasowo-częstotliwościowych poprzez uporządkowaną transmisję płaszczyzn bitowych (począwszy od najstarszego bitu każdego ze współczynników).

Istotność poszczególnych próbek czasowo-częstotliwościowej reprezentacji sygnału jest określana przez porównanie ich amplitudy z zestawem oktawowo opadających wartości progowych (dla niższych częstotliwości wartości progowe są niższe). Dodatkowo, algorytm EZW wykorzystuje własność łączącą współczynniki reprezentacji czasowo-częstotliwościowej o tej samej lokalizacji czasowej i różnych skalach. Jeżeli współczynnik  $A_{j,n}$  niższego zakresu częstotliwościowego został uznany za nieistotny, to współczynniki  $A_{j+1,2n}$  oraz  $A_{j+1,2n+1}$  prawdopodobnie również okażą się nieistotne (rys. 6.6). Jeśli tak jest, współczynnik  $A_{j,n}$  nazwany jest korzeniem drzewa zerowego (ang. *root of zerotree*) i wszystkie pochodzące od niego współczynniki (dzieci) wyższych oktaw również są nieistotne.

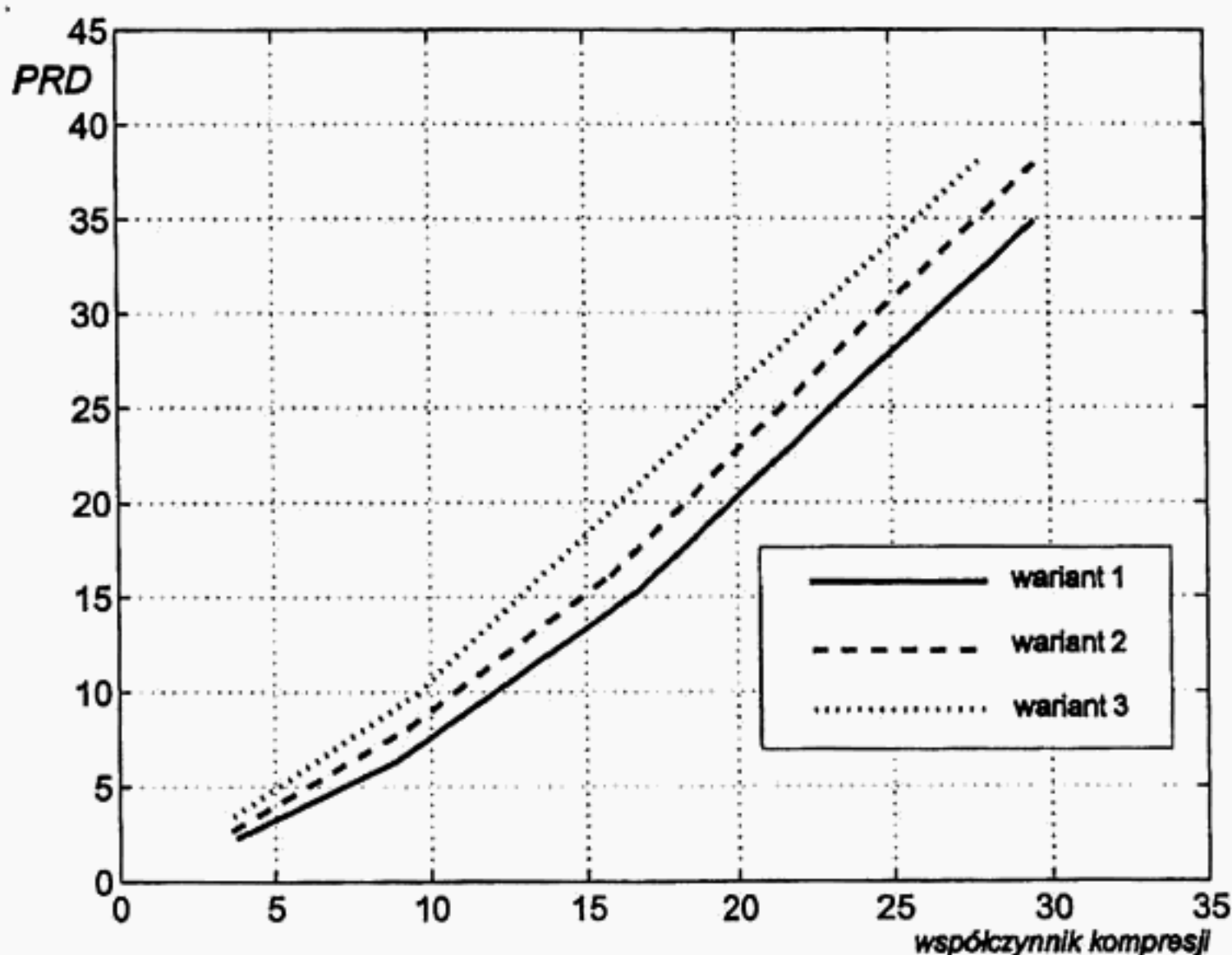


Rys. 6.6. Hierarchiczna struktura współczynników czasowo-częstotliwościowej reprezentacji sygnału. Jeżeli współczynnik  $A_{j,n}$  jest nieistotny, to jego dzieci  $A_{j+1,2n}$  oraz  $A_{j+1,2n+1}$  prawdopodobnie też są nieistotne



Zastosowanie algorytmu EZW jest możliwe zarówno do zakodowania współczynników powierzchni czasowo-częstotliwościowej otrzymanej przy użyciu klasycznej transformacji falkowej, jak i optymalnego drzewa dekompozycji *Wavelet Packets*. Zaskakujące na pierwszy rzut oka są nieco lepsze parametry kompresji osiągane przy wykorzystaniu tradycyjnej transformacji falkowej. Sytuacja taka znajduje swe uzasadnienie w jednorodności binarnego drzewa dekompozycji faworyzującej kodowanie drzew zerowych. Przeciwnie, proces optymalizacji drzewa dekompozycji skutkuje często wydzieleniem struktur podrzędnych, które nie mogą już być tak efektywnie zakodowane jako drzewa zerowe. Optymalne drzewo dekompozycji zawiera ponadto wiele istotnych wartości zlokalizowanych z dala od konturu drzewa, co dodatkowo ogranicza rozrost drzew zerowych.

Algorytm EZW w połączeniu z dekompozycją opartą na pakietach falkowych jest metodą bardzo efektywną. Doniesienia mówią o ośmiokrotnej redukcji objętości pliku wynikowego (baza MIT-BIH, częstotliwość próbkowania 250 Hz, kwantyzacja 12 bitów) przy zniekształceniach PRD rzędu 3,6% (rys. 6.7). Z punktu widzenia implementacji w rzeczywistym systemie akwizycyjno-telekomunikacyjnym niebagatelną zaletą tego algorytmu jest możliwość zadania wynikowego natężenia strumienia informacji jako parametru kontrolującego przebieg kompresji.



Rys. 6.7. Efektywność kompresji z użyciem algorytmu EZW w połączeniu z dekompozycją opartą na pakietach falkowych

### 6.3. Kompresja z użyciem dodatkowych transformacji

Kompresja czasowo-częstotliwościowej reprezentacji sygnału może być bardziej efektywna, jeśli transformacja falkowa zostanie złożona z inną transformacją wykorzystującą odmienny typ redundancji informacji w sygnale. Najczęściej są to: redundancja przestrzenna będąca rezultatem podobieństwa informacji pozyskiwanych przez położone we wzajemnym bliskim sąsiedztwie elektrody (zwłaszcza odprzewodzenia przedsercowe EKG) oraz redundancja czasowa będąca wynikiem lokalnie znacznie przeszacowanego pasma sygnału. Innym aspektem, w którym redundancja sygnału może być znaczna, jest jego pseudookresowość. Kolejne ewolucje serca są zwykle bardzo podobne, toteż zakodowanie jedynie informacji różnicowej może się okazać korzystne.

Ponieważ argumentem transformacji redukującej redundancję sygnału jest wektor danych, zależnie od wybranego kierunku redukcji może się on składać z:

- reprezentacji pojedynczej ewolucji serca w jednym kanale – w tym przypadku kolejne ewolucje będą reprezentowane przez kolejne wektory danych, a więc w wyniku transformacji zredukowany będzie nadmiar informacji wynikający z podobieństwa kolejnych ewolucji (dekorelacja długoczasowa);
- reprezentacji porcji sygnału EKG w jednym kanale – w tym przypadku kolejne kanały będą reprezentowane przez kolejne wektory danych, a więc w wyniku transformacji zredukowany będzie nadmiar informacji wynikający z podobieństwa zapisu w sąsiednich kanałach (dekorelacja przestrzenna).

Zastosowanie transformacji do dekorrelacji długoczasowej pojedynczego kanału zapisu EKG wymaga detekcji wszystkich ewolucji serca i zapisania każdej z nich jako odrębnego wektora danych. Istotne jest zachowanie synchronizacji zapisów poszczególnych ewolucji, jak również i to, żeby każda z nich miała jednakową długość. Synchronizację najlepiej osiągnąć poprzez wyznaczenie położenia prawdziwego wierzchołka załamka R jako wierzchołka paraboli najlepiej dopasowanej do tego załamka. Z kolei uzupełnianie długości każdego wektora danych może być zrealizowane przez dopisanie przed i za nim próbek o wartości równej średniej kilku skrajnych próbek. Z punktu widzenia dekorrelacji, wystarczające jest uzupełnianie do długości najdłuższego wektora danych, natomiast jeśli kolejnym przekształceniem ma być transformacja czasowo-częstotliwościowa, wektory danych powinny mieć długość możliwą do wyrażenia przy pomocy całkowitej potęgi liczby 2.

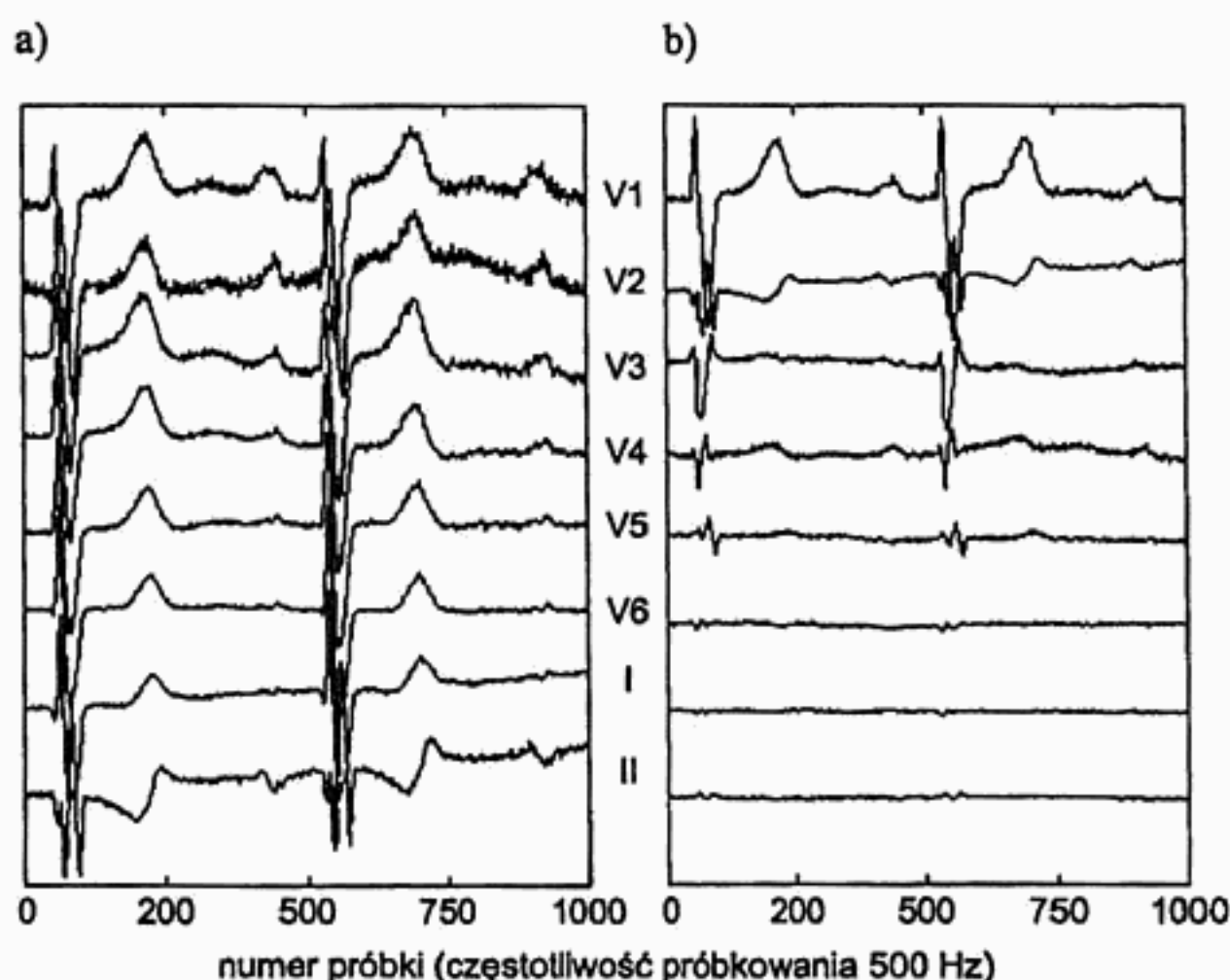
W przypadku użycia transformacji do dekorrelacji przestrzennej wieloodprzewodzeniowego zapisu EKG żadne procedury synchronizacyjne nie są wymagane, o ile poszczególne kanały są zapisane synchronicznie przez rejestrator. Długość porcji danych nie jest uwarunkowana przez transformację, ale podobnie jak w przypadku dekorrelacji długoczasowej, jeśli kolejnym przekształceniem ma być transformacja czaso-

wo-częstotliwościowa, powinna ona być możliwa do wyrażenia za pomocą całkowitej potęgi liczby 2.

W dalszej części podrozdziału 6.3 opisane są dwa warianty złożenia transformacji falkowej z przekształceniami realizującymi dekorrelację przestrzenną i długoczasową elektrokardiogramu. Szczegóły dotyczące użytych przekształceń zawiera dodatek D2 („Transformacja Karhunen–Loevego” – KLT) oraz dodatek D3 („Dyskretna transformacja kosinusowa” – DCT).

### 6.3.1. Reprezentacja czasowo-częstotliwościowa sygnałów zdekorelowanych

Pierwszy z wariantów złożenia transformacji KLT i DCT z transformacją falkową w celu kompresji elektrokardiogramów został zaproponowany przez A. Cetina w 1993 roku. Złożenie to polega na dokonaniu dekorrelacji przestrzennej wieloodprowadzeniowego zapisu EKG w dziedzinie czasu (rys. 6.8), a następnie użyciu transformacji czasowo-częstotliwościowej opartej na filtrach kwadraturowych oraz decymacji.

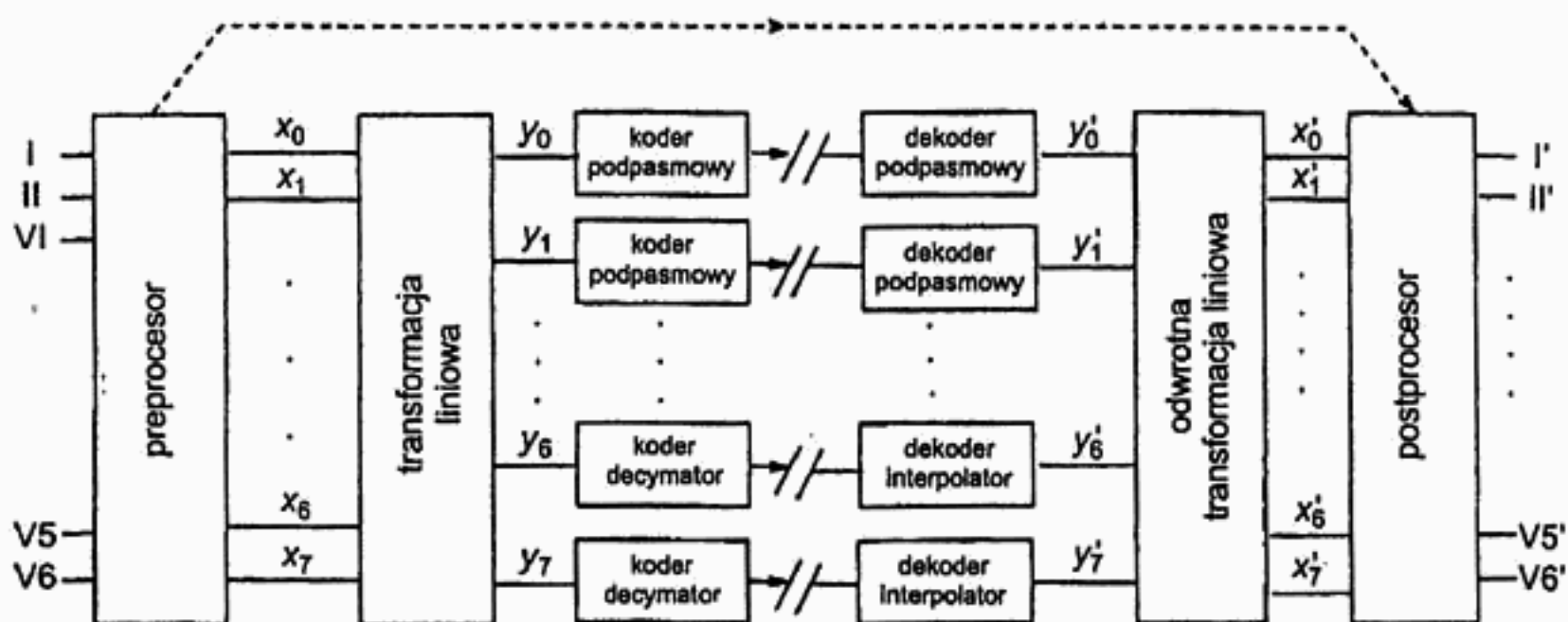


Rys. 6.8. Dekorrelacja przestrzenna wieloodprowadzeniowego zapisu EKG w dziedzinie czasu:  
a) sygnał oryginalny; b) sygnał zdekorelowany

Schemat blokowy kompresji elektrokardiogramów zdekorelowanych przedstawia rysunek 6.9. Dekorrelacja sygnałów w dziedzinie czasu jest wykorzystana podczas transformacji czasowo-częstotliwościowej w ten sposób, że kanały V1÷V5 zawierające

sygnał o znacznej amplitudzie są przetwarzane odmiennie od niskoenergetycznych kanałów V6÷II. Sygnały V1÷V5 są przetwarzane przez algorytm realizujący następujące kroki (rys. 6.10a):

1. Dekompozycja do reprezentacji czasowo-częstotliwościowej z użyciem kwadraturowych filtrów lustrzanych (dwa poziomy).
2. Dekorelacja krótkoczasowa aproksymacji drugiego poziomu (0÷64 Hz) za pomocą transformacji kosinusowej (zniesienie korelacji między sąsiednimi próbkami) najistotniejsze jej współczynniki są następnie kodowane algorytmem Huffmana.
3. Kodowanie stratne sygnałów detali pierwszego i drugiego poziomu (64÷250 Hz) poprzez eliminację wartości, których amplituda nie przekracza założonego progu.



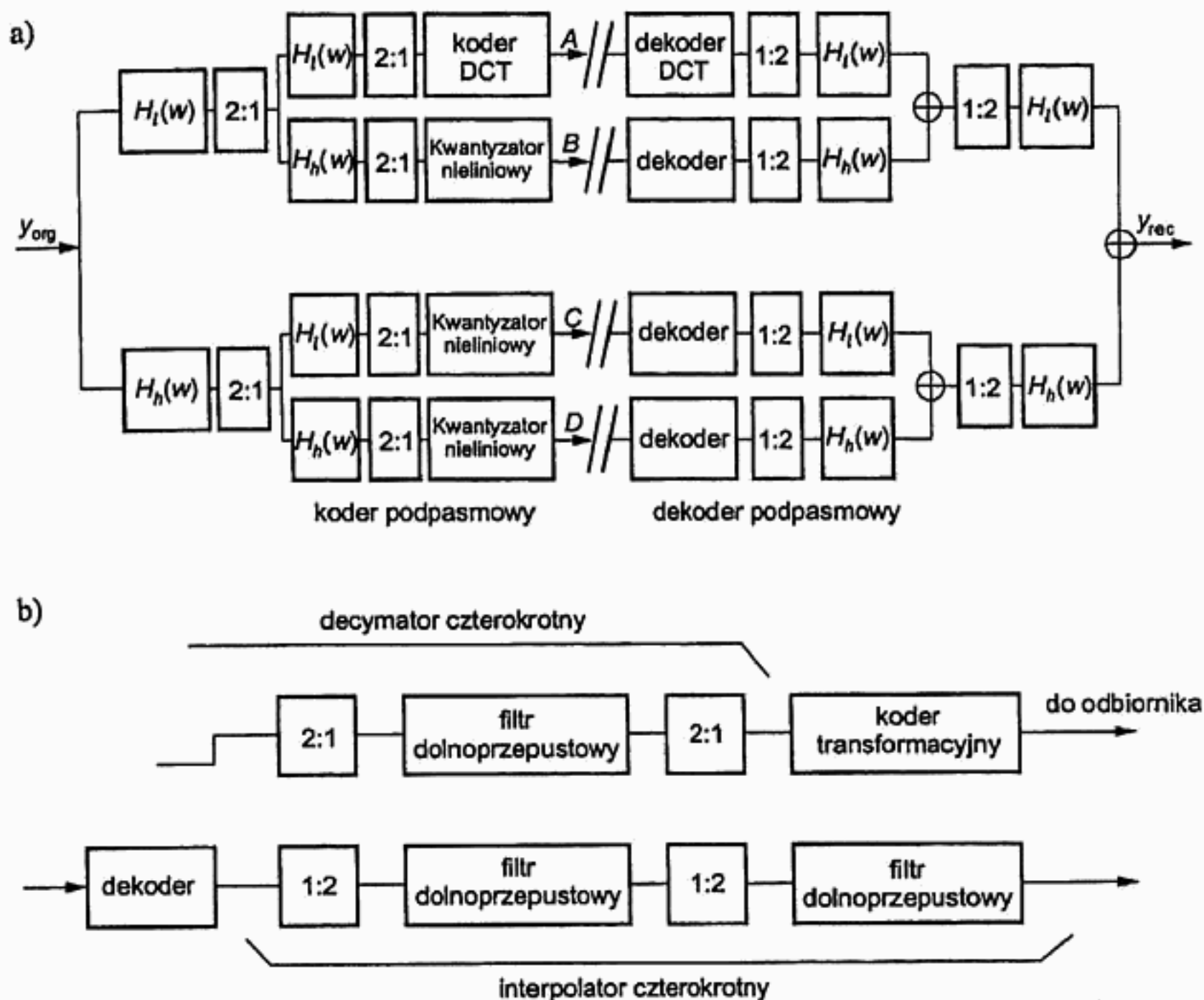
Rys. 6.9. Schemat blokowy kompresji elektrokardiogramów zdekorelowanych

Sygnały V6÷II są przetwarzane przez uproszczony algorytm nie uwzględniający w ogóle składowych wysokoczęstotliwościowych (rys. 6.10b):

1. Decymacja o dwa poziomy.
2. Dekorelacja krótkoczasowa za pomocą transformacji kosinusowej (zniesienie korelacji między sąsiednimi próbkami).
3. Kodowane najistotniejszych współczynników algorytmem Huffmana.

Według doniesień autorów metody efektywność kompresji zapewnia 6,2-krotną redukcję objętości pliku wynikowego (baza CSE, częstotliwość próbkowania 500 Hz, kwantyzacja 12 bitów) przy zniekształceniach PRD rzędu 3,7%. Próba zastosowania transformacji kosinusowej do dekodekacji sygnału skutkuje obniżeniem efektywności kompresji do wartości 5.





Rys. 6.10. Procedury kodowania sygnałów zdekorrelowanych: a) dla pięciu głównych wymiarów; b) dla pozostałych wymiarów

### 6.3.2. Dekorelacja czasowo-częstotliwościowych reprezentacji sygnałów

Odwrotny wariant złożenia transformacji KLT z transformacją falkową w celu kompresji elektrokardiogramów został zaproponowany przez B. Bradie w 1996 roku. Złożenie to polega na dokonaniu transformacji sygnału do dziedziny czasowo-częstotliwościowej i dekorrelacji długoczasowej pojedynczego odprowadzenia zapisu EKG.

Pierwszym krokiem jest utworzenie zespołu danych z wektorów danych, z których każdy odpowiada jednej ewolucji serca podczas 10 min zapisu. Wektory te są synchronizowane i o znormalizowanej długości 512 próbek.

Każdy z wektorów danych jest następnie całkowicie dekomponowany z użyciem falki Coiflets 12. rzędu za pomocą algorytmu *Wavelet Packets*. Reprezentacje poszcze-

gólnych wektorów danych leżące w odpowiadających sobie węzłach drzewa są następnie uśredniane, w wyniku czego powstaje drzewo wspólne reprezentatywne dla wszystkich wektorów zespołu danych.

Drzewo wspólne jest następnie przedmiotem optymalizacji, która polega na minimalizacji funkcji kosztu zdefiniowanej jako

$$C(\alpha) = \sum_i \log|\alpha_i| \quad (6.4)$$

gdzie  $\alpha$  oznacza węzeł drzewa dekompozycji, natomiast  $\alpha_i$  jest  $i$ -tym współczynnikiem reprezentacji w tym węźle. Funkcja kosztu została celowo wybrana tak, aby wybrane z jej użyciem drzewo dekompozycji było optymalne z punktu widzenia efektywności transformacji KLT.

Następnie węzły drzewa wspólnego, począwszy od najmniejszych wartości reprezentacji, są eliminowane aż do zmniejszenia globalnej energii sygnału o 2,5%. Tak obcięta struktura drzewa wspólnego zostaje narzucona wszystkim reprezentacjom wektorów danych. W rezultacie dotychczas pełne drzewa dekompozycji przyjmują jednakową strukturę zoptymalizowaną globalnie w poprzednim kroku. Jest to związane ze znacznym zmniejszeniem długości oraz utratą najmniej istotnych informacji w każdym z wektorów danych.

Kolejnym krokiem jest dekorelacja obciętych wektorów danych za pomocą transformacji KLT. Celem dekorelacji jest eliminacja nadmiarowych informacji, których źródłem jest podobieństwo kolejnych ewolucji serca. Macierz wynikowa zawiera składowe główne, z których te o najniższych indeksach reprezentują największą wariancję.

Ostatnim etapem jest eliminacja składowych głównych, począwszy od tych o najwyższych indeksach (a więc reprezentujących najniższą wariancję), aż do zmniejszenia globalnej energii sygnału o 1%.

Oczywiście tak złożony proces przetwarzania generuje na każdym etapie sporą liczbę współczynników, które muszą być przechowywane lub przesłane wraz z sygnałem w celu prawidłowej rekonstrukcji. Są to m.in. prawdziwe długości poszczególnych ewolucji, struktura drzewa wspólnego i macierz transformacji KLT. Pomimo dodatkowej ilości danych, których strumień stanowi prawie 25% pliku wynikowego, opisywana metoda kompresji należy do bardzo efektywnych, umożliwia ona średnio 21-krotną redukcję objętości pliku wynikowego (baza MIT-BIH, częstotliwość próbkowania 360 Hz, kwantyzacja 12 bitów) przy zniekształceniach sygnału nie przekraczających 2,8% (rys. 6.11). Wadą algorytmu jest niestety jego złożoność. Kompresja jest wykonywana na sygnale zarejestrowanym i dla jednego kanału z trudem mieści się w czasie odpowiadającym długości trwania zapisu (SPARC Workstation).



Rys. 6.11. Porównanie sygnałów oryginalnego i zrekonstruowanego z kompresji metodą dekorelacji najlepszych drzew dekompozycji *Wavelet Packets*: a) rytm zatokowy; b) bigeminia

## 6.4. Kompresja z modulacją dystrybucji zniekształceń

W metodzie opisywanej w tym rozdziale wykorzystano fakt niejednorodnej dystrybucji informacji diagnostycznej w elektrokardiogramie. Opisywany algorytm jest wprawdzie dedykowany do reprezentacji aktywności serca, ale może być łatwo zaadaptowany do każdego innego sygnału, w którym możliwe jest wyróżnienie interwałów bardziej i mniej istotnych. W elektrokardiogramie wyróżnienie takie może zostać dokonane automatycznie za pomocą algorytmu wyznaczającego granice załamków P, QRS i T, który należy do kanonu wspomagania diagnostyki medycznej i jest akceptowany w świecie medycznym.

Istotną nowością tego algorytmu kompresji jest lokalna kontrola poziomu zniekształceń, dzięki czemu mogą one zostać skumulowane w mniej istotnych diagnostycznie partiach sygnału.

### 6.4.1. Aspekt lokalnej istotności elektrokardiogramu

Elektrokardiogram reprezentuje zjawiska elektryczne związane z poszczególnymi fazami ewolucji serca. Dlatego dla diagnozującego kardiologa pewne partie sygnału są bardziej istotne, a inne mniej. Z tego powodu wyrażanie wielkości zniekształceń poprzez globalny współczynnik o charakterze energii, jakim jest PRD, nie jest właściwe w przypadku elektrokardiogramów. Bardziej uzasadnione jest użycie tej miary do porównywania dwóch analogicznych porcji sygnału (np. dwóch załamków P).

Przyjęcie tezy o nierównomiernym rozkładzie informacji diagnostycznych w sygnale skłania także do rewizji dotychczasowych kryteriów optymalizacji parametrów kompresji. Nowe wytyczne do konstruowania algorytmów powinny opierać się na założeniu minimalnego zniekształcenia informacji medycznych, zamiast założenia o zachowaniu energii. Konsekwentnie, zamiast pozostawiać wybór pomiędzy algorytmami bezstratnymi o niewielkiej efektywności i algorytmami stratnymi dowolnie zniekształcającymi informację, nowa grupa algorytmów „bezstratnych diagnostycznie” pozwala użytkownikowi decydować, jaki poziom zniekształceń jest dopuszczalny dla każdego segmentu sygnału. Dodatkowo, lokalna miara zniekształceń pozwala lepiej interpretować odchyłki sygnału w kontekście diagnostycznym.

Medyczna wiedza na temat istotności poszczególnych fragmentów zapisu EKG jest trudna do opisu, na dodatek każdy kardiolog inaczej postrzega cechy sygnału. Mimo rozbieżności, uwarunkowania sposobu interpretacji można sformułować następująco.

- Przeważająca większość elektrokardiogramów to badania wykonywane w celu diagnozy konkretnej jednostki chorobowej, mające dać odpowiedź na wcześniej postawione pytanie, tylko niewiele pozostałych to badania przesiewowe, prowadzone bez wcześniejszej wiedzy o pacjencie. Sposób interpretacji zapisu jest więc zależny od wiedzy na temat pacjenta.
- Odpowiednio do celu diagnostycznego badania używany jest właściwy sprzęt rejestrujący: systemy wysiłkowe, całodobowe rejestratory ambulatoryjne i kardiomonitoring przyłóżkowe pracują w odmiennych sytuacjach i dostarczają innego zakresu wiedzy kardioelektrycznej. Sposób interpretacji zapisu jest zależny od rejestratora, w którym algorytm jest zaimplementowany.
- Doświadczenie kardiologa, pomimo silnych ostatnio tendencji normalizacyjnych, odgrywa w dalszym ciągu istotną rolę w ocenie zapisu EKG. Kardiolodzy zdobywając doświadczenie przez lata praktyki, stają się specjalistami w coraz węższym zakresie dysfunkcji serca, co powoduje indywidualne preferencje poszczególnych odcinków zapisu. Wynika z tego, że sposób interpretacji jest zależny także od eksperta.

Konkluzją z przedstawionych powyżej rozważań jest brak spójności oczekiwań i niemożność jednoznacznego określenia stopnia istotności poszczególnych segmentów zapisu. W tej sytuacji skutecznym rozwiązaniem jest pozostawienie użytkownikowi (lekarzowi) wolnej ręki w decydowaniu o tym, które informacje są dla niego najważniejsze w danym przypadku diagnostycznymi i gdzie reprezentowane są one w zapisie. Decyzja taka zostaje podjęta na modelowej ewolucji serca zawierającej punkty początkowe i końcowe załamków, a następnie przeskalowana dla każdej ewolucji serca indywidualnie w oparciu o rzeczywiste granice załamków.

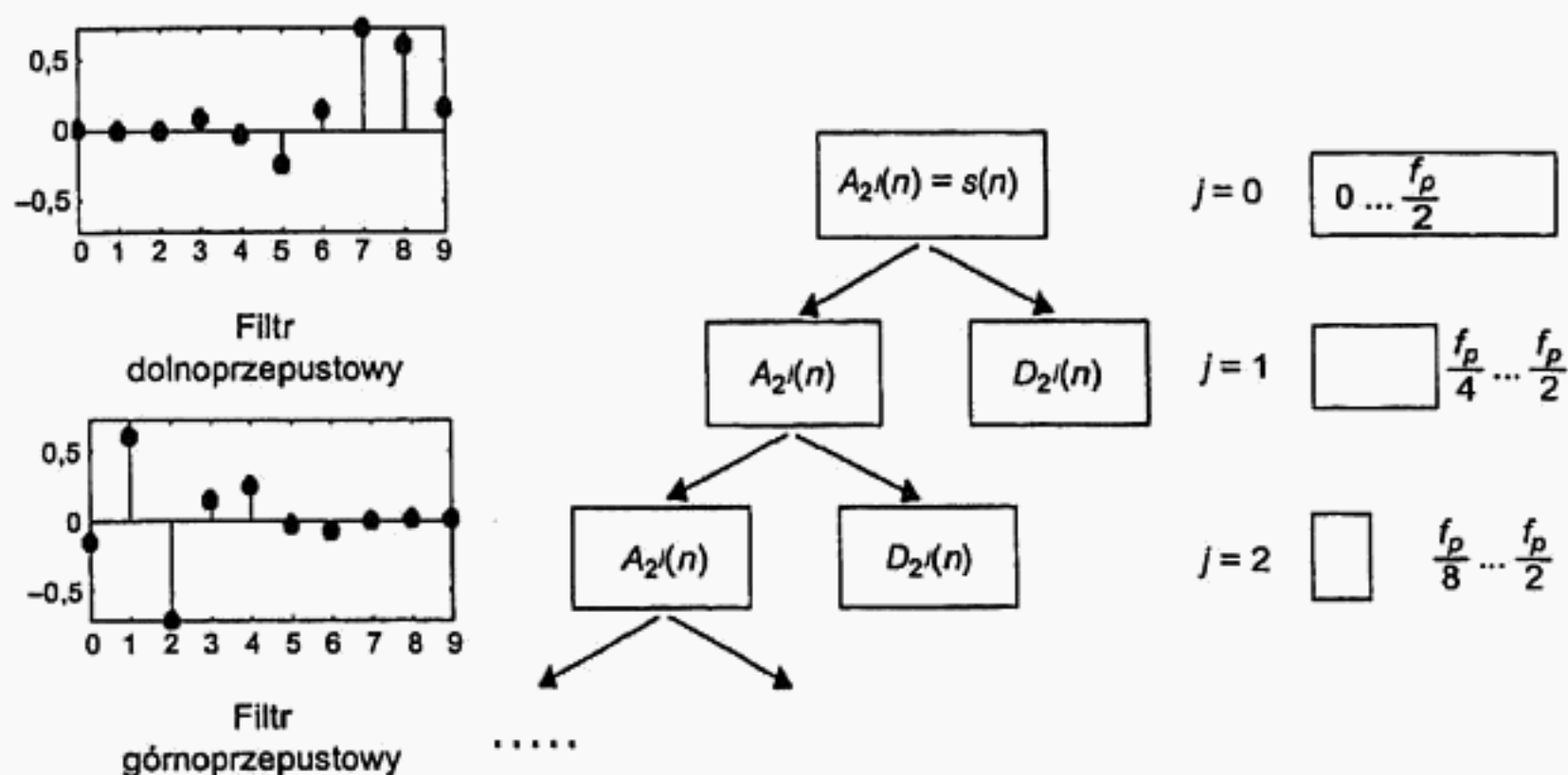
#### **6.4.2. Realizacja algorytmu kompresji z użyciem transformacji falkowej**

Algorytm kompresji oparty na transformacji falkowej korzysta z wyznaczonych wcześniej granic załamków P, QRS i T. Po dekompozycji sygnału do dziedziny czasowo-częstotliwościowej pozycje załamków są także przeskalowane do tej dziedziny. Dokładność pozycjonowania załamków nie odgrywa przy tym bardzo istotnej roli, gdyż jednostka czasu na płaszczyźnie czasowo-częstotliwościowej jest większa niż w dziedzinie czasu (dla niższych oktaw – wielokrotnie większa).

Transformacja użyta w tym algorytmie musi spełniać założenie dokładnej rekonstrukcji, gdyż jedynie w tym przypadku możliwa jest kontrola zniekształceń sygnału wynikowego za pomocą operacji na reprezentacji czasowo-częstotliwościowej. Drugim wymaganiem jest użycie falek o nośniku zwartym. Zbyt krótki nośnik zmniejsza wprowadzie zniekształcenia sygnału spowodowane efektem brzegowym, ale pogarsza



także separację pasm częstotliwości i prowadzi do rozproszenia energii na wiele komponentów. Ponieważ dokładna rekonstrukcja sygnału oryginalnego nie jest oczekiwana (algorytm stratny) i zaokrąglenia w dziedzinie reprezentacji są dopuszczalne, nie jest konieczne używanie algorytmu liftingowego w celu otrzymania całkowitoliczbowej reprezentacji sygnału. Transformacja użyta w opisywanym algorytmie kompresji wykorzystuje biortogonalne filtry Daubechies 5. rzędu i algorytm dekompozycji piramidowej (rys. 6.12).



Rys. 6.12. Współczynniki nośnika biortogonalnych filtrów Daubechies 5. rzędu i algorytm dekompozycji piramidowej wykorzystane w algorytmie kompresji falkowej z modulacją dystrybucji zniekształceń

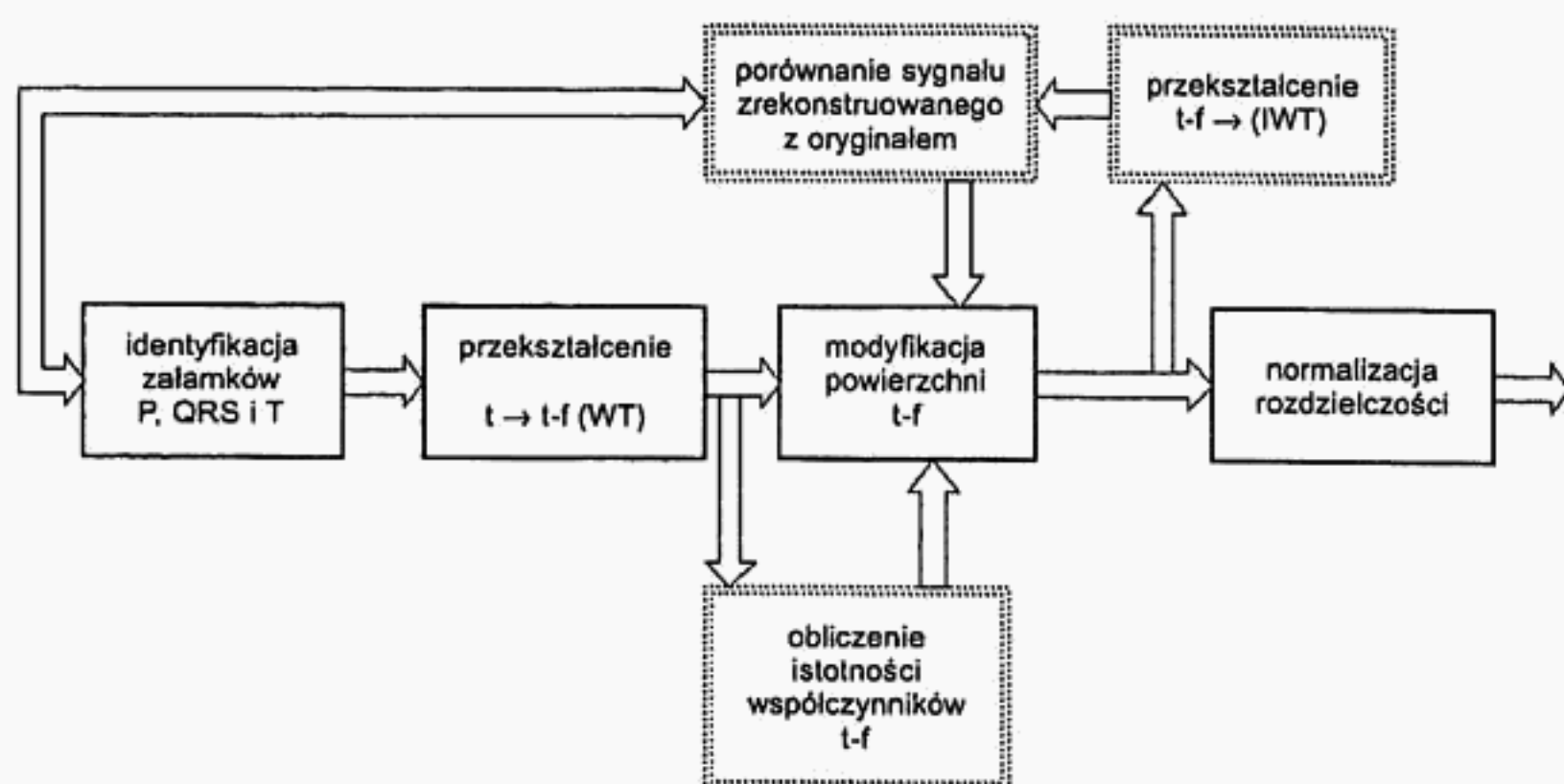
Dekompozycja sygnału o oryginalnej częstotliwości próbkowania 256 Hz jest przeprowadzana na trzech poziomach, co umożliwia wydzielenie oktaw o pasmach: 64÷128 Hz, 32÷64 Hz oraz 16÷32 Hz. Poniżej 16 Hz sygnał oryginalny jest zachowany, gdyż:

- wykazano mały związek zawartości informacyjnej sygnału z załamkami elektrokardiogramu,
- efektywna częstotliwość próbkowania wynosi 32 Hz i odpowiadający jej okres (ok. 31 ms) jest porównywalny z długością załamków,
- sygnał o paśmie 0÷16 Hz jest reprezentowany przez niewielką liczbę współczynników i ich ewentualna modyfikacja nie jest już tak interesująca z punktu widzenia kompresji.

Redukcja ilości informacji jest uzyskiwana poprzez eliminację współczynników reprezentacji czasowo-częstotliwościowej leżących poza pasmem chwilowym sygnału użytecznego. Pasma chwilowe jest to binarna funkcja dwóch zmiennych czasu i częstotliwości określająca przynależność komponentu na podstawie jego położenia względem punktów granicznych załamków. Funkcja ta rozdziela płaszczyznę czasowo-czę-

stotliwościową na dwie części, z których tylko jedna jest uważana za reprezentację sygnału użytecznego, a druga jest nadmiarowa. Spodziewane pasmo chwilowe elektrokardiogramu zostało wyznaczone eksperymentalnie, ale może być modyfikowane przez użytkownika.

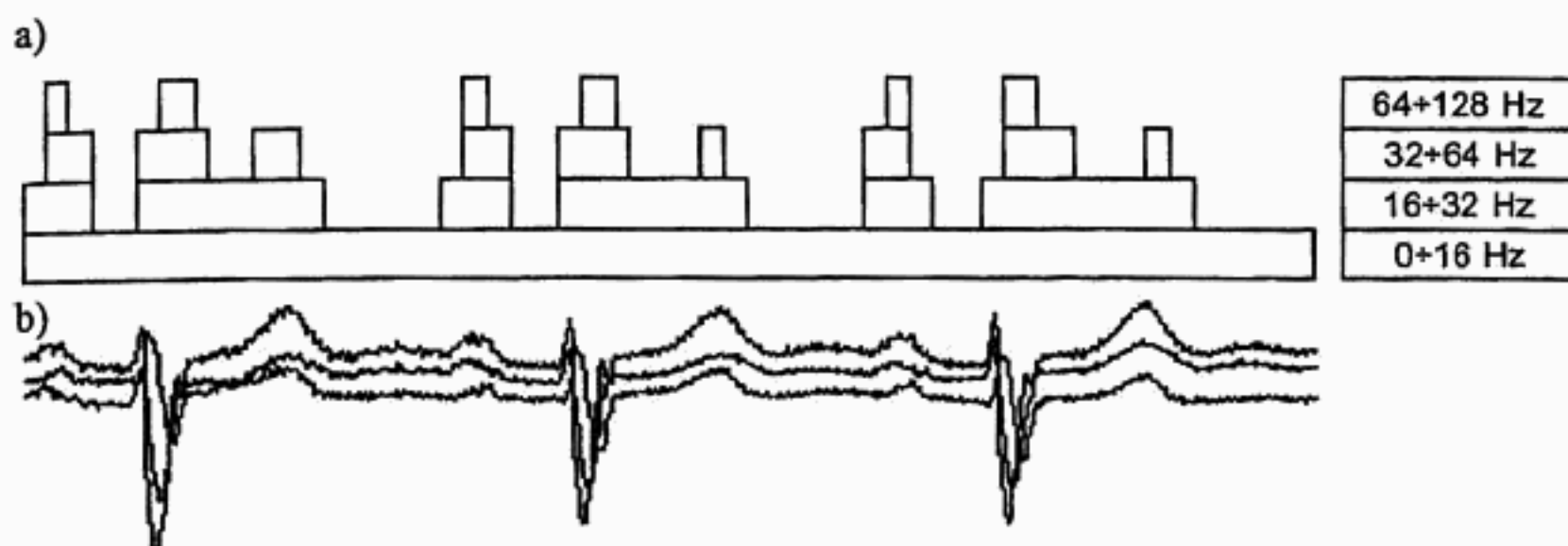
Weryfikacja poziomu zniekształceń sygnału jest wykonywana poprzez próbną rekonstrukcję sygnału skompresowanego z użyciem odwrotnej transformacji falkowej (rys. 6.13). Jeżeli różnica średniokwadratowa sygnału zrekonstruowanego i oryginalnego przekracza założoną wartość, eliminacja współczynników powierzchni czasowo-częstotliwościowej jest dokonywana ponownie z zachowaniem proporcji określonych w funkcji pasma chwilowego.



Rys. 6.13. Schemat blokowy algorytmu kompresji falkowej z modulacją dystrybucji zniekształceń

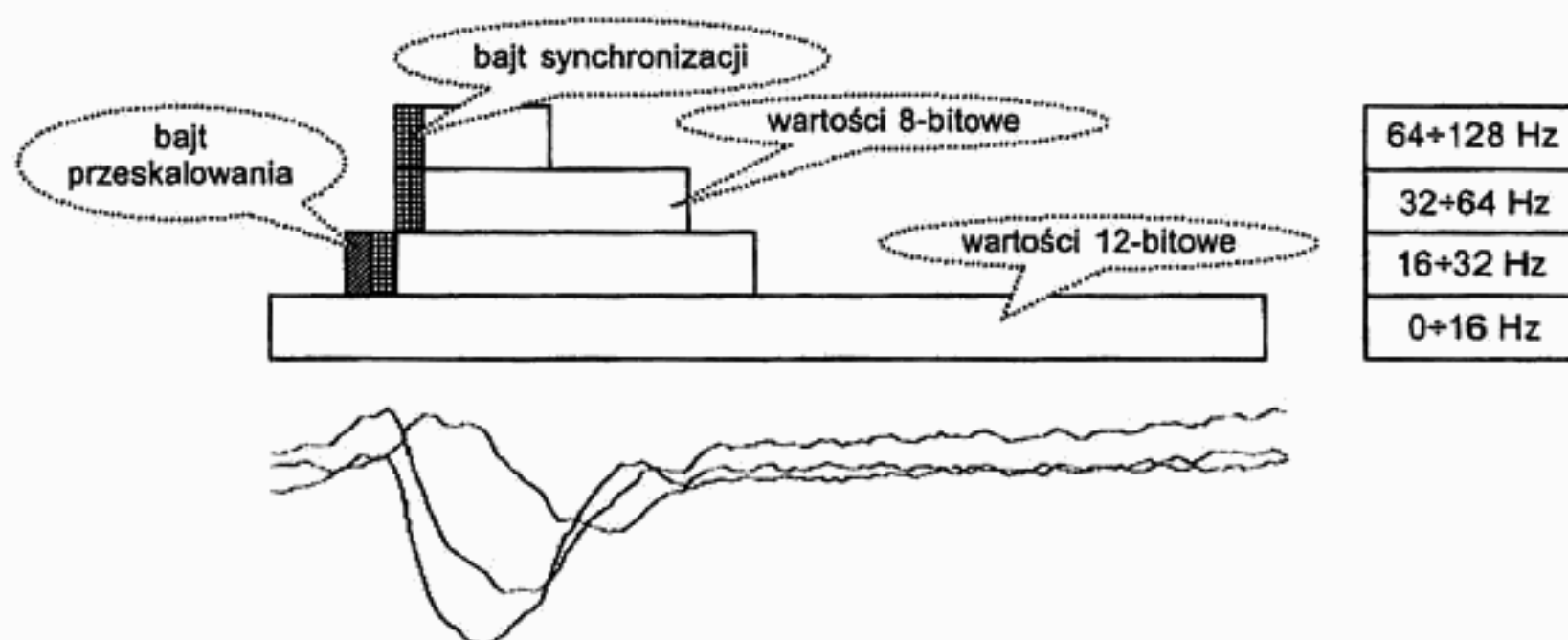
Użycie modyfikowanej funkcji chwilowego pasma prowadzi do uzyskania powierzchni czasowo-częstotliwościowej, na której sygnał poniżej ostatniego poziomu dekompozycji (16 Hz) będzie ciągły, natomiast sygnały wyższych oktaw będą tylko odcinkami niezerowe (rys. 6.14). Wystąpią one tylko tam, gdzie jest to istotne z punktu widzenia treści informacyjnej sygnału.

Niezerowe odcinki sygnałów wyższych oktaw są poprzedzone bajtem synchronizującym zawierającym pozycję ich początku względem sygnału ciągłego, natomiast zawartość odcinków jest przedmiotem normalizacji rozdzielczości. Normalizacja rozdzielczości polega na ustaleniu – dla każdego odcinka sygnału indywidualnie – nowej jednostki odpowiadającej 1 LSB, tak aby rozdzielczość nie przekraczała 8 bitów. W większości przypadków odcinki sygnałów wyższych oktaw niosą niewielką energię, najstarsze bity nie są wykorzystane i żadne przeskalowanie nie jest konieczne. W przeciwnym przypadku odszukiwane jest maksimum odcinka i wszystkie jego wartości są normalizowane tak, aby wartość maksymalna była reprezentowana przez 255.



**Rys. 6.14.** Powierzchnia czasowo-częstotliwościowa sygnału EKG zmodyfikowana przez funkcję pasma chwilowego: a) ciągły sygnał poniżej ostatniego poziomu dekompozycji i sygnały wyższych oktaw, odcinkami ciągle, występujące tylko tam, gdzie jest to niezbędne z punktu widzenia zachowania treści informacyjnej sygnału; b) oryginalny sygnał EKG

Jeżeli przeskalowanie amplitudy było wykonane, nowa wartość odpowiadająca 1 LSB jest zapisywana w bajcie poprzedzającym sygnał. Warto zwrócić uwagę, że jednostka ta może być inna dla kolejnych odcinków sygnału należących do tej samej oktawy. Sygnał ciągły w paśmie 0÷16 Hz jest przenoszony z pełną rozdzielczością 12 bitów. Struktura sygnału skompresowanego jest przedstawiona na rysunku 6.15.

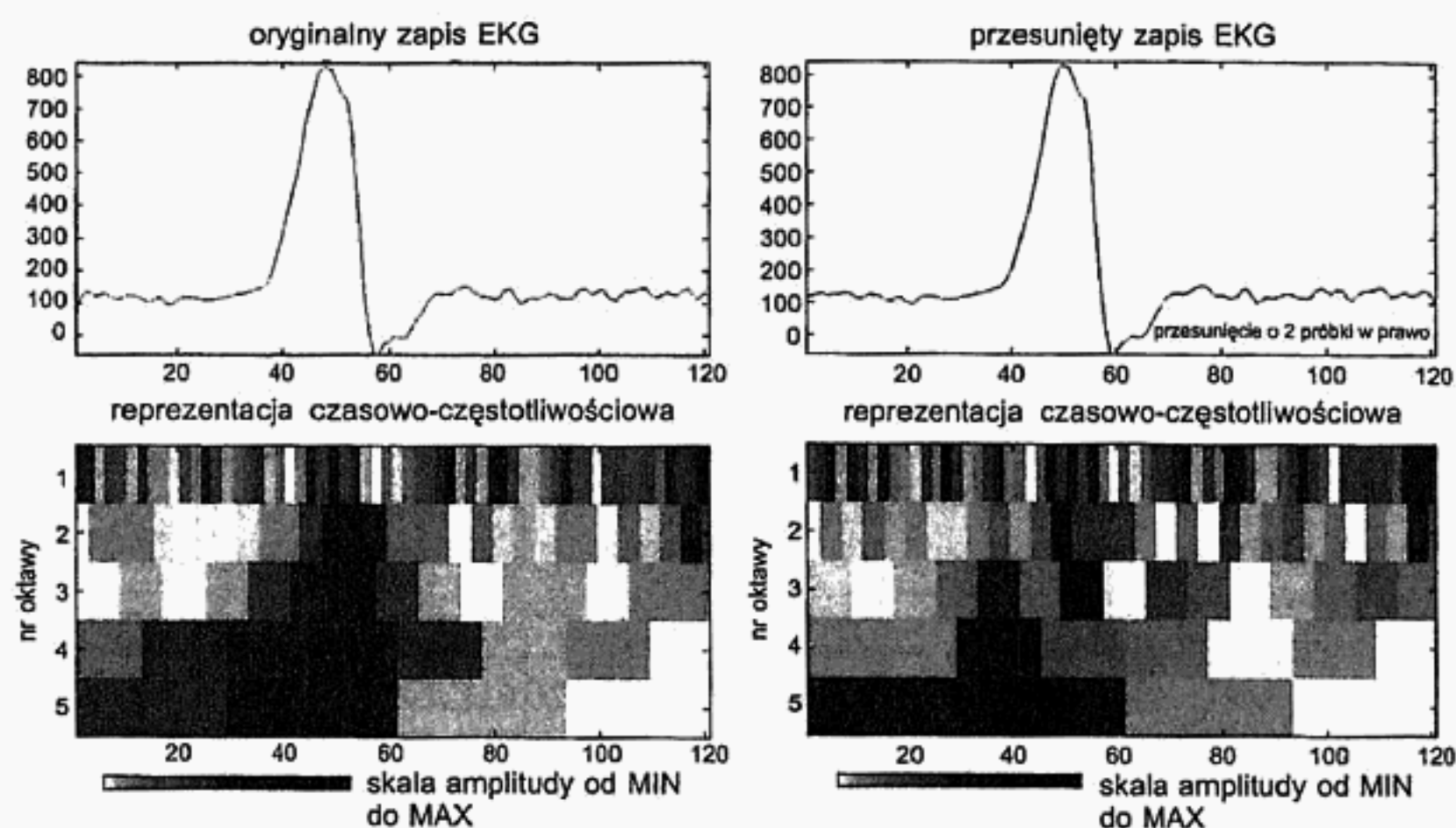


**Rys. 6.15.** Struktura sygnału skompresowanego z modulacją dystrybucji zniekształceń. Cechą charakterystyczną jest synchronizacja okazjonalnie występujących komponentów szczegółowych

Stopień kompresji sygnału EKG w dużej mierze zależy od lokalnych jego własności, toteż nie jest możliwe podanie dokładnej wartości obowiązującej dla wszystkich zapisów. Kompresja plików z podstawowego katalogu kardiologicznej bazy danych MIT-BIH (MITDB, po przepróbkowaniu na 256 Hz) pozwala oczekiwać 4,9-krotnej

redukcji objętości pliku wynikowego, co wydaje się wartością znaczną przy dopuszczeniu zniekształceń lokalnie tylko sięgających 5% (podawana zwykle średnia wartość PRD jest znacznie mniejsza).

Dalszą poprawę efektywności kompresji umożliwia modyfikacja powyższego algorytmu polegająca na zakodowaniu wysokoczęstotliwościowych składowych sygnału przy użyciu słownika. W ten sposób wykorzystana jest powtarzalność kolejnych ewolucji serca. Warunkiem powodzenia jest zapewnienie wzajemnie jednoznacznego przekształcenia sygnału w dziedzinie czasu do dziedziny czasowo-częstotliwościowej. Zagadnienie to wymaga uwagi, gdyż brak synchronizacji ewolucji serca z płaszczyzną czasowo-częstotliwościową może spowodować reprezentację tego samego przebiegu za pomocą różnych zestawów współczynników (rys. 6.16). W konsekwencji, z punktu widzenia dziedziny czasowo-częstotliwościowej założona powtarzalność ewolucji nie wystąpi, a zatem efektywność kodowania słownikowego byłaby niewielka.

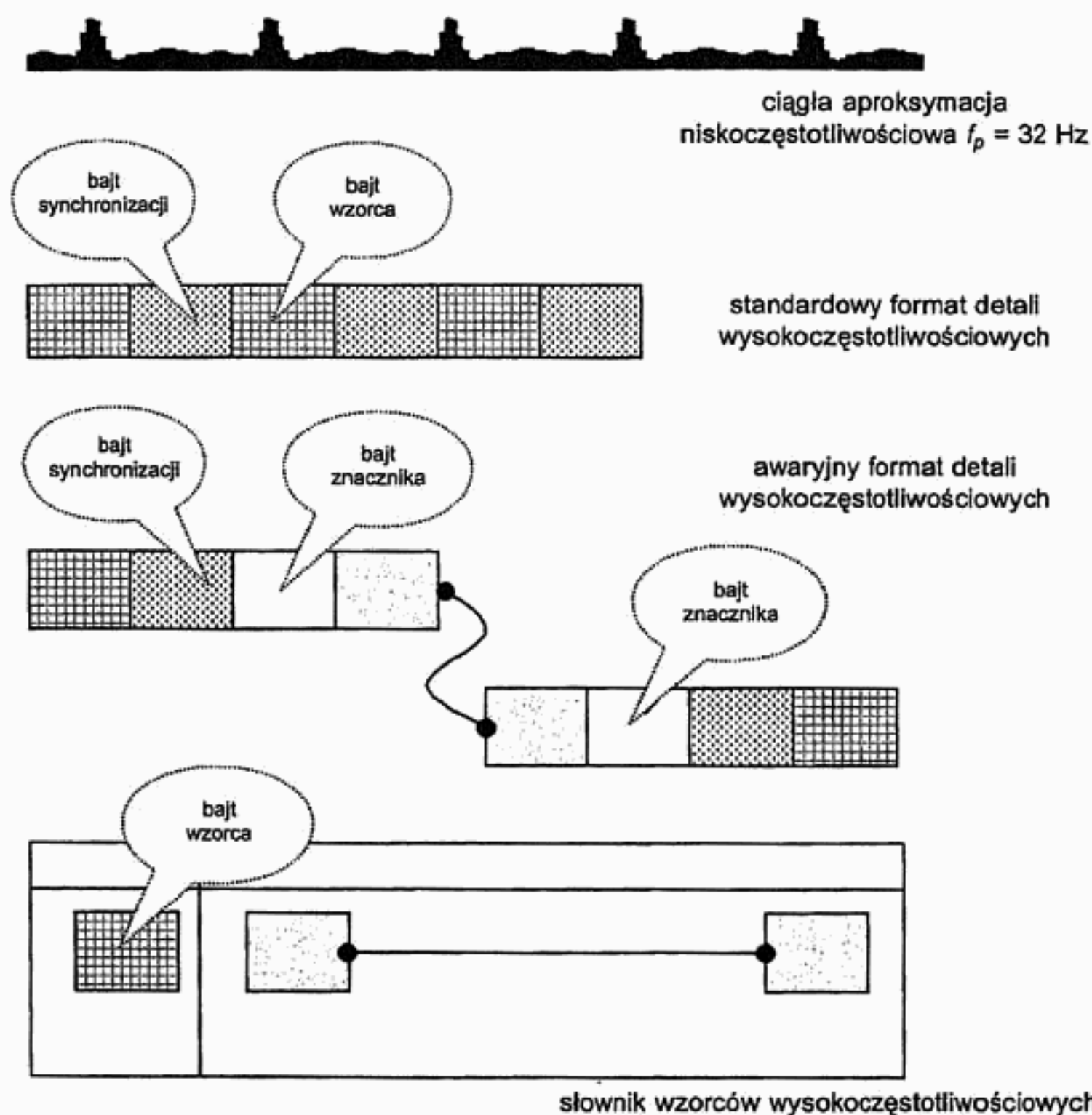


**Rys. 6.16.** Reprezentacja czasowo-częstotliwościowa dwóch identycznych sygnałów EKG. Kopia na rysunku z prawej strony jest przesunięta o dwie próbki na osi czasu względem początku odcinka dekompozycji

Kodowanie słownikowe polega na wyodrębnieniu wzorców czasowo-częstotliwościowych (maksymalnie 300) i umieszczaniu w pliku wynikowym numeru wzorca zamiast jego zawartości. Konieczne jest w tym przypadku uzupełnienie strumienia danych wyjściowych o tabelę translacji zawierającą reprezentację czasowo-częstotliwościową wszystkich użytych wzorców, tak więc znacznej efektywności należy oczekiwać dla sygnałów dłuższych, o przewadze rytmu zatokowego. W przypadku krótkiego odcinka rytmu zawierającego wiele zróżnicowanych morfologii podstawowe założenie kodowania słownikowego jest niespełnione i wynikowa efektywność jest niska.



Reprezentacja występujących okazjonalnie składowych wysokoczęstotliwościowych za pomocą wzorców jest związana z dodatkowym źródłem zniekształceń spowodowanych tolerancją wzajemnego dopasowania. Wymaga to ustalenia maksymalnej odchyłki ewolucji od wzorca, przy której dopuszczalne jest jej reprezentowanie przez ten wzorzec. Istnienie dodatkowego źródła zniekształceń nie wpływa jednak na wzrost ich poziomu. Procedura porównująca sygnał oryginalny z próbnie zrekonstruowanym używa już reprezentacji za pomocą wzorca, a więc kontrola poziomu zniekształceń obejmuje również proces kodowania. Konstrukcja algorytmu zapewnia, że po wyczerpaniu porównań z istniejącymi wzorcami ewolucji zakładany jest nowy wzorzec (wpis do słownika), a gdy maksymalna liczba wzorców jest wyczerpana – ewolucja jest reprezentowana dosłownie (za pomocą współczynników czasowo-częstotliwościowych) w pliku wynikowym (rys. 6.17).



Rys. 6.17. Zestaw formatów sygnału zakodowanego z użyciem słownika wysokoczęstotliwościowych wzorców w dziedzinie czasowo-częstotliwościowej

Zastosowanie słownika wysokoczęstotliwościowych wzorców w dziedzinie czasowo-częstotliwościowej skutkuje prawie dwukrotnym wzrostem efektywności kompresji. Przy zniekształceniach lokalnie tylko sięgających 5% oczekiwana efektywność kompresji wynosi 8,7 (baza MIT-BIH po przepróbkowaniu na 256 Hz, kwantyzacja 12 bitów, długość porcji sygnału 30 min). Przy porównywaniu tej metody z innymi należy jednak mieć na uwadze nie tylko ilościowe oszacowanie współczynnika kompresji i maksymalnej odchyłki sygnałów, ale także podstawową zaletę oferowaną przez niniejszy algorytm, jaką jest możliwość określenia nie zniekształconych stref sygnału i interwałów, w których zniekształcenia są dopuszczalne. Oczywiście modyfikacja funkcji chwilowego pasma prowadzi do zmiany efektywności kompresji, tak więc ograniczanie długości stref zniekształceń pogorszy redukcję objętości pliku wynikowego. Jest to jednak zgodne z określonym na wstępie priorytetem kontroli obok poziomu także czasowej dystrybucji zniekształceń.

### 6.4.3. Realizacja algorytmu kompresji z użyciem pakietów falkowych

Algorytm pozwalający na kontrolę dystrybucji zniekształceń w przypadku zastosowania dekompozycji *Wavelet Packets* także wykorzystuje rzuty wyznaczonych wcześniej granic załamków P, QRS i T na reprezentację czasowo-częstotliwościową. Zasadą jego działania jest użycie zmodyfikowanej definicji entropii, w myśl której poszczególne próbki reprezentacji nie są już traktowane równoprawnie, ale ważone w zależności od położenia względem punktów granicznych załamków. Podobnie jak w przypadku dekompozycji falkowej, wykorzystywana jest do tego związana z położeniem załamków funkcja lokalnej istotności zapisu, której wartości mogą być przedmiotem modyfikacji wprowadzonych przez użytkownika.

Wyjaśnienie idei kontroli dystrybucji zniekształceń w dekompozycji *Wavelet Packets* należy rozpocząć od przypomnienia, że poszukiwanie najlepszego drzewa dekompozycji w przeważającej większości przypadków ma na celu minimalizację entropii

$$E = \sum_i P(s_i) \quad (6.5)$$

czyli koncentrację jak największego odsetka energii sygnału w jak najmniejszej liczbie współczynników. W powyższym wzorze  $P$  oznacza operator funkcyjny (logarytm, kwadrat, wartość bezwzględna itp.).

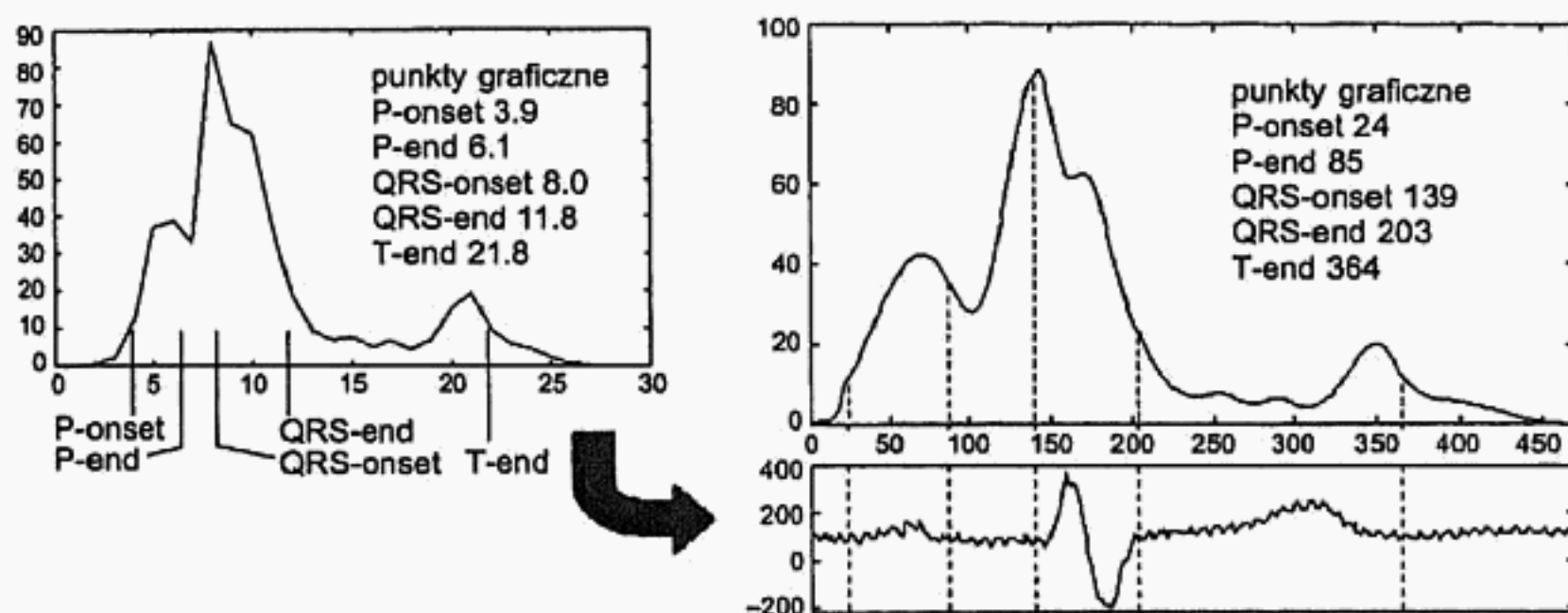
W obecnie opisywanym zastosowaniu chodzi jednak nie tyle o energetyczną bezstratność przekształcenia, ile o wierność reprezentacji informacji diagnostycznej. Ponieważ czasowy rozkład gęstości tej informacji jest skorelowany z położeniem załamków elektrokardiogramu, poszczególne próbki sygnału (a także reprezentacji w każdym węźle drzewa dekompozycji) powinny być przetwarzane z uwzględnieniem ich

położenia względem punktów granicznych załamków. Może to zostać zrealizowane za pomocą funkcji czasu współbieżnej z elektrokardiogramem zawierającej współczynniki wagowe  $w_i$  dla kolejnych próbek sygnału. W przypadku zapisu wielokanałowego funkcja taka jest wspólna dla wszystkich kanałów.

W celu obliczenia entropii kolejne próbki sygnału są dyskryminowane lub faworyzowane stosownie do ich położenia, poprzez wartość dyskretnej funkcji ważącej  $w(t)$  zawierającej współczynniki  $w_i$

$$E = \sum_i P(s_i \cdot w_i) \quad (6.6)$$

Funkcja ważąca jest obliczana dla każdej ewolucji serca bezpośrednio po określeniu granic załamków. Obliczenia polegają na odcinkowym skalowaniu standardowej funkcji istotności do interwałów wyznaczonych w bieżącej ewolucji. Skalowanie takie może się w najprostszym wariantie dokonać z użyciem interpolacji liniowej, jednak znacznie lepsze rezultaty odzwierciedlające rzeczywistą zmienność dystrybucji informacji diagnostycznej uzyskuje się dopiero przy zastosowaniu interpolacji z użyciem funkcji sklepanych 3. stopnia (rys. 6.18).



Rys. 6.18. Skalowanie standardowej funkcji lokalnej istotności elektrokardiogramu do funkcji ważącej udział poszczególnych próbek w entropii sygnału podczas poszukiwania najlepszego drzewa dekompozycji

Omówienia wymaga jeszcze sposób skalowania funkcji wagi dla węzłów położonych niżej w drzewie dekompozycji. Zarówno sygnał, jak i funkcja wagi są początkowo uzupełniane próbkami aż do długości, która jest możliwa do wyrażenia za pomocą całkowitej potęgi liczby 2. Podczas dekompozycji sygnał podlega decymacji, podobna procedura jest też stosowana wobec funkcji wagi. Dzięki temu próbki sygnału i funkcji wagi pozostają zawsze zsynchronizowane niezależnie od skalowania reprezentacji w poszczególnych węzłach drzewa.

Oczekiwane własności transformacji czasowo-częstotliwościowej pozwalają na szeroki wybór spośród wielu klasycznych metod:

- konieczne jest użycie transformacji bezstratnej (opartej na dekompozycji ortonormalnej), gdyż warunkuje to prawidłową rekonstrukcję (dekompresję) sygnału;
- transformacja powinna używać filtrów o możliwie krótkim nośniku i wprowadzających minimalne zniekształcenia fazowe.

Transformacja użyta w opisywanym algorytmie kompresji wykorzystuje biortogonalne filtry Daubechies 5. rzędu i algorytm dekompozycji piramidowej (patrz rys. 6.12). Poszukiwanie najlepszego drzewa dekompozycji przebiega z użyciem definicji entropii jako sumy logarytmu ważonej energii

$$E = \sum_i \log(s_i^2 \cdot w_i) \quad (6.7)$$

Podstawowymi składnikami skompresowanego elektrokardiogramu są zatem:

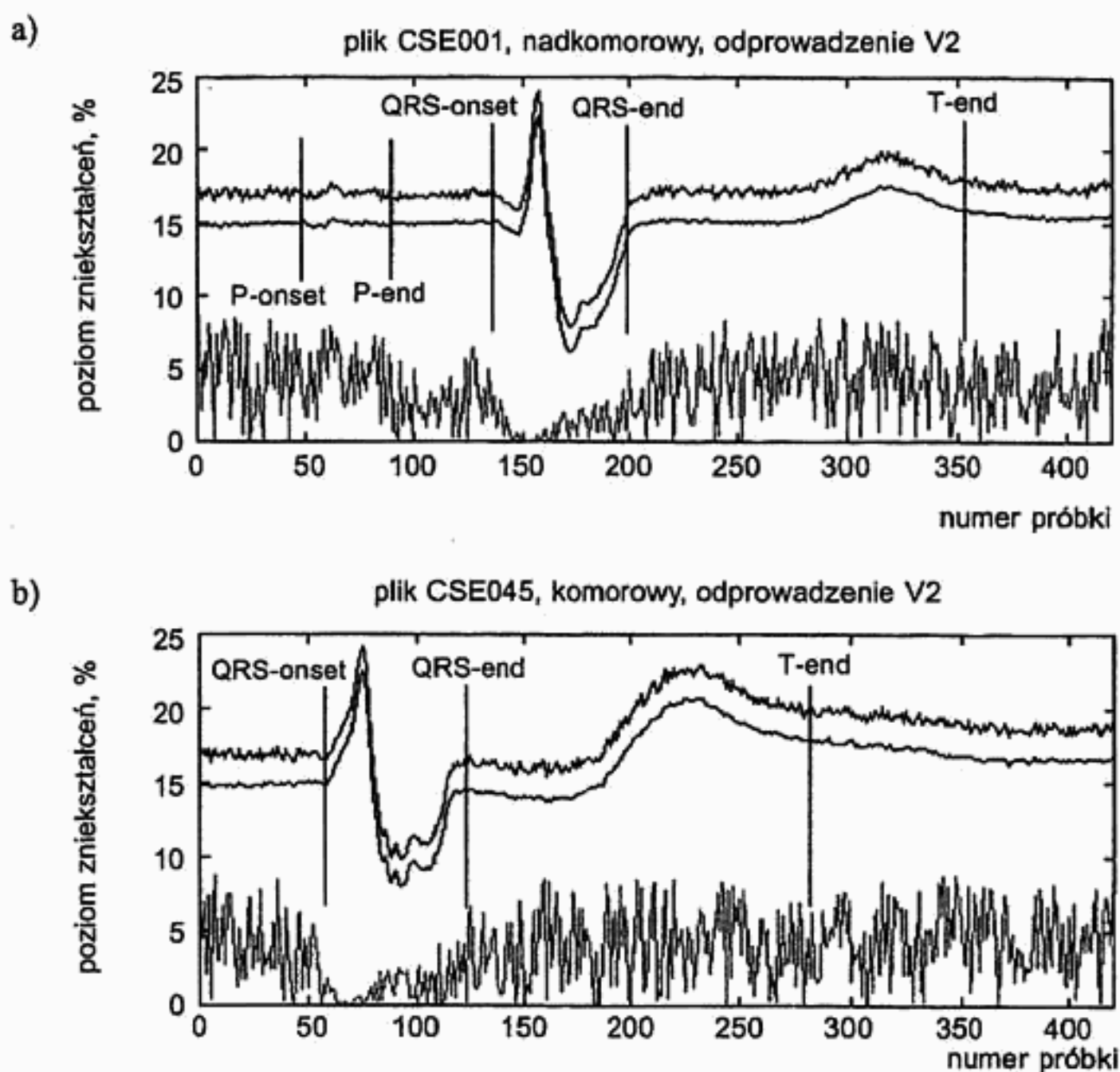
- wartości najlepszego drzewa dekompozycji,
- pozycje węzłów najlepszego drzewa dekompozycji,
- parametry synchronizacji i skali.

Pozycje węzłów najlepszego drzewa dekompozycji są zakodowane jako liczby 16-bitowe. Wartości współczynników dekompozycji, oryginalnie wyrażone liczbami rzeczywistymi, w efekcie zmiennoprzecinkowej transformacji falkowej i skalowania podlegają kwantyzacji i – analogicznie jak sygnał oryginalny – są reprezentowane w 12-bitowym formacie całkowitym. Kwantyzacja jest wprawdzie źródłem dodatkowych zniekształceń, ale biorąc pod uwagę koncentrację energii w niewielu węzłach najlepszego drzewa dekompozycji, błąd rekonstrukcji nie przekracza błędu kwantyzacji występującego podczas akwizycji sygnału. Dzięki operacji dodatkowego skalowania amplitudy, komponenty najlepszego drzewa zawsze wykorzystują pełny zakres dynamiki oferowany przez 12-bitową reprezentację danych.

Rezultaty testów tego algorytmu z użyciem bazy CSE (częstotliwość próbkowania 500 Hz, kwantyzacja 12 bitów) wykazują średnią efektywność kompresji 5,9 przy zniekształceniach PRD na poziomie 5,2%. Podobnie jak w przypadku algorytmu omawianego w podrozdziale 6.4.2 najistotniejszą zaletą jest dystrybucja zniekształceń do mniej diagnostycznie istotnych interwałów zapisu EKG. Wyróżnienie istotności diagnostycznej jest podane na podstawie testów wrażliwości parametrów diagnostycznych na lokalne zniekształcenia sygnału, ale może być przedmiotem modyfikacji użytkownika.

Przykłady kompresji ewolucji serca pochodzenia zatokowego oraz komorowego zostały przedstawione na rysunku 6.19. Górny zapis przedstawia sygnał oryginalny, a środkowy – sygnał zrekonstruowany po kompresji. Dolny przebieg reprezentuje lokalną procentową odchyłkę sygnałów i zarazem wyjaśnia, jak zniekształcenia zostały wyeliminowane ze stref o największym znaczeniu diagnostycznym.





**Rys. 6.19.** Oryginalny sygnał odprowadzenia V2, zrekonstruowany sygnał odprowadzenia V2 oraz czasowa dystrybucja zniekształceń uśredniona dla wszystkich 12 odprowadzeń. Przykłady dla ewolucji serca pochodzenia nadkomorowego (a) oraz komorowego (b)

## ROZDZIAŁ 7

# Odszumianie i modelowanie tła

Rozdział ten zawiera przykłady rozwiązań problemu identyfikacji i dyskryminacji szumu w dziedzinie czasowo-częstotliwościowej. Identyfikacja szumu możliwa jest dzięki odmiennym własnościom statystycznym źródeł szumu (pod tym pojęciem rozumiane są także sygnały ze wszystkich postronnych źródeł fizjologicznych) i źródła sygnału, którym jest obserwowany proces. W niektórych przypadkach globalny opis statystyczny sygnału jest niewystarczający do separacji szumu. Konieczne jest wtedy sformułowanie i użycie założeń na temat przebiegu obserwowanego procesu i lokalnych własności sygnału. Przykładowo dla elektrokardiogramu posiadana *a priori* wiedza mogąca znaleźć zastosowanie w odszumianiu to sekwencyjne występowanie załamków, narzucone przez fizjologię ograniczenia interwałów czasowych między nimi oraz spodziewana lokalna szerokość pasma częstotliwości sygnału EKG.

Dziedzina reprezentacji czasowo-częstotliwościowej jest dwuwymiarowa, co stanowi główny jej atut w porównaniu z dziedziną czasu oraz dziedziną częstotliwości. Komponenty dyskretnego ciągu próbek, z trudem separowalne w dziedzinie jednowymiarowej, mogą zostać łatwiej zidentyfikowane i zdyskryminowane na płaszczyźnie.

### 7.1. Redukcja reprezentacji czasowo-częstotliwościowej w poszczególnych pasmach

#### 7.1.1. Jednostajne progowanie reprezentacji czasowo-częstotliwościowej

Jedną z najwcześniej stosowanych metod redukcji czasowo-częstotliwościowych reprezentacji sygnału jest jednostajne progowanie (metoda Donoho). Polega ono na eliminacji współczynników falkowych, których wartość nie przekracza uprzednio zdefiniowanej wartości progowej. Regułą jest stosowanie indywidualnych wartości progowych w każdym zakresie częstotliwości zależnych od globalnej energii sygnału w tym zakresie. Eliminacja progowa jest związana z redukcją informacji zawartej w reprezentacji sygnału, toteż odwrotna transformacja falkowa nie dokona dokładnej rekonstrukcji sygnału oryginalnego. Zakładając, że sygnał użyteczny jest reprezentowany przez

niewielką liczbę komponentów o znaczącej energii, a także poprawny dobór wartości progowych, można tak zdefiniowany proces redukcji uznać za odszumianie sygnału w dziedzinie czasowo-częstotliwościowej. Zwraca przy tym uwagę podobieństwo metod kompresji sygnału i odszumiania, mimo że w tym ostatnim przypadku redukcja informacji jest prowadzona w innym celu.

Ze względu na sposób przetwarzania sygnału przez algorytm odszumiający można rozróżnić progowanie twarde i miękkie (rys. 7.1).

- Progowanie twarde polega na podstawieniu wartości zerowej za każdy współczynnik nie przewyższający wartości progowej

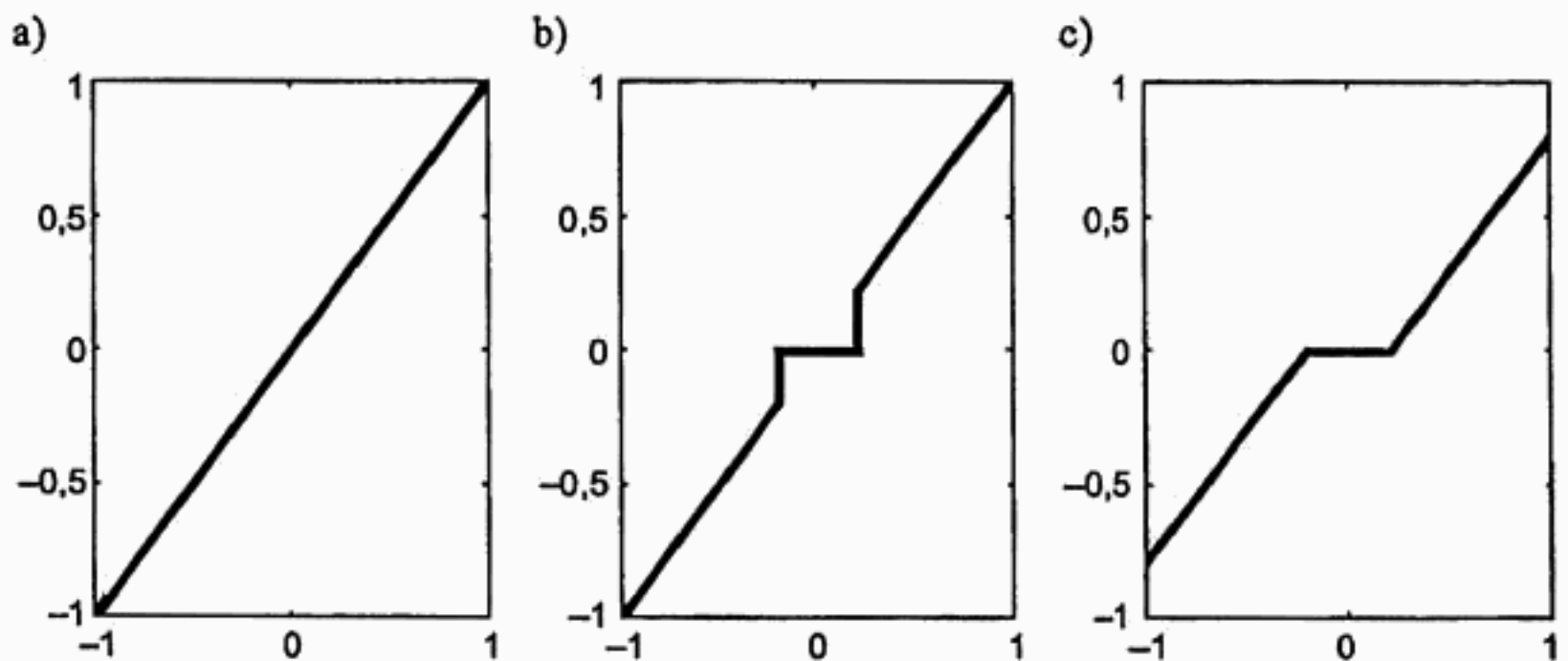
$$y(t) = \begin{cases} x(t) & \forall t : |x(t)| > \varepsilon \\ 0 & \forall t : |x(t)| \leq \varepsilon \end{cases} \quad (7.1)$$

Wadą progowania twardego jest brak w reprezentacji sygnału odszumionego wartości z otoczenia zera.

- Progowanie miękkie jest rozszerzeniem progowania twardego o dodatkową operację odjęcia wartości progowej od wszystkich zachowanych współczynników

$$y(t) = \begin{cases} \text{sign}(x(t)) \cdot (|x(t)| - \varepsilon) & \forall t : |x(t)| > \varepsilon \\ 0 & \forall t : |x(t)| \leq \varepsilon \end{cases} \quad (7.2)$$

co eliminuje nieciągłość zbioru wartości występującą w progowaniu twardym, kosztem znacznej odchyłki energii sygnału odszumionego od energii sygnału oryginalnego.



Rys. 7.1. Charakterystyka przekodowania współczynników reprezentacji czasowo-częstotliwościowej: a) bez progowania; b) progowanie twarde; c) progowanie miękkie; wartości progu przyjęto równe 20% w celach poglądowych

Zasadniczym elementem warunkującym optymalny wzrost stosunku sygnału do szumu przy jednoczesnym uniknięciu zniekształceń sygnału użytecznego jest poprawny wybór wartości progowych. Wybór ten zwykle jest poprzedzony analizą zawartości czasowo-częstotliwościowej reprezentacji sygnału, analizą spektralnych własności źródeł szumu oraz porównaniem reprezentacji sygnałów o zróżnicowanym stosunku sygnału do szumu. Porównanie takie pozwala ustalić komponenty powierzchni czasowo-częstotliwościowej reprezentujące główne składniki sygnału użytecznego i poprawnie, odrębnie dla każdej wartości współczynnika skali, obliczyć wartość progową jako odsetek całkowitej energii reprezentacji w odpowiadającym paśmie częstotliwości. Jedną z propozycji autorstwa D.L. Donoho opiera wybór wartości progowej na zróżnicowanej wariancji sygnału i szumu

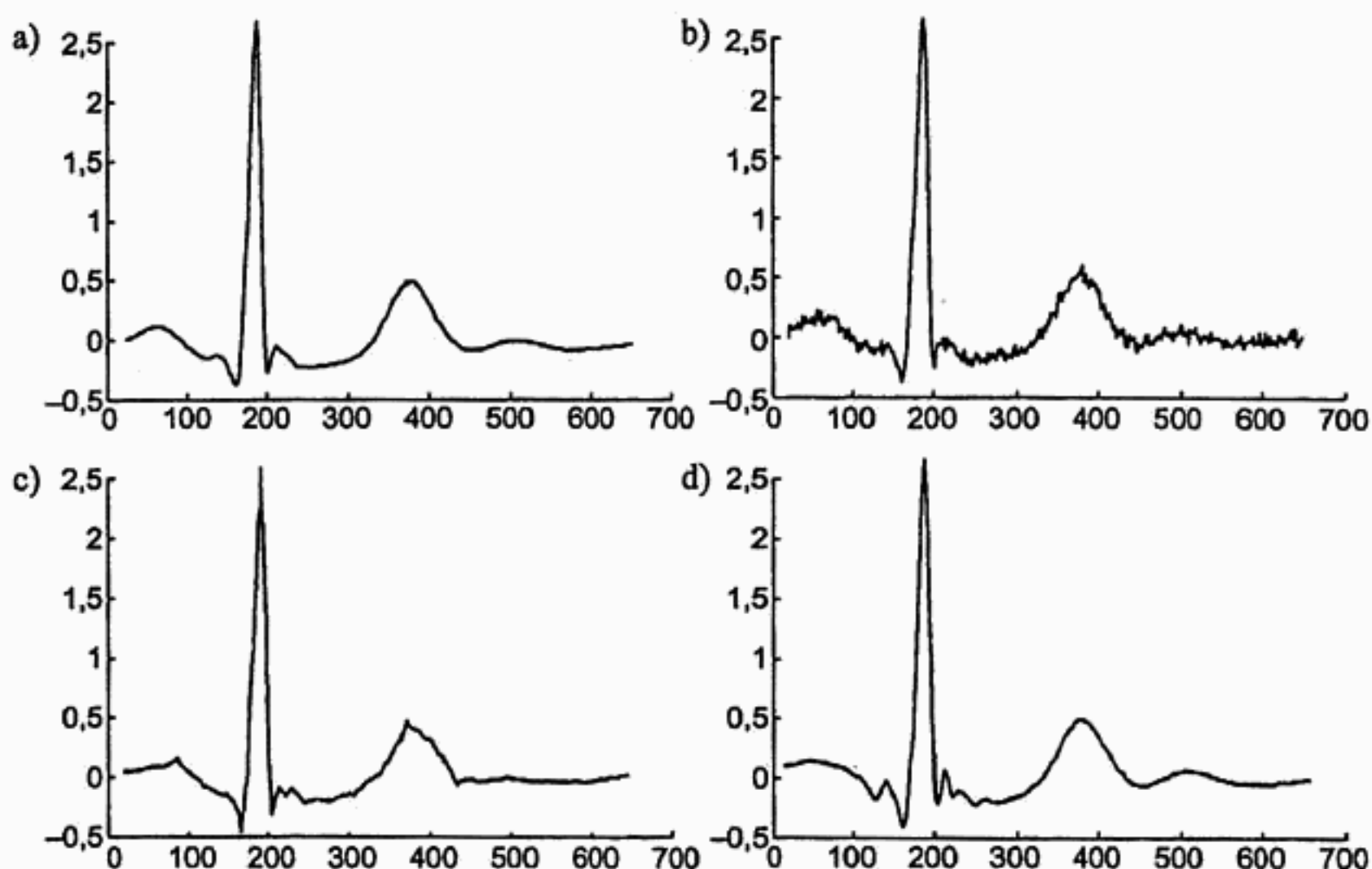
$$\varepsilon = \sigma \cdot \sqrt{2 \cdot \log(n)} \quad (7.3)$$

gdzie  $\sigma^2$  jest obliczane z wariancji szumu, a  $n$  jest liczbą próbek sygnału.

Progowanie reprezentacji czasowo-częstotliwościowej w celu eliminacji szumu jest binarną operacją decyzyjną dyskryminującą wybrane (uznane za szum) komponenty sygnału. Warto w tym miejscu poświęcić chwilę uwagi na analizę konsekwencji błędnych decyzji. Niezależnie od tego, czy omyłkowo usunięto komponent sygnału, czy też pozostawiono komponent szumu, istnienie nieciągłości nieliniowej funkcji decyzyjnej spowoduje pojawienie się w sygnale zrekonstruowanym w dziedzinie czasu zakłóceń przypominających kształtem funkcję skalującą. Ponieważ w praktyce energia reprezentacji czasowo-częstotliwościowej szumu jest rozproszona, należy poprzez odpowiedni dobór kształtu falki analizującej dążyć do jak najmniejszego rozproszenia energii sygnału użytecznego. Skupienie energii sygnału w niewielu współczynnikach o znacznych wartościach, np. poprzez wybór falki-matki zbliżonej kształtem do analizowanego przebiegu, ułatwi dyskryminację szumu. Przeciwnie, rozproszenie energii sygnału na wiele współczynników reprezentacji zmniejszy odstęp sygnału od szumu w tej dziedzinie. Jeżeli rozproszenie to jest spowodowane przez analizę z wykorzystaniem falek niskich rzędów, a więc o nieregularnym przebiegu czasowym, dodatkowo zniekształcenia będą widoczne nie tylko jako oscylacje, ale również jako nieciągłości w zrekonstruowanym sygnale czasowym.

Zniekształcenia spowodowane odszumianiem mogą być w pewnych warunkach znacznie bardziej uciążliwe niż szum. Przykładem jest elektrokardiogram, w którym nieciągłości o kształcie falki-matki są przyczyną błędów automatycznego określenia długości załameków (podczas interpretacji manualnej ekspert bez wahania odrzuci taki sygnał jako niewiarygodny). Inny przykład, również z dziedziny elektrokardiografii, ilustruje wystąpienie oscylacji, które mogą być przyczyną fałszywie pozytywnej detekcji późnych potencjałów komorowych (rys. 7.2).





Rys. 7.2. Przykłady zniekształceń sygnału EKG spowodowanych odszumianiem w dziedzinie czasowo-częstotliwościowej: a) sygnał surowy; b) sygnał zaszumiony; c) sygnał odszumiony z użyciem falki Coiflets 1 (pozostałości funkcji bazowej); d) sygnał odszumiony z użyciem falki Coiflets 5 (pasożytnicze oscylacje)

### 7.1.2. Adaptacyjne progowanie reprezentacji czasowo-częstotliwościowej

Jednostajne progowanie reprezentacji czasowo-częstotliwościowej używa stałej w czasie wartości progu dyskryminacji. Jest to słuszne w przypadku, gdy charakterystyka źródeł szumu jest niezmienna, a spodziewany sygnał charakteryzuje równomierna gęstość informacji. Najczęściej jednak progowanie jednostajne wykorzystuje się dlatego, że nie wymaga analizy czasowych charakterystyk szumu ani żadnych założeń odnośnie sygnału. W takich przypadkach uzyskiwany stopień poprawy stosunku sygnału do szumu jest nieoptymalny. Alternatywnym rozwiązaniem jest adaptacyjne progowanie reprezentacji czasowo-częstotliwościowej. Algorytm wykorzystuje co najmniej jedną z czasowych charakterystyk rozkładu (informacji w sygnale, energii szumu lub obie) do obliczenia zmiennej w czasie wartości progowej  $\varepsilon(t)$ .

Przypadek gdy znana jest czasowa zmienność charakterystyki szumu (zakłóceń), ma miejsce podczas analizy sygnału będącego odpowiedzią wystymulowaną. Dobrym przykładem jest analiza wzrokowych potencjałów wywołanych. Jeżeli pobudzenie było pojedynczym błyskiem, reakcja siatkówki oka powoduje powstanie dwóch grup odpowiedzi: reprezentującej olśnienie odpowiedzi fotopowej i reprezentującej treść

obrazu odpowiedzi skotopowej. W niektórych typach stymulacji udaje się te odpowiedzi rozgraniczyć: poprzez stosowanie stałego natężenia światła, co dyskryminuje bodziec fotopowy, poprzez stosowanie różnych barw albo poprzez dyskryminację czasową (odpowiedź fotopowa pojawia się zwykle jako pierwsza). Niestety w wielu przypadkach obie odpowiedzi są nałożone w czasie, a zainteresowany tylko jedną z nich użytkownik z konieczności traktuje drugą jako szum. W tym przypadku stosowanie progowania adaptacyjnego w dziedzinie czasowo-częstotliwościowej jest uzasadnione, gdyż znana jest charakterystyka czasowa sygnału zakłócającego. Jednocześnie, dzięki modyfikacji progu czasowo-częstotliwościowa reprezentacja sygnału zostanie zredukowana tylko w tych chwilach w których jest to konieczne. Zapewnia to oczekiwaną skuteczność odszumiania przy minimalnych zniekształceniach treści informacyjnej.

Analogiczne rozwiązanie okazuje się skuteczne w przypadku sygnału o zmiennej gęstości informacyjnej. Na wstępie należy tu rozróżnić pojęcie gęstości informacyjnej i energii – często sygnały elektrodiagnostyczne są z medycznego punktu widzenia bardziej istotne w tych fragmentach, które akurat nie niosą wysokiej energii. W tych przypadkach, stała wartość progu odcięcia ustalona w oparciu o kryterium amplitudowe może doprowadzić do dyskryminacji istotnych diagnostycznie komponentów zapisu. Jeżeli położenie czasowe takich komponentów jest znane lub może być wyznaczone automatycznie, modyfikacja progu dyskryminacji polegająca na lokalnym jego obniżeniu pozwala zachować oryginalną postać sygnału. Przykładem takiego zastosowania jest odszumianie elektrokardiogramu, w którym podejrzewane jest występowanie mikropotencjałów komorowych. Ponieważ są to zjawiskaisko niskoenergetyczne o wysokiej częstotliwości, a dodatkowo położone w bezpośrednim sąsiedztwie najwyższego w sygnale zespołu QRS, stosowanie jednostajnego progowania w dziedzinie czasowo-częstotliwościowej zniekształca ich przebieg w stopniu uniemożliwiającym detekcję. Tymczasem nieskomplikowana adaptacja funkcji progu dla zakresów częstotliwości 70÷250 Hz w interwale, którego początek następuje 50 ms, a koniec 170 ms po wierzchołku zespołu QRS pozwala uniknąć dyskryminacji tych delikatnych, lecz medycznie bardzo istotnych współczynników czasowo-częstotliwościowej reprezentacji elektrokardiogramu. W obrębie tak zakreślonego prostokąta na powierzchni czasowo-częstotliwościowej, fizjologicznie właściwego dla czasu i częstotliwości występowania mikropotencjałów, filtracja sygnału jest zabroniona, co zapewnia maksymalną wierność reprezentacji tych zjawisk.

Podane przykłady oczywiście nie wyczerpują wszystkich możliwych zastosowań progowania adaptacyjnego. Zaletą tego typu progowania i jego główną przewagą nad progowaniem jednostajnym jest faktyczne wykorzystanie czasowo-częstotliwościowej reprezentacji sygnału. Progowanie jednostajne nie wykorzystuje bowiem aspektu czasowego, czego najlepszym dowodem jest uzyskanie identycznych rezultatów odszumiania po zastosowaniu progowania w dziedzinie częstotliwościowej (progowanie transformaty Fouriera).

## 7.2. Modelowanie tła EKG z ekstrapolacją szumu izolinii

Rejestracja elektrokardiogramu coraz rzadziej odbywa się w ściśle zdefiniowanych warunkach laboratoryjnych. Wymaganiem współczesnej kardiologii jest przede wszystkim zapis całodobowego obrazu aktywności serca oraz w niektórych przypadkach zapis próby wysiłkowej polegający na obciążeniu organizmu wysiłkiem fizycznym. Rejestracja prowadzona w warunkach codziennego życia pacjenta stawia nowe wymagania algorytmom analizy oraz procedurom identyfikacji i eliminacji zakłóceń. Zakłócenia pochodzą zasadniczo z trzech źródeł, z których tylko jedno, aparatura pomiarowa, może być kontrolowane. Pozostałe zakłócenia o nieznanych charakterystykach mają swoje źródło w środowisku elektromagnetycznym, w którym przebywa pacjent, oraz aktywności elektrycznej pozostałych jego organów. Standardowo, źródła te reprezentowane są przez pojęcia „zakłócenia mięśniowe” i „przódźwięk sieci”, które *de facto* oznaczają cały zespół zakłóceń o trudnych do zdefiniowania charakterystykach. Wynika stąd trudność zbudowania modelu matematycznego zakłóceń w elektrokardiogramie, który byłby dostatecznie bliski rzeczywistości.

### 7.2.1. Wprowadzenie

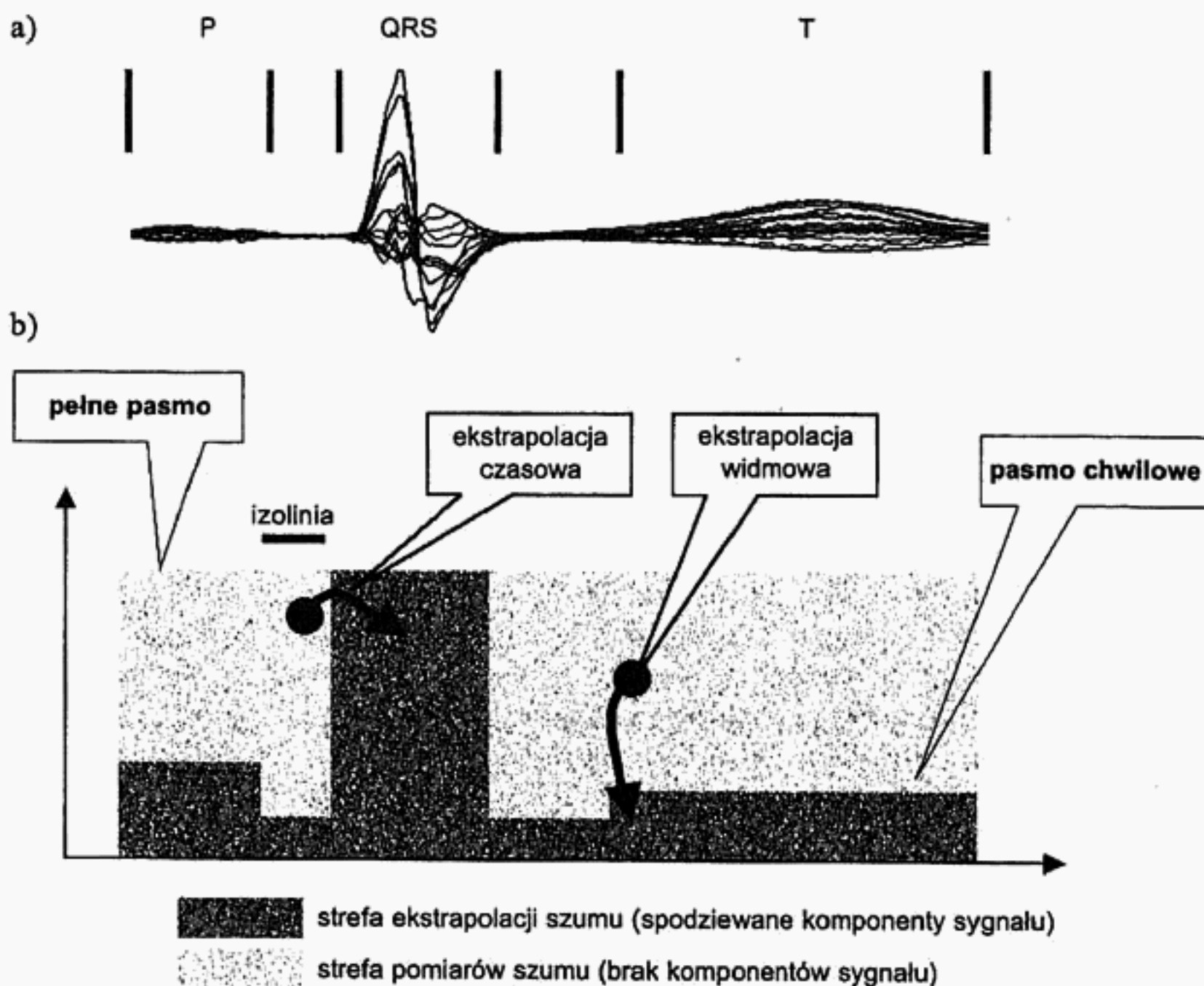
Modelowanie tła elektrokardiogramu opiera się na wykorzystaniu **niejednakowej gęstości informacji diagnostycznych** na poszczególnych odcinkach elektrokardiogramu. Wykorzystanie tych niejednorodności wymaga uprzedniego określenia funkcji chwilowej gęstości informacyjnej sygnału EKG, co jest zadaniem szczególnie trudnym wobec różnorodności zjawisk reprezentowanych przez ten sygnał. Pewnym standardem przyjętym w przetwarzaniu sygnałów jest utożsamianie gęstości strumienia informacji z widmem sygnału, co jest ściśle reprezentowane w przypadku sygnałów dyskretnych. Określenie chwilowego natężenia strumienia informacji wymaga więc zdefiniowania chwilowego widma sygnału EKG. Użycie w bezpośrednim sąsiedztwie słów: „chwilowe” i „widmo”, implikuje zastosowanie najbardziej do tego celu naturalnej reprezentacji sygnału – powierzchni czasowo-częstotliwościowej.

Algorytm ten pozwala określić widmowe parametry zakłóceń w elektrokardiogramie, co w większości zastosowań odszumiania i kompresji jest wystarczające. W wielu podobnych zastosowaniach definicja czasowej postaci szumu nie jest niezbędna. Ponieważ zaproponowany model szumu jest ciągły i parametry uaktualniane są na bieżąco, możliwe jest także zbudowanie czasowo-częstotliwościowej postaci modelu szumu (rys. 7.3).

Do estymacji parametrów widmowych szumu wykorzystano odcinek referencyjny tzw. linii izoelektrycznej pomiędzy końcem załamka P, a początkiem zespołu QRS, który reprezentuje udokumentowany fizjologicznie brak aktywności elektrycznej serca. Niewątpliwą zaletą jest fakt, że współcześnie istniejące oprogramowanie automatycznej analizy EKG pozwala z dużą precyzją i niemal stuprocentową pewnością automatycznie wyznaczyć położenie tego odcinka. Dodatkowym atutem jest fakt, że elektrokardiogram nie zawiera składowych o częstotliwościach przekraczających 32 Hz poza



zespołem QRS, natomiast częstotliwość próbkowania jest zwykle wyższa niż 128 Hz. Istnienie dwóch oktaw nie zawierających praktycznie użytecznych składowych sygnału przez ponad 80% czasu trwania pozwala modyfikować widmowe parametry modelu szumu także poza odcinkiem izoliny. Częstotliwość uaktualniania parametrów modelu jest 4-krotnie niższa niż częstotliwość próbkowania elektrokardiogramu z wyjątkiem krótkiej, trwającej zwykle ok. 100 ms przerwy na zespół QRS. Uprawnione jest zatem nazwanie zaproponowanego modelu szumu – ciągłym.



Rys. 7.3. Ciągły model szumu elektrokardiogramu w dziedzinie czasowo-częstotliwościowej: a) elektrokardiogram; b) strefy sygnału i szumu w płaszczyźnie czasowo-częstotliwościowej

### 7.2.2. Szczegóły implementacji

Wybór użytej transformacji czasowo-częstotliwościowej został dokonany w oparciu o analizę pożądanych własności transformacji, takich jak:

- wykorzystywanie filtrów o nośniku zwartym, gdyż od długości niezerowej części filtrów zależała długość atomów powierzchni czasowo-częstotliwościowej i precyzja lokalizacji energii w dziedzinie czasu; ponieważ celem jest zdefiniowanie



własności elektrokardiogramu (pasma chwilowego) w oparciu o wyznaczone wcześniej punkty charakterystyczne załamków, celowe wydaje się zatrzymanie dekompozycji po osiągnięciu przez atomy powierzchni czasowo-częstotliwościowej długości porównywalnej z długościami załamków;

- jak najlepsza separacja sąsiednich zakresów częstotliwościowych, opisywana jako stromość nachylenia zboczy widmowych charakterystyk filtrów – zrealizowana poprzez wybór falki-matki o jak najbardziej stromych zboczach charakterystyki widmowej;
- ortogonalność funkcji filtrów zapewniająca ściśle odwzorowanie informacji oraz zachowanie energii sygnału podczas przekształcenia go do dziedziny czasowo-częstotliwościowej.

Transformacja użyta w opisywanym algorytmie kompresji wykorzystywała biortogonalne filtry Daubechies 5. rzędu i algorytm dekompozycji piramidowej. Dekompozycja sygnału o oryginalnej częstotliwości próbkowania 500 Hz jest przeprowadzana na trzech poziomach, co umożliwia wydzielenie oktafów o pasmach:  $250 \div 125$  Hz,  $64 \div 125$  Hz i  $32 \div 64$  Hz. Poniżej 32 Hz sygnał oryginalny jest zachowywany, gdyż:

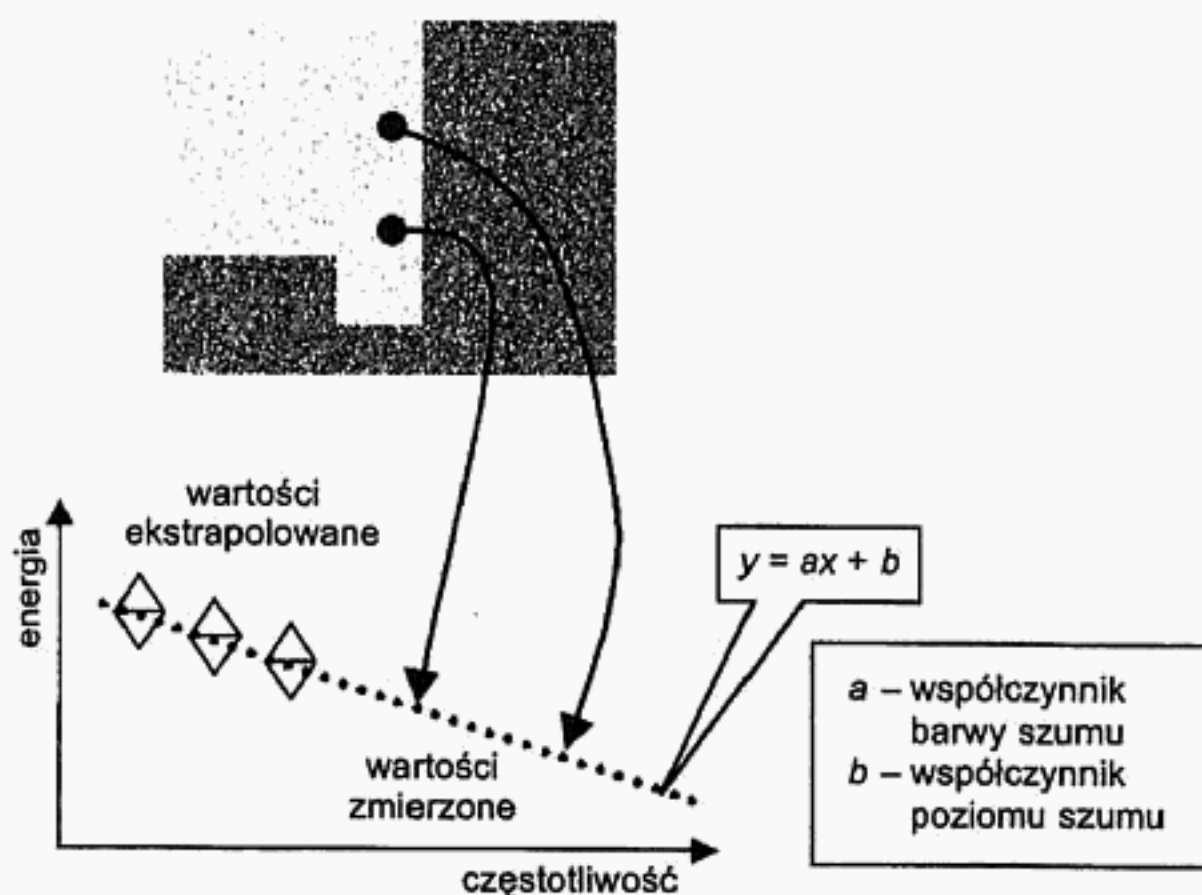
- przeprowadzone badania wykazały mały związek zawartości informacyjnej sygnału z załamekami elektrokardiogramu, istotne komponenty sygnału użytecznego występują w tym zakresie zarówno w obrębie załamków, jak i poza nimi;
- efektywna częstotliwość próbkowania wynosi 64 Hz i odpowiadający jej okres (16 ms) jest porównywalny z długością załamków;
- trzy pasma częstotliwościowe estymacji parametrów widmowych szumu uznano za wystarczające do określenia parametrów modelu (poziomu i barwy).

Dla celów pomiaru szumów w górnych zakresach częstotliwości przyjęto, że sygnał linii izoelektrycznej (pomiędzy końcem załamka P a początkiem zespołu QRS lub około 30 ms przed początkiem zespołu QRS w przypadku skurczów pochodzenia komorowego) stanowi zapis udokumentowanego braku aktywności elektrycznej serca. Sygnały na linii izoelektrycznej reprezentują zatem pozakardiologiczne zjawiska elektryczne rejestrowane wraz z elektrokardiogramem (zakłócenia i szumy). Przy założeniu stacjonarności tych zjawisk, a przynajmniej braku korelacji z aktywnością serca, stanowią one podstawę do określenia poziomu szumów w sygnale załamków.

Ponieważ linia izoelektryczna jest stosunkowo krótka (znormalizowana długość jest równa 32 ms, co odpowiada 16 próbkom sygnału), pomiar szumów może przebiegać tylko dla pasm częstotliwości powyżej 32 Hz. Niestety niewielka liczba próbek izolinii w poszczególnych pasmach częstotliwości wyklucza stosowanie jakichkolwiek testów statystycznych w celu określenia istotności różnic reprezentacji czasowo-częstotliwościowych załamków i reprezentacji izolinii

Ponieważ zapis linii izoelektrycznej jest uznawany za udokumentowaną fizjologicznie reprezentację braku aktywności elektrycznej serca, rozkład widmowy energii szumów, zmierzony w zakresie wyższych oktafów na izolinii może być podstawą do eks-

trapolacji widma szumów dla niskich częstotliwości. Wynikająca z częstotliwości próbkowania sygnału i z długości linii izoelektrycznej głębokość dekompozycji pozwala na zmierzenie maksymalnie trzech punktów lokalnej widmowej charakterystyki szumów. Te punkty są następnie podstawą do wyznaczenia krzywej ekstrapolacji (najwyżej 2. stopnia) przybliżającej poziom szumów w zakresie niższych oktaw. Zasadę działania ekstrapolacji przedstawia rysunek 7.4.



Rys. 7.4. Szczegóły ekstrapolacji widmowej parametrów modelu szumu

Ekstrapolację częstotliwościową poziomu szumów na zakres niskich częstotliwości można również przedstawić za pomocą wykresu widmowego.

Ekstrapolacja wymaga uwzględnienia następujących warunków:

- jednostka czasu jest inna dla każdego zakresu częstotliwościowego, dla obliczenia poziomu szumów dla czwartej oktawy należy użyć dwóch uśrednionych wartości zmierzonych w oktawie trzeciej, czterech uśrednionych wartości zmierzonych w oktawie drugiej oraz ośmiu uśrednionych wartości zmierzonych w oktawie pierwszej – jest to wynikiem zmiennej rozdzielczości czasowo-częstotliwościowej zastosowanej transformacji falkowej;
- zmienna jednostka częstotliwości wynika ze zmiennej rozdzielczości czasowej zastosowanej transformacji czasowo-częstotliwościowej, zatem na wykresie liniowym widma ekstrapolacja będzie prowadzona wzdłuż krzywej wykładniczej;
- ponieważ tylko linia izoelektryczna zapewnia reprezentację braku aktywności i może stanowić odcinek odniesienia, w pozostałych częściach ewolucji serca należy przyjąć wartości poziomu szumów ekstrapolowane na podstawie odcinków izolacji.

Wahania widmowych własności szumów na kolejnych izoliniiach uznano za miarę stacjonarności sygnału szumu, określają one wiarygodność, z jaką wartości zmierzone na izolii można ekstrapolować w czasie na pozostałe segmenty ewolucji serca.

Współczynniki prostej  $y = ax + b$  określającej lokalne własności widmowe szumów noszą następujące nazwy:

- współczynnik kierunkowy  $a$  to współczynnik barwy, gdyż definiuje stosunek składowych nisko- i wysokoczęstotliwościowych widma szumu;
- stała  $b$  to współczynnik poziomu, gdyż określa lokalną wartość energii szumów.

W najprostszej wersji modelu szumu założeniem jest aktualizacja parametrów tylko na izolii, co określa maksymalną częstotliwość aktualizacji – raz na uderzenie. Ograniczona długość izolii uniemożliwia dekompozycję sygnału poniżej poziomu, na którym długość atomu powierzchni czasowo-częstotliwościowej przekracza 16 ms. W przypadku bazy CSE próbkowanej z częstotliwością 500 Hz jest to poziom trzeci. Ilość poziomów dekompozycji, równa ilości pasm częstotliwościowych, wyznacza natomiast maksymalną liczbę punktów będących podstawą wyznaczania lokalnych własności widmowych szumu. Mając do dyspozycji trzy punkty można rozważyć estymację chwilowego widma szumu z użyciem funkcji kwadratowej i liniowej. Ponieważ brak dostatecznie umotywowanych przesłanek dla konieczności użycia funkcji kwadratowej, zaimplementowano jedynie estymację liniową w oparciu o chwilową energię sygnału w dwóch najwyższych oktawach, pozostawiając na przyszłość analizę ewentualnych korzyści płynących z zastosowania estymacji wyższego rzędu. Dodatkowym argumentem technicznym jest tutaj fakt, że częstotliwość próbkowania elektrokardiogramów długoczasowych rzadko przekracza 250 Hz, pozostawiając nie więcej niż dwie oktawy zapasu pomiędzy pasmem zarejestrowanym a pasmem sygnału użytecznego.

Zmierzony w strefach udokumentowanego braku aktywności serca poziom szumu elektrokardiogramu jest następnie ekstrapolowany na sąsiednie interwały czasowe. Pomiar bezpośredni nie jest tam możliwy, gdyż w interwałach tych sygnał wykorzystuje pełną szerokość pasma, toteż współczynniki powierzchni czasowo-częstotliwościowej są uważane za reprezentację sygnału użytecznego. Niemniej, przy założeniu krótkookresowej stabilności źródła szumu estymacja jego parametrów metodą ekstrapolacji jest poprawna.

Ekstrapolacja poziomu szumów na sąsiednie interwały czasowe może być zrealizowana:

- jako ekstrapolacja liniowa,
- z wykorzystaniem funkcji sklepanych.

Ekstrapolacja liniowa jest najprostszym obliczeniowo sposobem wyznaczenia poziomu szumów w punktach, w których bezpośredni pomiar nie jest możliwy. W przypadku szumu jednostajnego wartości ekstrapolowane nie odbiegają od wartości prawdziwych. Weryfikacji takiej można dokonać poprzez dodatkowo wykonaną ekstrapolację wartości na odcinkach pomiarowych (np. izolii), a następnie – porównanie



z wartościami zmierzonymi. Z powodu zmiennej energii szumu, a taką obserwujemy zwykle w przypadku źródeł rzeczywistych, wartości przewidziane z użyciem ekstrapolacji liniowej znacznie odbiegają od wartości zmierzonych. Nieciągłości powstające na styku kolejnych odcinków ekstrapolacji sugerują, że charakterystyka źródła szumu zmienia się skokowo, co zakłóca prawidłowy jego opis zwłaszcza w dziedzinie widmowej.

Opisana wada ekstrapolacji liniowej może być wyeliminowana przy zastosowaniu funkcji sklejaných trzeciego stopnia, co niestety związane jest ze znacznym wzrostem nakładów obliczeniowych. Ekstrapolację tę stosuje się odrębnie do każdego z pasm częstotliwościowych reprezentujących sygnał na płaszczyźnie czasowo-częstotliwościowej. Każde z pasm traktowane jest w tym przypadku jak sygnał w dziedzinie czasu charakteryzowany przez częstotliwość próbkowania zależną od wartości współczynnika skali. Każdy z takich sygnałów jest niekompletny, ponieważ pomiar poziomu szumu jest możliwy tylko w tych interwałach czasu – właściwych odrębnie dla każdego z pasm – w których komponenty elektrokardiogramu nie są spodziewane. Próbkowanie modelu szumu położone w interwałach zawierających komponenty elektrokardiogramu nie są uwzględniane, a więc sygnały reprezentujące każde z pasm są traktowane jako próbkowane nieregularnie. Zastosowanie aproksymacji funkcją sklejaną trzeciego stopnia ma za zadanie odtworzenie wartości brakujących próbek na podstawie wartości, w których poziom szumu został zmierzony. W ten sposób zmienność charakterystyki modelowanego źródła szumu jest bardziej płynna (wszystkie pochodne są ciągłe), co bardziej odpowiada szumom indukowanym i generowanym w układach pomiarów biopotencjałów.

### **7.2.3. Własności algorytmu odszumiania z ekstrapolacją szumu izolinii**

Głównymi celem eksperymentalnej weryfikacji algorytmu odszumiania jest określenie stabilności parametrów widmowych modelu szumu przy zaszumieniu elektrokardiogramu testowego szumem stacjonarnym o wartościach 10, 25 i 50% amplitudy zespołów QRS. Źródłem sygnałów testowych była kardiologiczna baza danych CSE zawierająca 10-sekundowe odcinki sygnału zarejestrowane z częstotliwością próbkowania 500 Hz. Zaletą tej bazy jest udostępnianie rezultatów wstępnego przetwarzania zapisów potwierdzonych przez kardiologów – w tym także punktów początkowych i końcowych załamków.

Stabilność określono jako wariancję współczynników barwy i poziomu szumu. Otrzymane wartości wynoszą odpowiednio: 27% i 21%.

Przeprowadzony dodatkowo eksperyment polegający na porównaniu (pomiar PRD) sygnału oryginalnego, zaszumionego (jak poprzednio 10, 25 i 50% amplitudy zespołów QRS) i odszumionego z użyciem zaproponowanego ciągłego modelu szumu w dziedzinie czasowo-częstotliwościowej, pozwolił określić skuteczność i użyteczność zaproponowanego modelu do odszumiania. Średnia wartość PRD zapisów odszumio-



nych w stosunku do zapisów oryginalnych wyniosła 17,8% średniej wartości PRD zapisów zaszumionych w stosunku do zapisów oryginalnych. Oznacza to, że sygnały odsumione były ponad pięciokrotnie bliżej oryginału niż sygnały zaszumione.

Próba zastosowania ekstrapolacji kwadratowej w dziedzinie częstotliwości prowadzi do znacznych wahań własności widmowych szumów z uderzenia na uderzenie. Ponieważ nie znajduje to uzasadnienia w zmianie jakości sygnału, jedynym wytłumaczeniem tego zjawiska może być przypadkowe przesunięcie w fazie pozycji funkcji analizującej i komponentów sygnału. Wpływu tego przesunięcia można uniknąć przez uśrednianie większej liczby współczynników falkowych na izolinii, jest to jednak możliwe tylko przez rezygnację z pomiarów szumu w trzeciej oktawie (na długości izolinii przypadają tam zwykle dwa współczynniki) i oparciu estymacji na dwóch punktach pomiaru szumu uśrednionych ze współczynników falkowych oktawy drugiej (4 wartości) i pierwszej (8 wartości). Tak przeprowadzona ekstrapolacja liniowa charakteryzuje się znaczną stabilnością własności widmowych (barwy) szumów z uderzenia na uderzenie.

## Zakończenie

Jako datę narodzin transformacji falkowej powszechnie przyjmuje się moment wydania publikacji A. Grossmanna i J. Morleta (lipiec 1984 r.), zatem obchodziła ona niedawno swoje osiemnaste urodziny. I pomimo że niektóre koncepcje są znacznie starsze, wtedy właśnie starannie pielęgnowane narzędzie opuściło zacisza pracowni matematycznych. Dzięki ogromnej przydatności praktycznej szybko zaczęło zdobywać uznanie inżynierów różnych specjalności, by na długo przed osiągnięciem pełnoletności stać się jednym z filarów nowoczesnych metod przetwarzania sygnałów.

Transformacja falkowa oraz transformacje korzystające z jej rozwiązań były wśród najbardziej dynamicznie rozwijanych narzędzi analizy w dekadzie 1984–1993. W tym okresie trudno znaleźć choć rok, w którym nie ukazałaby się publikacja stanowiąca kamień milowy na drodze rozwoju tej metody. Następnie przyszedł czas na zastosowania, najpierw w zakresie analizy, a kolejno w dziedzinach coraz bardziej zaawansowanego przetwarzania sygnałów. Obecnie transformacja falkowa jest coraz częściej spotykana w złożeniach z innymi metodami przetwarzania i rozpoznawania sygnałów: innymi transformacjami, metodami statystycznymi, sieciami neuronowymi i metodami widmowymi. Nic dziwnego, jako pełnoletnia, nie chce już dłużej pojawiać się samotnie.

Falki prawie od swych narodzin są obecne w aplikacjach biomedycznych, a zwłaszcza są wykorzystywane jako efektywne narzędzie obróbki zapisów elektrodiagnostycznych. To właśnie w sygnałach generowanych przez narządy naszego organizmu wielki jest udział składników niestacjonarnych, w których metody falkowe wykazują szczególnie swą przydatność. Wiele parametrów diagnostycznych reprezentowanych jest w niuansach zapisu aktywności serca i mózgu, wcześniej rozróżnianych jedynie wprawnym okiem doświadczonego eksperta. Niuansach, które mogą niekiedy decydować o życiu.

Jednak to nie tradycyjny podział technik elektrodiagnostycznych, ale właśnie historia rozwoju teorii i zastosowań rodziny transformacji falkowych wyznacza kolejność rozdziałów tego podręcznika. Rozdziały pierwszy i czwarty mówią o podstawach: o zapleczu teoretycznym i najprostszych, a jednocześnie najszerszych w sensie obszaru i najlepiej ugruntowanych zastosowaniach. Rozdziały drugi i trzeci poświęcone są rozszerzeniom metod falkowych i technicznym problemom implementacji przekształceń

falkowych w maszynie cyfrowej. Rozdział piąty podaje przykłady przetwarzania sygnałów na powierzchni czasowo-częstotliwościowej, których efektem jest zmodyfikowana postać czasowa sygnału. Rozdział szósty przedstawia metody efektywnego kodowania sygnału w dziedzinie czasowo-częstotliwościowej i wreszcie siódmy wprowadza w zagadnienia detekcji i dyskryminacji szumu wykorzystującej zmienność widma chwilowego sygnału,

Z racji swej dojrzałości, czasowo-częstotliwościowa reprezentacja sygnału jest już na równych prawach obecna w wielu zastosowaniach technicznych. Ciąg próbek detali wybranej skali traktowany jak niezależny sygnał dyskretny, operacje statystyczne w obrębie gałęzi drzewa dekompozycji czy też transformacja falkowa przetworzonego wstępnie sygnału wciąż jednak kryją tajemnice jeszcze nie zbadane.

Intencją autora jest zainspirowanie obecnych i przyszłych projektantów i użytkowników metod czasowo-częstotliwościowych. Zadanie to zostanie spełnione, jeżeli właśnie Państwo, szanowni Czytelnicy, dopiszecie dalsze rozdziały tej książki.

## Dodatek

Dodatek zawiera uproszczone opisy transformacji wykraczających poza główny wątek rozważań przedstawionych w książce. Stanowią one doraźną pomoc dla bardziej dociekliwego Czytelnika, zanim sięgnie On do bardziej wyczerpującej literatury.

### D.1. Próbkowanie, ograniczanie i warianty transformacji Fouriera

Próbkowanie związane jest z ograniczeniem pasma częstotliwości od góry, gdyż twierdzenie Shannona zapewnia pełną rekonstrukcję sygnału o częstotliwości sięgającej połowy częstotliwości próbkowania. Niejednoznaczność rekonstrukcji następuje już dla przypadku częstotliwości sygnału równej częstotliwości próbkowania, gdyż amplituda sygnału zrekonstruowanego wynika ze wzajemnego przesunięcia w fazie sygnału i taktu próbkowania. Rzeczywiste sygnały charakteryzują się trudną do zdefiniowania szerokością pasma, jako że oprócz składowej użytecznej o przewidywalnych własnościach uwzględnienia wymagają także zakłócenia, które ze swej natury charakteryzuje przypadkowość. Wystąpienie w próbkowanym sygnale składowych o częstotliwościach wyższych niż połowa częstotliwości próbkowania prowadzi do zjawiska aliasingu, polegającego na odwzorowaniu (odbiciu) zakresu  $(f_p/2 \div f_p)$  w zakres  $(0 \div f_p/2)$ . Dysponując ciągiem próbek, nie można już rozróżnić składowych obu zakresów, dochodzi więc do zniekształcenia informacji zawartych w paśmie częstotliwości  $(0 \div f_p/2)$ . Wynika stąd konieczność stosowania filtrów dolnoprzepustowych (antyaliasingowych) zapewniających skuteczne tłumienie składowych, których obecność mogłaby zagrozić jednoznaczności odwzorowania. Stosowanie tych filtrów jest ściśle zalecane nawet wtedy, gdy z charakterystyki źródła sygnału użytecznego wynika niemożliwość przekroczenia połowy częstotliwości próbkowania.

Ograniczanie długości zarejestrowanego sygnału jest związane z ograniczaniem od dołu zakresu rejestrowanych częstotliwości. Jest to bezpośrednia konsekwencja zasady nieoznaczoności Heisenberga, która mówi, że do określenia częstotliwości potrzebny jest odcinek sygnału o długości co najmniej odpowiadającej okresowi tej częstotliwości. Określenie zawartości energii składowej widma możliwe jest tylko dla



składowych będących wielokrotnością tej najniższej częstotliwości reprezentowanej w sygnale, gdyż dla wszystkich wartości pośrednich funkcja analizująca  $e^{j\omega t}$  nie scala się do zera na odcinku równym długości sygnału. Wynika z tego, że widmo sygnału ograniczonego istnieje tylko dla pewnych częstotliwości, których okres zmieści się w długości sygnału całkowitą ilość razy, jest więc widmem dyskretnym. Najniższa częstotliwość reprezentowana w sygnale (równa odwrotności jego czasu trwania) jest jednocześnie równa kwantowi częstotliwości widma dyskretnego.

Wzajemną jednoznaczność reprezentacji sygnału w dziedzinie czasu i częstotliwości najlepiej ukazuje przekształcenie Fouriera w różnych odmianach (rys. D.1):

- ciągłe przekształcenie Fouriera (sygnał jest nieograniczony i ciągły oraz widmo jest nieograniczone i ciągłe);
- ciągły szereg Fouriera (sygnał jest ograniczony i ciągły, a widmo jest nieograniczone i dyskretne);
- dyskretne przekształcenie Fouriera (sygnał jest nieograniczony i dyskretny, a widmo jest ograniczone i ciągłe);
- dyskretny szereg Fouriera (sygnał jest ograniczony i ciągły oraz widmo jest ograniczone i ciągłe).

## D.2. Transformacja Karhunen–Loevego

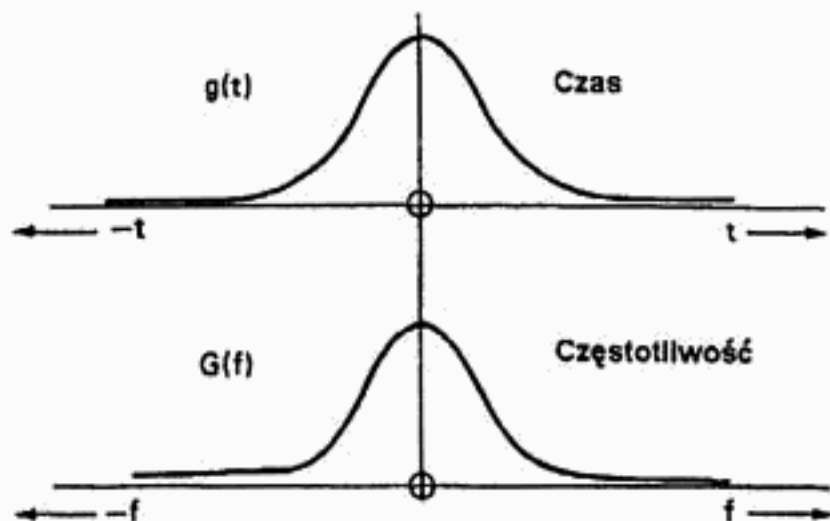
Metody reprezentacji danych z użyciem minimalnej liczby parametrów przy założonym poziomie zniekształceń bazują na takim przekształceniu układu współrzędnych, aby kierunki osi pokrywały się z głównymi cechami reprezentowanymi w sygnale. Podstawowe z tych metod to analiza składowych głównych PCA (ang. *principal component analysis*) oraz analiza składowych niezależnych ICA (ang. *independent component analysis*). Obie z nich należą do tzw. technik multiwariacyjnych, czyli wykorzystujących współzmiennność danych w więcej niż jednym wymiarze. Analiza składowych głównych jest obecnie dobrze znaną metodą redukcji wymiarowości danych, a ponieważ główne składowe reprezentowane są w początkowych wymiarach, możliwa jest łatwa eliminacja nieistotnych składowych sygnału.

Z analizą składowych głównych związana jest transformacja Karhunen–Loevego (KLT), zwana też optymalną transformacją liniową, przekształcająca oryginalną reprezentację sygnału w nowy układ współrzędnych, w którym reprezentacja jest optymalna. Sama transformacja nie zmienia liczby wymiarów, ale w nowym układzie współrzędnych całkowita entropia sygnału jest zminimalizowana, więc często okazuje się, że istotność informacji reprezentowanych przez dalsze wymiary jest pomijalna. Dzieje się tak wtedy, gdy w reprezentacji oryginalnej informacje w poszczególnych wymiarach są skorelowane. Dlatego transformacja Karhunen–Loevego zwana jest także dekorelacją sygnału. Transformacja jest wprowadzicie zdefiniowana dla ciągłych procesów stacjonarnych, ale jej użycie do wolnozmiennych sygnałów pseudookresowych jest także dopuszczalne.

a)

$$G(f) = \int_{-\infty}^{\infty} g(t) e^{-j 2\pi f t} dt$$

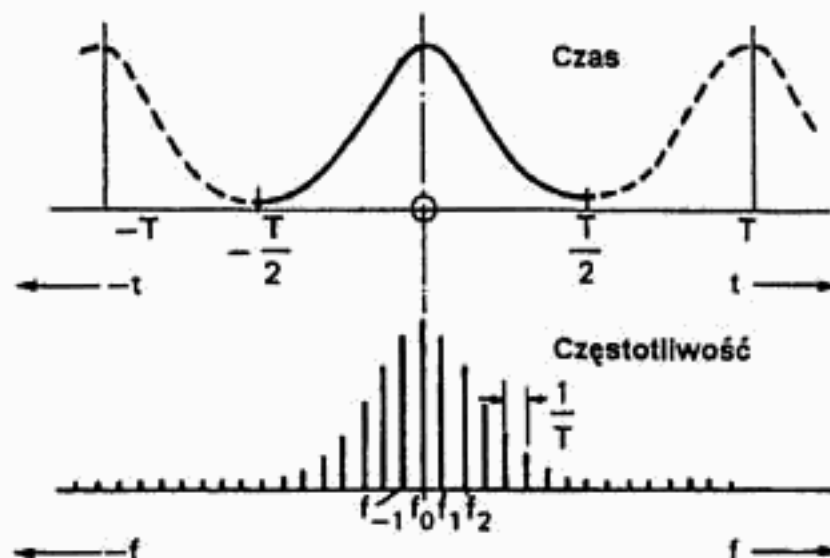
$$g(t) = \int_{-\infty}^{\infty} G(f) e^{j 2\pi f t} df$$



b)

$$G(f_k) = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} g(t) e^{-j 2\pi f_k t} dt$$

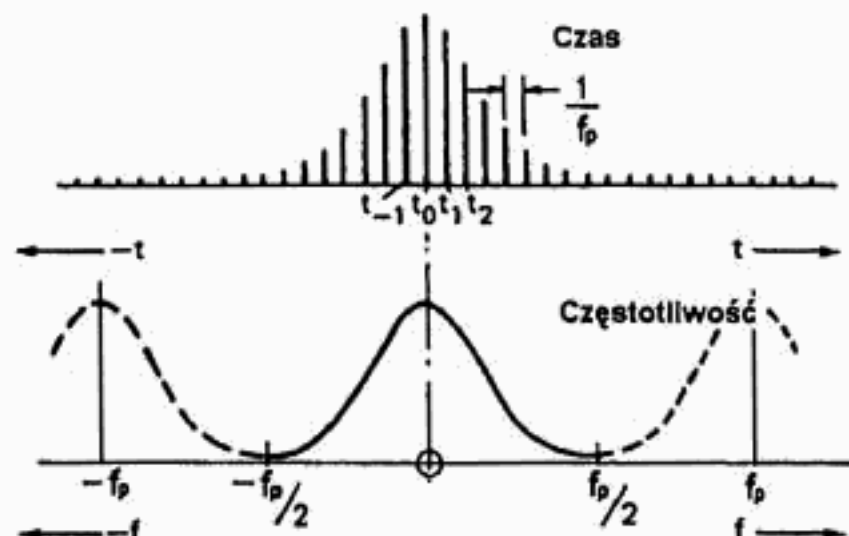
$$g(t) = \sum_{k=-\infty}^{\infty} G(f_k) e^{j 2\pi f_k t}$$



c)

$$G(f) = \sum_{n=-\infty}^{\infty} g(t_n) e^{-j 2\pi f t_n}$$

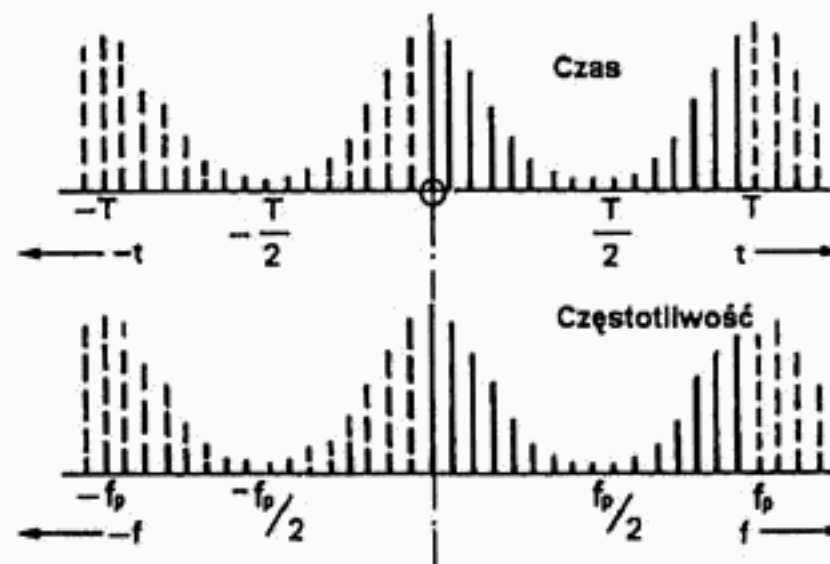
$$g(t_n) = \frac{1}{f_s} \int_{-f_s/2}^{f_s/2} G(f) e^{j 2\pi f t_n} df$$



d)

$$G(f_k) = \frac{1}{N} \sum_{n=0}^{N-1} g(t_n) e^{-j \frac{2\pi n k}{N}}$$

$$g(t_n) = \sum_{k=0}^{N-1} G(f_k) e^{j \frac{2\pi n k}{N}}$$



**Rys. D.1.** Ciągłe i dyskretne wersje transformacji Fouriera: a) ciągłe przekształcenie Fouriera; b) ciągły szereg Fouriera; c) dyskretne przekształcenie Fouriera; d) dyskretny szereg Fouriera

Transformacja polega na liniowej kombinacji  $n$  współczynników macierzy  $A$  i próbek sygnału oryginalnego  $X_n = [x_1(n), \dots, x_C(n)]^T$

$$Y_n = A \cdot X_n \quad (D.1)$$

We wzorze tym  $Y_n = [y_1(n), \dots, y_C(n)]^T$  jest reprezentacją sygnału w dziedzinie składowych głównych, natomiast  $A$  jest kwadratową macierzą transformacji o boku  $C$ , gdzie  $C$  reprezentuje liczbę wektorów danych (ewolucji serca lub kanałów rejestracji).

Pierwszym krokiem jest obliczenie macierzy autokowariancji  $V_x$  zespołu danych  $X_n$

$$V_X = \frac{1}{M} \sum_{i=0}^{M-1} \begin{bmatrix} x_1(i) \\ \vdots \\ x_C(i) \end{bmatrix} \begin{bmatrix} x_1(i) & \dots & x_C(i) \end{bmatrix} \quad (D.2)$$

gdzie  $M$  jest liczbą próbek w wektorze danych, a  $C$  jest liczbą takich wektorów. Wektory własne macierzy  $V_x$  są nazywane składowymi głównymi i tworzą bazę transformacji wpisane jako wiersze macierzy transformacji  $A$ . Kolejne wartości własne macierzy  $V_x$  są natomiast miarą wariancji zawartej przez odpowiadające im wektory własne.

Składowa główna, związana z największą wartością własną, reprezentuje kierunek w przestrzeni danych, w którym wariancja danych jest największa. Składowe główne są wzajemnie ortogonalne, ale nie są całkowicie niezależne. Problem statystycznej niezależności rozwiązuje dopiero analiza składowych niezależnych zakładająca zerową korelację wzajemną nie tylko drugiego rzędu, ale i wszystkich wyższych rzędów.

Transformacja Karhunen–Loevego, mimo że dokonuje optymalnej dekorelacji sygnałów, jest niechętnie używana w praktycznych implementacjach. Powodem jest szereg problemów technicznych, z których najważniejszym jest, paradoksalnie, adaptacja macierzy  $A$  do przetwarzanej porcji sygnału. Najważniejsze z nich wymieniono poniżej:

- konieczne jest przechowywanie macierzy transformacji  $A$  wraz z sygnałem (jej wartości są odrębnie właściwe dla każdej porcji sygnału);
- nie ma algorytmu szybkiej transformacji KLT (złożoność obliczeniowa dla macierzy  $N \times N$  wynosi  $N^3$ );
- przetwarzanie dłuższych sekwencji sygnału zmniejsza proporcje udziału macierzy transformacji w pliku wyjściowym, ale także założenia stacjonarności są coraz gorzej spełnione;
- dziedziną macierzy transformacji jest zbiór liczb rzeczywistych – w celu otrzymania reprezentacji stałoprzecinkowej można albo zaokrąglić jej wartości, co prowadzi do niedokładnej rekonstrukcji sygnału, albo obliczyć suboptymalną macierz o wartościach całkowitych, co z kolei psuje własności dekorelacji i wymaga dodatkowych nakładów obliczeniowych.

### D.3. Dyskretna transformacja kosinusowa

Alternatywne podejście do problemu dekorelacji polega na wykorzystaniu dyskretnej transformacji kosinusowej DCT (ang. *discrete cosine transform*). Wejściowym zestawem próbek  $X = [x_1(i) \ \dots \ x_C(i)]$  jest ciąg próbek o tych samych numerach we wszystkich  $C$  wektorach danych. Dla takiej sekwencji DCT jest zdefiniowana jako

$$G(k) = \sqrt{\frac{\alpha_k}{C}} \sum_{s=0}^{C-1} x(s) \cdot \cos \frac{(2 \cdot s + 1) \cdot k \cdot \pi}{2 \cdot C} \quad (\text{D.3})$$

gdzie  $k = 0, 1, 2, \dots, C-1$ ,  $\alpha_k = 1$  dla  $k = 0$ , a dla pozostałych  $k$   $\alpha_k = 2$ .

Transformacja kosinusowa jest transformacją suboptymalną i w granicy zaledwie przybliża rezultaty dekorelacji osiągane za pomocą KLT. Niemniej, dzięki wysokiej efektywności obliczeniowej i łatwości implementacji to ona jest powszechnie stosowana do realizacji dekorelacji sygnałów.

Do zalet DCT należą:

- użycie standardowej funkcji cosinus, co wyklucza konieczność każdorazowego obliczania i przechowywania macierzy transformacji;
- możliwość obliczania za pomocą szybkiego algorytmu o złożoności  $(N \log N)$ ;
- realizacja sprzętowa przez niektóre procesory sygnałowe.



## Bibliografia

- [1] Akay M. (ed.): *Time Frequency and Wavelets in Biomedical Signal Processing*. New Jersey, IEEE Press 1998
- [2] Akay M.: *Wavelets in Biomedical Engineering*. *Annals of Biomed. Eng.*, vol. 23, 1995, 531–542
- [3] Aminoff M. (ed.): *Electrodiagnosis in Clinical Neurology*. New York, Churchill Livingstone 1986
- [4] Augustyniak P.: *Przetwarzanie sygnałów elektrodiagnostycznych*. Kraków, UWND AGH 2001
- [5] Augustyniak P.: *Przybornik falkowej analizy sygnałów*. [w:] *Mat. I Krajowej Konferencji użytkowników MATLAB-a*, Kraków 14–15.11.1995, 250–257
- [6] Augustyniak P., Tadeusiewicz R.: *Improve the Quality of Diagnostic Parameters of an Electronystagmogram Using Signal Filtration in the Time-Frequency Domain and Adaptatively Adjusted Characteristics*. [w:] *Mat. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Paris 18–21.06.1996, 381–384
- [7] Augustyniak P., Tadeusiewicz R.: *The Bandwidth Variability of a Typical Electrocardiogram*. [w:] *Mat. European Medical and Biological Engineering Conference EMBEC '99*, Wien 04–07.11.1999, 394–395
- [8] Augustyniak P.: *Odrębna kompresja wyższych oktaw elektrokardiogramu*. [w:] *Mat. konferencji „Biocybernetyka i Inżynieria Biomedyczna”*, Warszawa 02–04.12.1999, 423–427
- [9] Augustyniak P.: *Compression and Denoising of the ECG Using Standard Bandwidth Variability Function*. [w:] *Mat. First International Conference on Advances in Medical Signal and Information Processing*, Bristol 4–6.09.2000, 48–53
- [10] Augustyniak P.: *Optimal Pattern-to-Signal Synchronization for Time-Frequency Template Matching ECG Compression Technique*. [w:] *Mat. Vth International Conference on Medical Informatics and Technologies MIT 2000*, Ustroń 8–10.11.2000, CS41–CS48
- [11] Augustyniak P.: *The Dynamic Range of an ECG in the Time-Frequency Domain Used for the Lossless Signal Compression*. [w:] *Mat. Vth International Conference on Medical Informatics and Technologies MIT 2000*, Ustroń 8–10.11.2000, BI29–BI37

- [12] Augustyniak P.: *Frequency Band Dependent Quantization Level for Adaptive ECG-Dedicated Signal Compression*. [w:] Mat. International Conference on Image and Signal Processing, Agadir 3–5.05.2001, 1013–1025
- [13] Augustyniak P.: *Controlling the Distortions Distribution in a Wavelet Packet Based ECG Compression*. [w:] Mat. International Conference on Image and Signal Processing, Agadir 3–5.05.2001, 267–277
- [14] Barlas G.D., Skordalakis E.S.: *A Novel Family of Compression Algorithms for ECG and Other Semiperiodical One-Dimensional Biomedical Signals*. IEEE Trans. Biomed. Eng., vol. 43, No. 8, 1996
- [15] Białasiewicz J.T.: *Falki i aproksymacje*. Warszawa, WNT 2000
- [16] Bhaskaran V., Konstantinides K.: *Image and Video Compression Standards: Algorithms and Architectures*. Boston, Kluwer 1995
- [17] Blanchett T., Kember G.C., Fenton G.A.: *KLT-Based Quality Controlled Compression of Single-Lead ECG*. IEEE Trans. Biomed. Eng., vol. 45, No. 7, 1998, 942–945
- [18] Bradie B.: *Wavelet Packet-Based Compression of Single Lead ECG*. IEEE Trans. Biomed. Eng., vol. 43, 1996, 493–501
- [19] Calderbank A.R., Daubechies I., Sweldens W., Yeo B.L.: *Wavelet Transforms that Map Integers to Integers*. Techn. Rep., Department of Mathematics, Princeton University 1996
- [20] Calderbank A.R., Daubechies I., Sweldens W., Yeo B.L.: *Lossless Image Compression Using Integer to Integer Wavelet Transforms*. [w:] Mat. IEEE International Conference of Image Processing, Santa Barbara, 10.10.1997, 596–599
- [21] Cardenas-Barrera J.L., Lorenzo-Ginori J.V.: *Mean-Shape Vector Quantizer for ECG Signal Compression*. IEEE Trans. Biomed. Eng., vol. 46, No. 1, 1999, 62–70
- [22] Cetin A.E., Koymen H., Aydin M.C.: *Multichannel ECG Data Compression by Multirate Signal Processing and Transform Domain Coding Techniques*. IEEE Trans. on Biomed. Eng., vol. 40, No. 5, 1993, 495–499
- [23] Chen J. Itoh S.: *A Wavelet Transform-Based ECG Compression Method Guaranteeing Desired Signal Quality*. IEEE Trans. on Biomed. Eng., vol. 45, No. 12, 1998, 1414–1419
- [24] Chen L., Chen W., Itoh S.: *Lossy to Lossless Compression of ECG Data Using Integer Wavelet Transform*. [w:] Mat. European Medical & Biological Engineering Conference EMBEC'99, Vienna 4–7.11.1999, 1610–1611
- [25] Chui Ch.K.: *An Introduction to Wavelets*. New York, Academic Pres 1992
- [26] Chui Ch.K.: *Wavelets: A Mathematical Tool for Signal Analysis*. Philadelphia SIAM 1997
- [27] Cohen A.: *Ondelettes, Analyses Multirésolution et Traitement Numérique du Signal*. Ph.D. Thesis, University of Paris IX, DAUPHINE 1992
- [28] Cohen A., Zigel Y.: *Compression of Multichannel ECG Through Multichannel Long-Term Prediction*. IEEE Eng. Med. Biol., vol. 17, No. 1, 1998, 109–115

- [29] Coifman R.R., Wickerhauser M.V.: *Entropy-Based Algorithms for Best Basis Selection*. IEEE Trans. on Information Theory, vol. 38, No. 2, 1992
- [30] Daubechies I.: *Ten Lectures on Wavelets*. Philadelphia SIAM 1992
- [31] Daubechies I., Sweldens W.: *Factoring Wavelet Transforms into Lifting Steps*. J. Fourier Anal. Appl., vol 4, No. 3, 1998, 245–267
- [32] Dąbrowska B., Dąbrowski A.: *Podręcznik elektrokardiografii*. Warszawa, PZWL 1993
- [33] Dąbrowski A., Dąbrowska B., Piotrowicz R.: *Elektrokardiografia holterowska*. Warszawa–Oxford–Polska, Wydawnictwo Medyczne 1994
- [34] Djohan A., Nguyen T.Q., Tompkins W.J.: *ECG Compression Using Discrete Symmetric Wavelet Transform*. [w:] Mat. 17th Int. Conf. IEEE Medicine and Biology, 1995
- [35] Donoho D.L.: *Interpolating Wavelet Transforms*. J. of Appl. and Comput. Harmonic Analysis, 1994
- [36] Drozdek A.: *Wprowadzenie do kompresji danych*. Warszawa, WNT 1999
- [37] Duda K.: *Projektowanie biortogonalnych układów falkowych w wersji predykcyjnej z uwzględnieniem zjawisk brzegowych*. Elektronika i Elektrotechnika, t. 19, z. 2, 2000, 125–131
- [38] Duda K., Zieliński T.P.: *Zastosowanie adaptacyjnej transformaty falkowej w wersji predykcyjnej do bezstratnej kompresji sygnału EKG*. [w:] Mat. Konferencji „Modelowanie i pomiary w medycynie”, Krynica 7–11.05.2001
- [39] Duda K., Turcza P., Zieliński T.P.: *Lossless ECG Compression with Lifting Wavelet Transform*. [w:] Mat. IEEE Instrumentation and Measurement Technology Conference, Budapest 21–23.05.2001
- [40] Duda K.: *Lossless ECG Compression with Adaptive Lifting Wavelet Transform*. [w:] Mat. TICSP Workshop on Spectral Methods and Multirate Signal Processing SMMSP, Pula 16–18. 06. 2001
- [41] Durka P.J.: *Time-Frequency Analyses of EEG*. Rozprawa doktorska, Uniwersytet Warszawski 1996 (<http://durka.info/dissert.html>)
- [42] Fayn J., Rubel P.: *An Improved Method for the Precise Measurement of Serial ECG Changes in QRS Duration and QT Interval*. J. Electrocardiol., vol. 24 sup., 1992, 123–127
- [43] Heim K.: *Metody kompresji danych*. Warszawa, MIKOM 2001
- [44] Hilton M.L.: *Wavelet and Wavelet Packet Compression of Electrocardiograms*. IEEE Trans. on Biomed. Eng., vol. 44, No. 5, 1997, 394–402
- [45] Holter N.: *New Method for Heart Studies: Continuous Electrocardiography of Active Subjects over Long Periods is now Practical*. Science, 134, 1961, 1214–1220
- [46] Horspool R.N., Windels W.J.: *ECG Compression Using Ziv-Lempel Techniques*. Comput. Biomed. Res., 28, 1995, 67–68
- [47] Howard P.G., Vitter J.S.: *Practical Implementations of Arithmetic Coding*. [w:] Storer J.A. (red.), Image and Text Compression, Boston, Kluwer 1992, 85–112



- [48] Huffman D.A.: *A Method for the Construction of Minimum-Redundancy Codes*. Proceedings of the Institute of Radio Engineers, 40, 1952
- [49] Jaleleddine S., Hutchens C., Coberly W.: *ECG Data Compression Techniques – A Unified Approach*. IEEE Trans. Biomed. Eng., vol. 37, 1990, 329–343
- [50] Kuzume K., Nijima K.: *Design of Optimal Lifting Wavelet Filters for Data Compression*. New Jersey, IEEE Press 1998
- [51] *Lossless Data Compression*. Proceedings of the IEEE, vol. 88, No. 11, 2000
- [52] Lu Z., Kim D.Y., Pearlman W.A.: *Wavelet Compression of ECG Signals by the Set Partitioning in Hierarchical Trees Algorithm*. IEEE Trans. on Biomed. Eng., vol. 47, No. 7, 2000
- [53] Mallat S.G.: *A Theory for Multiresolution Signal Decomposition: The Wavelet Representation*. IEEE Trans on PAMI, vol. 11, No. 7, 1989
- [54] Mallat S.G.: *A Wavelet Tour of Signal Processing*. San Diego, Academic Press 2001
- [55] Mallat, S., Zhang, Z.: *Matching Pursuit with Time-Frequency Dictionaries*. IEEE Trans. on Signal Processing, vol. 41, 1993, 3397–3415
- [56] Misiti M., Misiti Y., Oppenheim G., Poggi J.-M.: *Wavelet Toolbox User's Guide*. The MathWorks, 1995
- [57] Moody G., Mark R.: *MIT-BIH Arrhythmia Database Directory*. Massachusetts Institute of Technology, Biomedical Engineering Center 1988
- [58] Moss A. J., Stern S. (ed.): *Noninvasive Electrocardiology – Clinical Aspects of Holter Monitoring*. W.B. Saunders Co. Ltd., Cambridge University Press 1996
- [59] Nagarajan K., Kresch E., Rao SS., Kresh Y.: *Constrained ECG Compression Using Best Adapted Wavelet Packet Bases*. IEEE Signal Processing Letters, vol. 3, No. 10, 1996, 273–275
- [60] Nave G., Cohen A.: *ECG Compression Using Long Term Prediction*. IEEE Trans. Biomed. Eng., vol. 40, 1993, 877–885
- [61] Pivik R.T., Joncas S., Busby K.A.: *Sleep Spindles and Arousal: The Effects of Age and Sensory Stimulation*. SleepResearch Online, vol. 2(4), 1999, 89–100
- [62] Ramakrishnan A.G.: *ECG Coding by Wavelet-Based Linear Prediction*. IEEE Trans. Biomed. Eng., vol. 44, No. 12, 1997, 1253–1261
- [63] Randall R.B., Tech B.A.: *Frequency analysis*. Amsterdam, Bruel & Kjaer 1987
- [64] Rechtschaffen A., Kales A.: *A Manual of Standardized Terminology, Techniques and Scoring Systems for Sleep Stages in Human Subjects*. Los Angeles, BIS/BRI, UCLA 1968
- [65] Shannon C.E.: *A Mathematical Theory of Communication*. Bell System Technical Journal, vol. 27, 1948, 379–423, 623–656
- [66] Shapiro J.M.: *Embedded Image Coding Using Zerotrees of Wavelet Coefficients*. IEEE Trans. Signal Processing, vol 41, No. 12, 1993, 3445–3462
- [67] Shaou-Gang M, Heng-Lin Y.: *Quality Driven Gold Washing Adaptive Vector Quantization and Its Application to ECG Data Compression*. IEEE Trans. Biomed. Eng., vol. 47, No. 2, 2000, 209–218



- [68] Skarbek W.: *Multimedia Algorytmy i Standardy Kompresji*. Warszawa, PLJ 1998
- [69] Sweldens W., Fernández G., Periaswamy S.: *Liftpack: A Software Package for Wavelet Transforms Using Lifting* (<ftp://ftp.math.sc.edu/pub/>)
- [70] Sweldens W.: *The Lifting Scheme: A Construction of Second Generation Wavelets*. Techn. Rep., Department of Mathematics, University of South Carolina 1995
- [71] Sweldens W.: *The Lifting Scheme: a Custom-Design Construction of Biorthogonal Wavelets*. J. of Appl. and Comput. Harmonic Analysis, vol. 3(2), 1996, 186–200
- [72] Sweldens W., Schröder P.: *Building Your Own Wavelets at Home*. [in:] *Wavelets in Computer Graphics*. ACM SIGGRAPH Course notes, 1996, 15–87
- [73] Tadeusiewicz R.: *Podstawy elektroniki medycznej*. Kraków, UWND AGH 1978
- [74] Teiwes W.: *Video-okulografie – Registrierung von Augenbewegungen in Drei Freiheitsgraden zur Erforschung und Medizinischen Diagnostik des Gleichgewichtssystems*. Dissertation TU Berlin 1991
- [75] Tkacz E.: *Nowe możliwości diagnostyczne analizy zmienności rytmu serca (HRV)*. Warszawa, prace IBIB-PAN, nr 45, 1996
- [76] Tompkins W.J. (ed.): *Biomedical Digital Signal Processing – C-languages Examples and Laboratory Experiments for the IBM PC*. New Jersey, Prentice Hall 1993
- [77] Vetterli M., Kovacevic J.: *Wavelets and Subband Coding*. New Jersey, Prentice Hall, Englewood Cliffs 1995
- [78] Villasenor J.D., Belzer B., Liao J.: *Wavelet Filter Evaluation for Image Compression*. IEEE Trans. Image Processing, vol. 4, No. 8, 1995, 1053–1060
- [79] Vitter J.S.: *Algorithm 673: Dynamic Huffman Coding*. ACM Trans. on Mathematical Software, vol. 15, 1989
- [80] Wickerhauser M.V.: *Adapted Wavelet Analysis from Theory to Software*. New Jersey, IEEE Press 1994
- [81] Willems J.L.: *Common Standard for Quantitative Electrocardiography – 10<sup>th</sup> and Final Report*. Leuven, Commission of the EC – Medical and Public Health Research, 1990
- [82] Willems J.L.: *Common Standard for Quantitative Electrocardiography Multilead Atlas – Measurements Results Data Set 3*. Leuven, Commission of the EC – Medical and Public Health Research 1988
- [83] Wojtaszczyk P.: *Teoria falek*. Warszawa, PWN 2000
- [84] Xiong Z., Herley C.: *Flexible Tree-Structured Signal Expansions Using Time-Varying Wavelet Packets*. IEEE Trans. on Signal Processing, vol. 45, No. 2, 1997, 333–345
- [85] Young R.K.: *Wavelet Theory and its Applications*. Boston, Kluwer Academic Publishers 1993
- [86] Zareba W., Maison-Blanche P., Locati E.H.: *Noninvasive Electrocardiology in Clinical Practice*. New York, Futura Publishing Co. 2001
- [87] Zigel Y., Cohen A., Katz A.: *ECG Signal Compression Using Analysis by Synthesis Coding*. IEEE Trans. Biomed. Eng., vol. 47, No. 10, 2000

- [88] Zigel Y., Cohen A., Katz A.: *The Weighted Diagnostic Distortion (WDD) Measure for ECG Signal Compression*. IEEE Trans. Biomed. Eng., vol. 47, No. 11, 2000
- [89] Zigel Y.: *ECG Signal Compression*. Thesis M.Sc. Degree, Ben Gurion University 1998
- [90] Ziv J., Lempel A.: *A Universal Algorithm for Sequential Data Compression*. IEEE Trans. on Information Theory, IT-23, 1977, 337-343
- [91] Ziv J., Lempel A.: *Compression of Individual Sequences via Variable-Rate Encoding*. IEEE Trans. on Information Theory, IT-24, 1978, 530-536
- [92] Zygierewicz J., Blinowska K.J., Durka P.J. Szelenberger W., Niemcewicz S., Androsiuk W.: *High Resolution Study of Sleep Spindles*. Electroencephalogr. Clin. Neurophysiol., vol. 110, 1999, 2136-2147s