

Universitatea "Politehnica" din Timișoara
Facultatea de Electronică și Telecomunicații
Departamentul Comunicații

*Contribuții la compresia semnalului vocal în sistemele de
comunicații numerice*

Teză de doctorat

2002

Autor Andrei Cubițchi

Conducător științific Prof. dr. ing. Ioan Naforniță

Cuprins

Capitolul 1. Introducere	1
Capitolul 2. Transformări ortogonale	4
2. 1. Câteva transformări ortogonale utilizabile la compresia de date	4
2.1.1. Legătura dintre teoria funcțiilor "wavelet" și codarea în subbenzi	5
2.1.1.1. Codarea subbandă cu reconstrucție perfectă folosind sisteme cu structură arborescentă cu filtre realizabile	13
2.1.1.1.1. Metode de proiectare a filtrelor CQF	17
2.1.2. Legătura dintre sistemele de codare în subbenzi și teoria seriilor de funcții "wavelet"	17
2.1.3. Transformarea "wavelet" discretă TUD	26
2.1.4. Pachete de funcții "wavelet"	29
2.1.4.1. Alegerea celei mai bune baze	31
2.1.4.2. Pachete de funcții "wavelet" de tip Malvar	34
2.2. Utilizarea transformărilor ortogonale prezentate la compresia de date	36
2.2.1. Analiza statistică a TUD	36
2.2.2. Analiza statistică a TPWD	47
2.2.3. Analiza statistică a TPC	51
2.3. Alegerea celei mai bune funcții de tipul "wavelets mother"	52
Capitolul 3. Detectorul de prag	57
3.1. Detecția adaptivă de prag	57
3.2. Analiza statistică a detectorului de prag	59
Capitolul 4. Sistemul de cuantizare	61
4.1. Proprietățile psiho-acustice ale semnalului de vorbire	61
4.1.1. Fenomenul de mascare	61
4.1.2. Benzi critice	63
4.1.2.1. Pragul de mascare	64
4.1.2.2. Utilizarea pragului de mascare la compresia semnalului de vorbire	65
4.2. Cuantizarea adaptivă în domeniul TPC	66
4.3. Celelalte blocuri ale sistemului de compresie	68
Capitolul 5. Simulări ale metodei de compresie	69
5.1. Primul experiment	69
5.2. Al doilea experiment	74
5.3. Al treilea experiment	75
5.4. Al patrulea experiment	79
Capitolul 6. Concluzii	83
Capitolul 7. Contributii originale	87
Bibliografie	90

Capitolul 1. Introducere

Una dintre metodele cele mai importante de prelucrare a semnalelor utilizate în telefonia numerică este compresia de date. Acest domeniu de prelucrare a semnalelor are o dinamică accentuată. În prezent se utilizează două categorii mari de tehnici de compresie: cele cu pierdere de informație și cele fără pierdere de informație. În cazul semnalelor vocale, cele care se vehiculează în telefonia numerică, având în vedere redundanța lor sporită, se preferă utilizarea tehnicilor de compresie cu pierdere de informație. Dintre acestea trebuie amintite cele bazate pe predicție (așa cum sunt cele folosite în standardul GSM), cele bazate pe transformări ortogonale, precum și cele bazate pe codarea în subbenzi.

Indiferent la care dintre aceste tehnici de compresie ne referim trebuie remarcat că ea se bazează pe aplicarea unei succesiuni de metode de bază de prelucrare a semnalelor cum ar fi: eșantionarea, cuantizarea sau codarea.

Pentru a fi eficiente, metodele de compresie, trebuie să exploateze specificul datelor ce urmează a fi comprimate. Particularitățile semnalului de vorbire sunt exploatate în standardele de compresie a vorbirii în vigoare cum ar fi de exemplu standardul GSM.

Semnalul de vorbire este un semnal foarte complex, care posedă numeroase caracteristici. În scopul compresiei este util să se sublinieze că acesta nu este un semnal staționar dar că el poate fi considerat ca local staționar pe durate de ordinul a câtorva zeci de milisecunde. Frecvent, semnalul de vorbire este împărțit în segmente de 20 ms. Fiecare astfel de segment poate fi privit ca și un semnal staționar. Compresia semnalului de vorbire este favorizată de existența unui model de producere a vorbirii simplu și eficient. Este vorba despre un model autoregresiv care presupune că semnalul de vorbire se obține prin filtrarea unui zgomot alb sau a unui tren de impulsuri Dirac folosind un filtru care nu are decât poli. Acest model se pretează la compresie prin predicție linară și stă la baza standardului GSM. Există și un model de percepție a vorbirii, numit psiho-acustic. Cunoașterea sa poate fi exploatată pentru creșterea factorului de compresie.

În scopul prelucrării semnalelor de vorbire este utilă o clasificare a acestora. În funcție de aplicație, se vorbește despre semnal de vorbire în banda telefonică, respectiv despre semnal de vorbire de bandă largă. În cazul semnalului de vorbire în banda telefonică, semnalul în timp continuu achiziționat se filtrează în banda [300, 3300] Hz și apoi se eșantionează cu frecvența de 8 kHz. Din această categorie face parte semnalul care este transmis în rețeaua telefonică publică. Există numeroase normalizări și recomandări referitoare la semnalul de vorbire în banda telefonică.

Rețeaua telefonică publică

După 1972 norma internațională G.711 precizează pentru semnalul de vorbire o cuantizare (codare) PCM (Pulse Code Modulation), corespunzătoare unui debit de 64 kbit/s (este vorba despre o cuantizare uniformă pe 8 biți). După 1984 norma G.721 a definit o cuantizare cu modulația impulsurilor în cod, diferențială, adaptivă, ADPCM (Adaptive Differential Pulse Code Modulation). Nu se mai cuantizează direct semnalul de vorbire eșantionat ci diferența dintre acesta și o variantă a sa obținută printr-o predicție bazată pe filtrare adaptivă. Debitul corespunzător acestei metode de compresie este de 32 kbit/s.

Un codor bazat pe tehnici de modelare și de cuantizare vectoriale a fost selecționat de ITU în anul 1991. Acesta corespunde unui debit de 16 kbit/s. Performanțele sale sunt specificate de norma G.728. Metoda care stă la baza construcției acestui codor se numește LD-CELP (Low Delay Code

Excited Linear Prediction Coding). Este vorba despre o codare bazată pe predicția liniară prezentând o întârziere mică de reconstrucție, proprietate foarte importantă pentru o legătură telefonică.

După 1994 s-a simțit nevoia introducerii unui codor care să aibă un debit de 8 kbit/s. La acesta se lucrează în cadrul mai multor echipe de cercetare în prezent.

Comunicații mobile

Natura canalului de transmisiuni, legătură radio, cere economisirea la maximum a benzii semnalului transmis pentru a fi posibil un număr cât mai mare de utilizatori. În 1989 a fost introdusă norma europeană GSM (Groupe Special Mobile). Această normă folosește tehnicile de acces multiplu prin divizare în timp, TDMA (Time Division Multiple Access) și codarea sursei, RPE-LTP (Regular Pulse Excitation-Long Term Prediction) făcând posibil un debit de 13 kbit/s. Rețeaua de comunicații bazată pe standardul GSM are o capacitate de 3 ori mai mare decât rețeaua analogică. După 1994 a fost introdus standardul $\frac{1}{2}$ GSM care permite construcția unei rețele cu o capacitate de 10 ori mai mare decât capacitatea unei rețele analogice corespunzătoare.

Se preconizează introducerea în viitor a unui sistem de comunicații mobile de generația a 3-a, bazat pe tehnici de acces multiplu prin divizarea codurilor CDMA (Code Division Multiple Access) care să permită debite variabile. De exemplu, societatea americană Qualcomm a realizat codorul QCELP, care selecționează dinamic din 20 în 20 de ms un anumit debit dintre valorile posibile: 8, 4, 2, și 1 kbit/s (ceea ce corespunde la un debit mediu de 4 kbit/s).

În cazul semnalului de bandă largă se face o prefiltrare în banda [50, 7000] Hz urmată de o eşantionare care folosește frecvența de 16 kHz. Acest tip de semnal de vorbire se folosește în aplicațiile care cer o calitate superioară a semnalului reconstruit.

Există două metode de compresie a acestor semnale:

- CELP,
- codarea în subbenzi (asemănătoare celei folosite la compresia muzicii).

OBSERVAȚIE. Toate debitele amintite mai sus se referă la codarea sursei. În calculul debitului efectiv al sistemului de compresie trebuie să se țină cont și de operația de codare a canalului. De aceea debitele reale necesare sunt mai mari decât cele amintite mai sus.

Pentru a realiza o compresie de calitate este necesar să se îndeplinească două deziderate: să se obțină un factor de compresie cât mai mare și semnalul reconstruit în urma *decompresiei* să aibă distorsiuni cât mai mici. Din păcate aceste două deziderate sunt antagonice, cu cât unul dintre ele este mai deplin satisfăcut cu atât celălalt este satisfăcut în mai mică măsură.

Metodele de compresie bazate pe predicție liniară au fost elaborate în urmă cu peste douăzeci de ani. Între timp teoria prelucrării semnalelor a evaluat foarte mult. De aceea scopul acestei teze este de a prezenta metode mai recente de compresie a semnalelor și de a investiga posibilitățile de utilizare a acestor metode la compresia semnalului de vorbire. Printre aceste metode, un loc aparte îl ocupă cele bazate pe utilizarea teoriei funcțiilor "wavelet". Deși de dată mai recentă, această teorie poate fi utilizată cu succes în compresia semnalelor. Cea mai bună dovadă este că foarte recent a fost elaborat un nou standard de compresie a imaginilor statice, numit JPEG 2000, bazat pe utilizarea funcțiilor "wavelet".

Această teorie a fost utilizată și pentru realizarea unor programe mai performante de căutare în bazele de date conținând amprente ale FBI-ului. Timpul de căutare a unei anumite amprente a fost redus, datorită compresiei imaginii acesteia, de la câteva ore la câteva minute.

Teoria funcțiilor "wavelet" se utilizează și la compresia semnalelor biomedicale, ca de exemplu electrocardiogramele. În prezent se lucrează și la elaborarea unor noi metode de compresie a semnalelor audio bazate pe utilizarea acestei teorii.

Iată de ce scopul acestei teze este de a prezenta un studiu al posibilităților de utilizare a teoriei funcțiilor "wavelet" la compresia semnalului de vorbire. În continuare se descrie schema unui sistem de compresie.

De obicei la compresie se utilizează transformări ortogonale, deoarece acestea sunt neredundante. Rolul transformării ortogonale este de a decorela semnalul care trebuie comprimat. În urma aplicării transformării ortogonale se obțin numeroase eşantioane de valoare foarte mică. Acestea pot fi neglijate fără a fi afectat prea mult conținutul informațional al semnalului de prelucrat. În urma efectuării acestei operații se obține un nou semnal care este cuantizat. Următoarea operație este una de codare (compresie fără pierderi) care contribuie și ea la creșterea factorului de compresie. Semnalul de la ieșirea codorului reprezintă varianta comprimată a semnalului de prelucrat.

Structura acestei lucrări este bazată pe schema de compresie descrisă mai sus. În capitolul 2 se prezintă câteva transformări ortogonale, insistându-se asupra transformărilor "wavelet" discrete. Se evidențiază efectul de decorelare al acestor transformări.

În capitolul 3 se studiază sistemele de rejecție a eşantioanelor, din domeniul transformatei, mai mici decât un anumit prag. Se prezintă o strategie de inițializare a valorii acestui prag, precum și o metodă adaptivă de alegere a valorii pragului.

În capitolul 4 se face o prezentare a metodelor de cuantizare, care pot fi folosite în aplicațiile de compresie, urmărindu-se obținerea unui echilibru între factorul de compresie, pe care-l pot produce aceste metode și volumul de calcul necesar.

Capitolul 5 este dedicat simulării unor metode de compresie a semnalului de vorbire, precum și analizei rezultatelor acestor simulări.

În capitolul 6 se prezintă concluziile acestei teze, iar în capitolul 7 principalele contribuții ale autorului.

Capitolul 2. Transformări ortogonale

2. 1. Câteva transformări ortogonale utilizabile la compresia de date

După cum s-a arătat și în introducere, schema unui sistem de compresie bazat pe utilizarea unei transformări ortogonale este cea prezentată în figura 2.1.1.

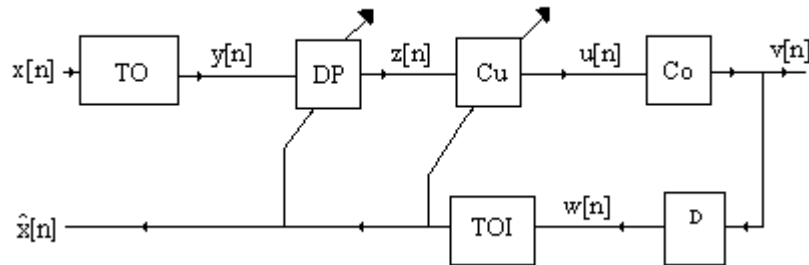


Figura 2.1.1. Schema unui sistem de compresie bazat pe utilizarea unei transformări ortogonale.

S-au utilizat următoarele prescurtări:

- TO - sistem de calcul al transformării ortogonale;
- DP - detector de prag (este sistemul care rejectează eşantioanele de valoare mică, toate eşantioanele inferioare unui prag sunt anulate);
- Cu - Sistem de cuantizare;
- Co - Sistem de codare;
- D - Sistem de decodare, inversul sistemului Co;
- TOI - Sistem de calcul al transformării ortogonale inverse celei calculate de TO.

Principalele semnale din această figură sunt: semnalul de prelucrat $x[n]$, semnalul obținut în urma compresiei $v[n]$ și semnalul reconstruit, obținut prin efectuarea operației de decompresie, $\hat{x}[n]$.

Eşantioanele semnalului $x[n]$ sunt corelate. Asta înseamnă că de informația conținută în eşantionul curent sunt responsabile și eşantioanele vecine și reciproc. De aceea prin înlăturarea eşantionului curent este afectată nu numai informația conținută în el ci și informația conținută în eşantioanele vecine. Rolul transformării ortogonale este de a decorela semnalul. În urma aplicării transformării ortogonale se obține un nou semnal. Dependența conținutului de informație al eşantionului curent al acestui nou semnal, $y[n]$, de informația conținută în eşantioanele vecine este mai slabă. De aceea prin înlăturarea eşantionului curent informația conținută în eşantioanele vecine este mai puțin afectată decât în cazul semnalului $x[n]$. Înlăturarea eşantionului curent conduce la o pierdere de informație cu atât mai mică cu cât valoarea sa este mai mică. Un exemplu este prezentat în figura 2.1.2.

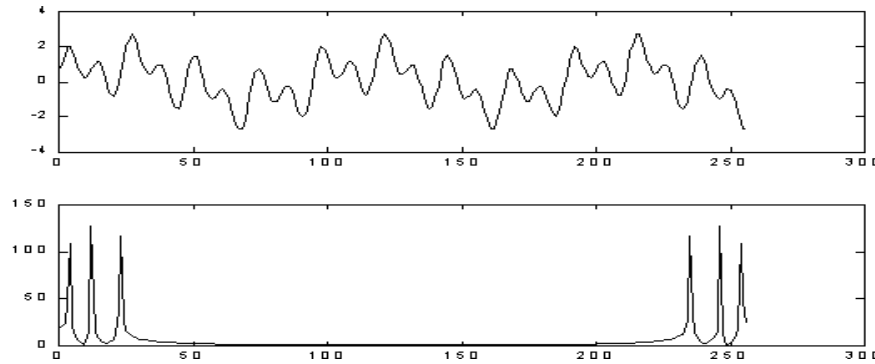


Figura 2.1.2. Un exemplu de semnal $x[n]$ (sus) și $y[n]$ (jos). Transformarea ortogonală folosită este modulul transformării Fourier discretă.

Analizând această figură se constată că prin aplicarea transformării Fourier discretă toată informația s-a grupat în șase eșantioane de valoare semnificativă. Anulând orice grup de eșantioane dintre cele cu indicele cuprins între 60 și 190, conținutul informațional al semnalului $y[n]$ nu se modifică. Evident nu se poate spune același lucru pentru semnalul $x[n]$, toate eșantioanele cu indicele cuprins între 60 și 190 având o contribuție importantă la forma acestuia.

După cum s-a arătat în [1], în paragraful 3.1.1. sau în [2] sau în [4], transformarea care realizează decorelarea maximă a unui semnal este transformarea Karhunen-Loeve. Din păcate nu există algoritmi rapizi pentru calculul acestei transformări deoarece acesta presupune inversarea unei matrici. Există cazuri de semnale când această matrice este singulară. De aceea în practică se utilizează transformări suboptimale, ca de exemplu transformarea cosinus discretă sau transformarea "wavelet" discretă, așa cum s-a arătat în paragraful 3.2. din [1]. Deși aceste două transformări converg asimptotic la transformarea Karhunen-Loeve totuși aplicarea transformării "wavelet" discretă la compresie are anumite avantaje. Aceste avantaje vor fi evidențiate în această lucrare. Vor fi analizate, de asemenea, și alte transformări discrete, bazate pe teoria funcțiilor "wavelet", ca de exemplu, transformarea cu pachete de funcții "wavelet" discretă, TPWD, sau transformarea cu pachete cosinusoidale discretă, TPC. Se va demonstra că și aceste transformări converg asimptotic la transformarea Karhunen-Loeve. Pentru aceasta este însă necesar să se prezinte câteva aspecte ale teoriei funcțiilor "wavelet". O astfel de prezentare este făcută în capitolul 2 din [5]. În continuare se vor prezenta, fără demonstrații, principalele rezultate obținute în ultima referință bibliografică citată.

2.1.1. Legătura dintre teoria funcțiilor "wavelet" și codarea în subbenzi

O introducere detaliată în teoria funcțiilor "wavelet", în limba română, evidențiând aspectele matematice semnificative ale acesteia poate fi găsită în [4] sau în [6]. În continuare nu vor fi prezentate decât aspectele interesante ale acestei teorii din punctul de vedere al compresiei de date. Pentru început se analizează funcționarea unui codor în două subbenzi, celulă de bază a structurilor de prelucrare *multirating*, care se folosesc la compresia semnalelor audio, conform standardului MPEG. Se consideră sistemul din figura 2.1.1.1.b).

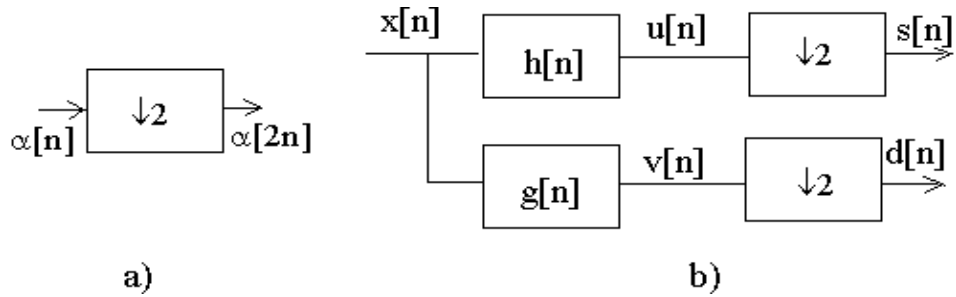


Figura 2.1.1.1. a) Simbol pentru un decimator; b) schema unui codor cu două subbenzi.

Pentru început se consideră cazul cel mai simplu, în care cele două filtre folosite sunt ideale. Răspunsurile în frecvență ale filtrelor numerice cu răspunsurile la impuls $h[n]$ și $g[n]$ din figura 2.1.1.1. sunt prezentate în figura 2.1.1.2.

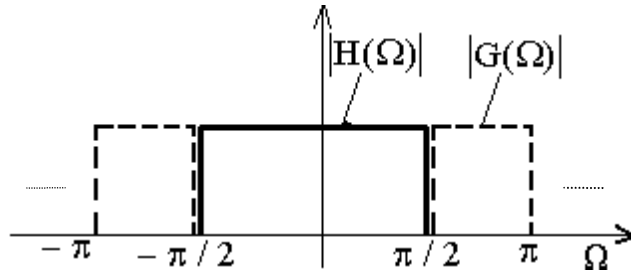


Figura 2.1.1.2. Răspunsurile în frecvență ale filtrelor din figura 2.1.1.1.

Pentru a analiza codorul în subbenzi se calculează transformatele z ale semnalelor $s[n]$ și $d[n]$. În acest scop se constată că:

$$U(z) = X(z) H(z) \quad ; \quad V(z) = X(z) G(z)$$

Conform definiției transformatei z :

$$S(z) = \sum_n s[n] z^{-n} = \sum_n u[2n] z^{-n}$$

$$U(z) = \sum_n u[n] z^{-n} = \sum_n u[2n] z^{-2n} + \sum_n u[2n+1] z^{-(2n+1)}$$

$$U(-z) = \sum_n u[2n] z^{-2n} - \sum_n u[2n+1] z^{-(2n+1)}$$

și se observă că putem scrie:

$$\frac{1}{2} [U(z) + U(-z)] = \sum_n u[2n] z^{-2n} = \sum_n u[2n] (z^2)^{-n} = S(z^2)$$

Revenind la expresia lui $S(z)$:

$$S(z) = \frac{1}{2} \left[U \left(z^{\frac{1}{2}} \right) + U \left(-z^{\frac{1}{2}} \right) \right] \quad (2.1.1.1)$$

sau :

$$S(z) = \frac{1}{2} \left[X \left(z^{\frac{1}{2}} \right) H \left(z^{\frac{1}{2}} \right) + X \left(-z^{\frac{1}{2}} \right) H \left(-z^{\frac{1}{2}} \right) \right] \quad (2.1.1.2)$$

În mod analog se demonstrează că :

$$D(z) = \frac{1}{2} \left[X \left(z^{\frac{1}{2}} \right) G \left(z^{\frac{1}{2}} \right) + X \left(-z^{\frac{1}{2}} \right) G \left(-z^{\frac{1}{2}} \right) \right] \quad (2.1.1.3)$$

Pentru a calcula spectrele semnalelor $s[n]$ și $d[n]$ se folosește substituția:

$$z = e^{j\Omega}$$

în relațiile (2.1.1.2) și (2.1.1.3), obținându-se :

$$S(\Omega) = \frac{1}{2} \left[X \left(\frac{\Omega}{2} \right) \cdot H \left(\frac{\Omega}{2} \right) + X \left(\frac{\Omega}{2} + \pi \right) \cdot H \left(\frac{\Omega}{2} + \pi \right) \right]$$

$$D(\Omega) = \frac{1}{2} \left[X \left(\frac{\Omega}{2} \right) \cdot G \left(\frac{\Omega}{2} \right) + X \left(\frac{\Omega}{2} + \pi \right) \cdot G \left(\frac{\Omega}{2} + \pi \right) \right]$$

Fie, de exemplu, spectrul $X(\Omega)$, cel trasat în figura 2.1.1.3.

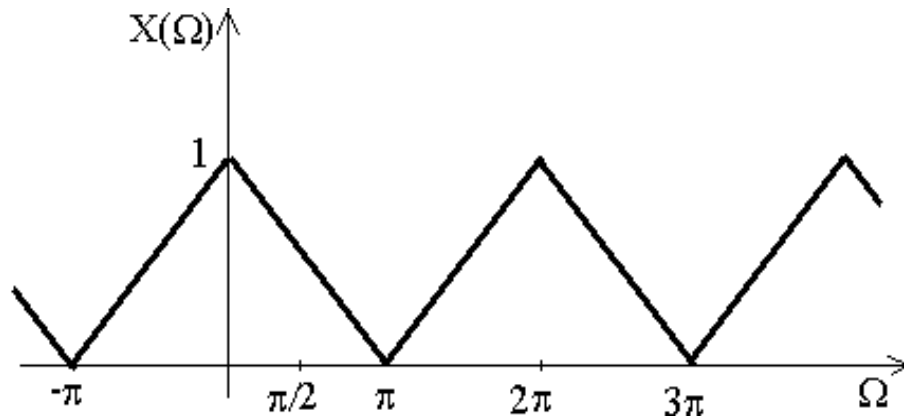


Figura 2.1.1.3. Un exemplu de spectru de semnal de intrare.

Spectrele semnalelor $s[n]$ și $d[n]$ sunt prezentate în figurile 2.1.1.4 și 2.1.1.5.

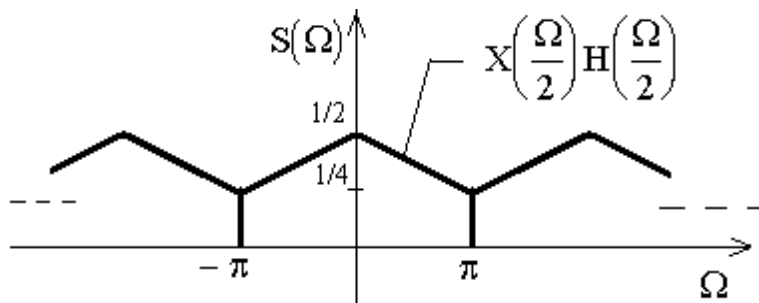


Figura 2.1.1.4. Spectrul semnalului $s[n]$.

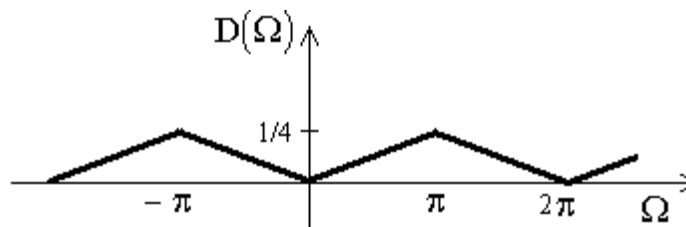


Figura 2.1.1.5. Spectrul semnalului $d[n]$.

Se constată că spectrul $S(\Omega)$ este asemenea cu spectrul $X(\Omega)$ în banda $[-\pi/2, \pi/2]$.
 Se constată că porțiunea din spectrul $D(\Omega)$ în banda $[-2\pi, -\pi] \cup [\pi, 2\pi]$ este asemenea cu spectrul $X(\Omega)$ în banda $[-\pi, \pi] - [-\pi/2, \pi/2]$.

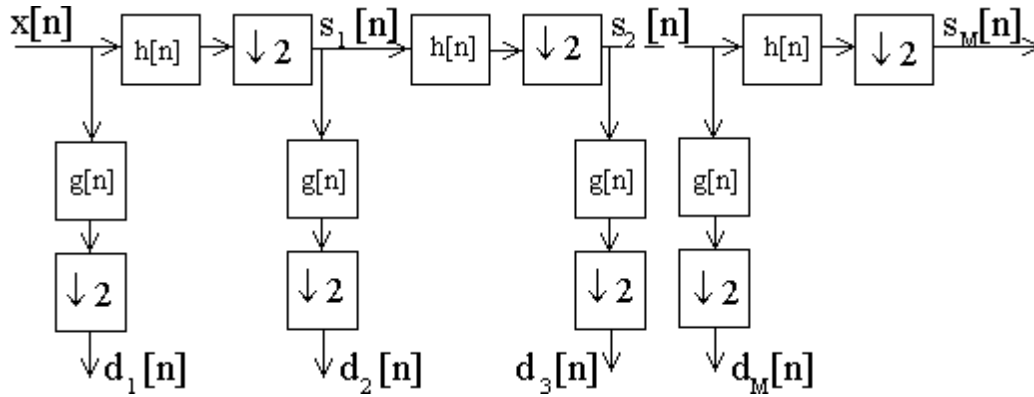


Figura 2.1.1.6. Structură arborescentă de codare în subbenzi.

Se poate deci afirma că semnalul $x[n]$ a fost codat în două subbenzi, componentele sale de joasă frecvență regăsindu-se în semnalul $s[n]$ iar componentele sale de înaltă frecvență, în semnalul $d[n]$. Sistemul de codare în subbenzi din figura 2.1.1.1 poate fi privit și ca un sistem de proiecție. Dacă secvența $x[n]$ reprezintă coeficienții dezvoltării unui semnal $x(t)$ într-o bază a unui spațiu Hilbert V_0 atunci secvențele $s[n]$ și $d[n]$ reprezintă coeficienții dezvoltărilor semnalelor $s(t)$ și $d(t)$, proiecțiile semnalului $x(t)$ pe două subspații Hilbert închise ale lui V_0 , V_{-1} și W_{-1} , în bazele acestor subspații. Spațiile V_0 și V_{-1} reprezintă elementele unei analize multirezoluție.

Pentru a crește numărul de subbenzi se poate utiliza o structură arborescentă așa cum se vede în figura 2.1.1.6. Acest sistem poate fi utilizat, pentru calculul transformării "wavelet" discretă.

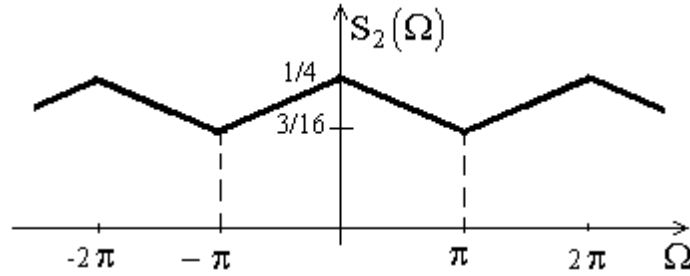
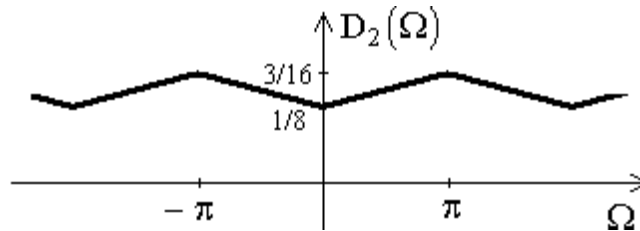
Se calculează transformatele z ale semnalelor $s_k[n]$ și $d_k[n]$, $k = 1 \div M$. Se observă (conform figurii 2.1.1.1) că:

$$s_1[n] = s[n] \quad ; \quad d_1[n] = d[n]$$

și astfel se poate scrie :

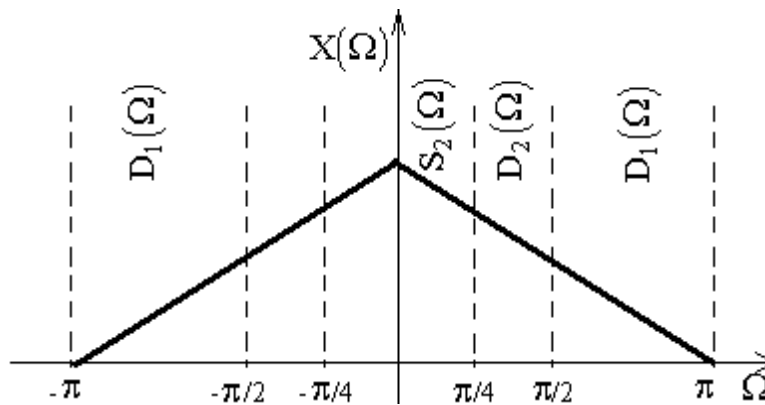
$$S_2(z) = \frac{1}{2} \left[S_1 \left(z^{\frac{1}{2}} \right) H \left(z^{\frac{1}{2}} \right) + S_1 \left(-z^{\frac{1}{2}} \right) H \left(-z^{\frac{1}{2}} \right) \right]$$

$$D_2(z) = \frac{1}{2} \left[S_1 \left(z^{\frac{1}{2}} \right) G \left(z^{\frac{1}{2}} \right) + S_1 \left(-z^{\frac{1}{2}} \right) G \left(-z^{\frac{1}{2}} \right) \right]$$

Figura 2.1.1.7. Spectrul semnalului $s_2[n]$.Figura 2.1.1.8. Spectrul semnalului $d_2[n]$.

Continuând exemplul considerat anterior, spectrele semnalelor $s_2[n]$ și $d_2[n]$ iau formele din figurile 2.1.1.7 și 2.1.1.8.

Se constată că spectrul $S_2(\Omega)$ este asemenea cu spectrul $X(\Omega)$ din banda $[-\pi/4, \pi/4]$ și că spectrul $D_2(\Omega)$ este asemenea cu spectrul $X(\Omega)$ din banda $[-\pi/2, \pi/2] - [-\pi/4, \pi/4]$.

Figura 2.1.1.9. Corespondența dintre spectrul $X(\Omega)$ și spectrele $S_k(\Omega)$, $D_k(\Omega)$, $k = 1+2$.

Procedând similar se constată că spectrul $S_M(\Omega)$ este asemenea cu spectrul $X(\Omega)$ din banda $[-\pi/2^M, \pi/2^M]$ și că spectrul $D_M(\Omega)$ este asemenea cu spectrul $X(\Omega)$ din banda $[-\pi/(2^{M-1}), \pi/(2^{M-1})] - [-\pi/2^M, \pi/2^M]$.

Cu alte cuvinte fâșii din spectrul $X(\Omega)$ au fost puse în corespondență cu semnalele $s_k[n]$ și $d_k[n]$. Această corespondență este evidențiată în figura 2.1.1. 9.

Se constată că folosind sistemul din figura 2.1.1.6, banda spectrului semnalului $x[n]$ este divizată în octave. Se poate deci afirma că sistemul cu structură arborescentă din figura 2.1.1.6 este într-adevăr un codor în subbenzi. În continuare se analizează operația de decodare.

Se pune problema refacerii semnalului $x[n]$ pornind de la semnalele $s[n]$ și $d[n]$. Se consideră în acest scop sistemul din figura 2.1.1.10 b).

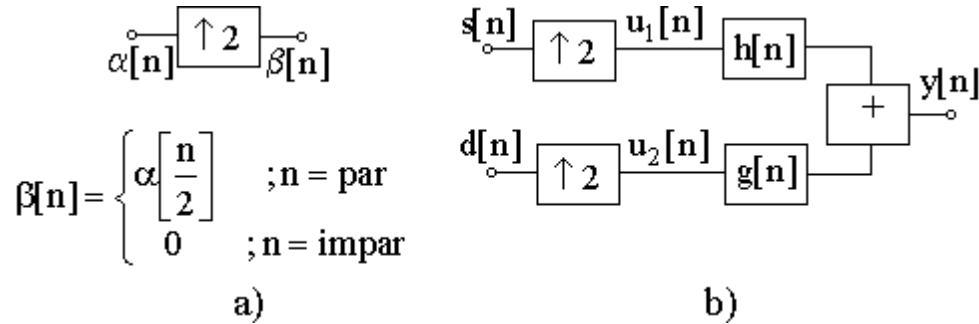


Figura 2.1.1.10. a) Interpolator și definiția semnalului de la ieșirea sa; b) sistem de decodare corespunzător celui din figura 2.1.1.1.

Se calculează transformata z a semnalului $b[n]$ (figura 2.1.1.10 a.) pe baza transformatei z a semnalului $a[n]$:

$$\alpha(z) = \sum_n \alpha[n]z^{-n}$$

$$\beta(z) = \sum_n \beta[n]z^{-n} = \sum_n \beta[2n]z^{-2n} + \sum_n \beta[2n+1]z^{-(2n+1)} =$$

$$= \sum_n \alpha[n]z^{-2n} = \alpha(z^2)$$

astfel încât se pot scrie transformatele z pentru celelalte semnale ce apar în sistemul de codare:

$$U_1(z) = S(z^2) \quad ; \quad U_2(z) = D(z^2);$$

sau, ținând seama de relațiile (2.1.1.2) și (2.1.1.3):

$$Y(z) = H(z) \cdot \frac{1}{2} [X(z) \cdot H(z) + X(-z) \cdot H(-z)] +$$

$$+ G(z) \cdot \frac{1}{2} [X(z) \cdot G(z) + X(-z) \cdot G(-z)] \quad (2.1.1.4)$$

Pe baza acestei relații se determină spectrul semnalului $y[n]$:

$$Y(\Omega) = H(\Omega) \cdot \frac{1}{2} [X(\Omega) \cdot H(\Omega) + X(\Omega + \pi) \cdot H(\Omega + \pi)] +$$

$$+ G(\Omega) \cdot \frac{1}{2} [X(\Omega) \cdot G(\Omega) + X(\Omega + \pi) \cdot G(\Omega + \pi)] \quad (2.1.1.5)$$

Dacă se folosesc filtrele cu răspunsurile în frecvență cu caracteristicile de modul din figura 2.1.1.2 atunci sunt valabile relațiile :

$$H(\Omega) \cdot H(\Omega + \pi) = G(\Omega) \cdot G(\Omega + \pi) = 0$$

$$H^2(\Omega) + G^2(\Omega) = 1$$

Pe baza acestor relații, (2.1.1.5) devine :

$$\begin{aligned} Y(\Omega) &= \frac{1}{2} \cdot X(\Omega) \cdot H^2(\Omega) + \frac{1}{2} \cdot X(\Omega) \cdot G^2(\Omega) = \\ &= \frac{1}{2} \cdot X(\Omega) \cdot [H^2(\Omega) + G^2(\Omega)] = \frac{1}{2} \cdot X(\Omega) \end{aligned} \quad (2.1.1.6)$$

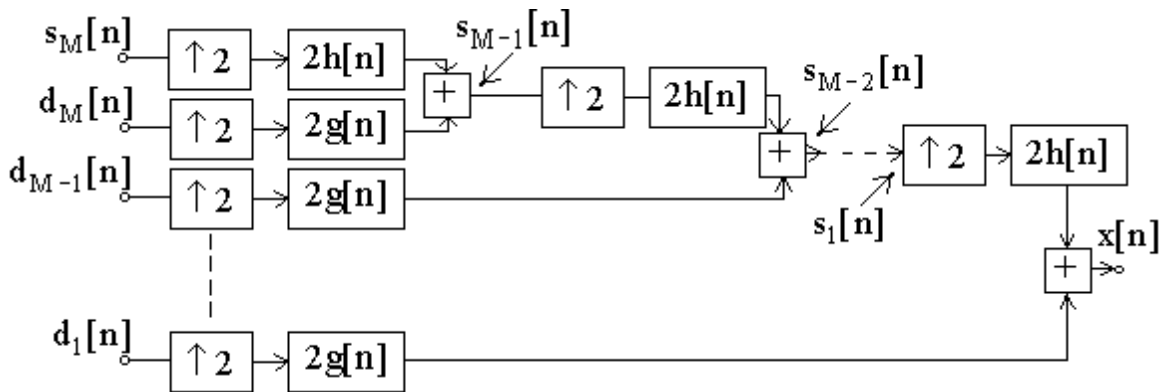


Figura 2.1.1.11. Schema unui decodor pentru semnale codate în M subbenzi.

Deci, cu excepția unei constante multiplicative (egală cu 1/2), semnalele $x[n]$ și $y[n]$ sunt identice. Se spune că sistemul de decodare din figura 2.1.1.11 este cu reconstrucție perfectă. De aceea sistemul din figura 2.1.1.11 poate fi utilizat pentru reconstrucția perfectă a semnalului prelucrat de sistemul din figura 2.1.1.6, în ipoteza că se folosesc filtrele ideale cu răspunsurile în frecvență din figura 2.1.1.2. Sistemul din figura 2.1.1.11 permite calculul transformării "wavelet" discretă inversă.

OBSERVAȚII

- O1.** O analiză similară poate fi făcută și pentru cazul în care interpolarea și decimarea nu se fac folosind constanta 2 ci o alta, de exemplu, M , $M \in \mathbb{N}$. În acest caz nu se va mai obține o descompunere în octave a benzii B a semnalului $u[n]$ ci în subbenzi a căror lățime va scădea cu puteri ale lui M .
- O2.** Pentru structurile care utilizează arbori simetrici se poate face o analiză similară. Aceasta corespunde noțiunii de pachete de funcții "wavelet", introdusă în [7].
- O3.** Principala limitare a sistemelor de codare și decodare în subbenzi cu structură arborescentă prezentate până acum este că filtrele cu răspunsurile în frecvență din figura 2.1.1.2 nu sunt realizabile.

În continuare se vor determina clase de filtre realizabile care permit codarea în subbenzi, cu structură arborescentă și cu reconstrucție perfectă.

2.1.1.1. Codarea subbandă cu reconstrucție perfectă folosind sisteme cu structură arborescentă cu filtre realizabile

Se consideră în continuare că $h_r[n]$ și $g_r[n]$ sunt filtre realizabile. Un sistem, echivalent celui din figura 2.1.1.10, destinat reconstrucției perfecte, este prezentat în figura 2.1.1.12.

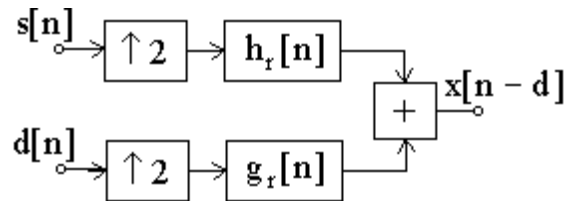


Figura 2.1.1.12. Sistemul de reconstrucție corespunzător unui codor în două subbenzi.

Conform acestei figuri rezultă că semnalul de la ieșirea decodorului este o variantă întârziată cu d a semnalului de la intrare.

Trebuie determinate răspunsurile la impuls $h_r[n]$ și $g_r[n]$ precum și condițiile pe care trebuie să le îndeplinească răspunsurile la impuls $h_r[n]$ și $g_r[n]$ pentru ca la ieșirea sistemului din figura 2.1.1.12 să se poată obține semnalul $x[n-d]$. În acest scop se rescrie relația (2.1.1.4):

$$z^{-d} \cdot X(z) = H_r(z) \cdot \frac{1}{2} [X(z) \cdot H(z) + X(-z) \cdot H(-z)] + G_r(z) \cdot \frac{1}{2} [X(z) \cdot G(z) + X(-z) \cdot G(-z)] \quad (2.1.1.7)$$

sau, regrupând în membrul drept :

$$z^{-d} \cdot X(z) = X(z) \cdot \frac{1}{2} [H_r(z) \cdot H(z) + G_r(z) \cdot G(z)] + X(-z) \cdot \frac{1}{2} [H_r(z) \cdot H(-z) + G_r(z) \cdot G(-z)]$$

Această ecuație este satisfăcută și de soluțiile sistemului de ecuații :

$$H_r(z) \cdot H(z) + G_r(z) \cdot G(z) = 2z^{-d}$$

$$H_r(-z) \cdot H_r(z) + G_r(-z) \cdot G_r(z) = 0$$

În continuare se rezolvă acest sistem, considerându-se cunoscute transformatele z notate cu $H(z)$ și $G(z)$. Determinantul sistemului este :

$$\Delta = \begin{vmatrix} H(z) & G(z) \\ H(-z) & G(-z) \end{vmatrix} = H(z) \cdot G(-z) - H(-z) \cdot G(z)$$

Determinanții corespunzători celor două necunoscute sunt de forma:

$$\Delta H_r = \begin{vmatrix} 2z^{-d} & G(z) \\ 0 & G(-z) \end{vmatrix} = 2z^{-d} \cdot G(-z)$$

Deci soluțiile sunt date de relațiile următoare :

$$H_r(z) = \frac{2 \cdot z^{-d} \cdot G(-z)}{H(z) \cdot G(-z) - H(-z) \cdot G(z)} \quad (2.1.1.8)$$

$$G_r(z) = \frac{-2 \cdot z^{-d} \cdot H(-z)}{H(z) \cdot G(-z) - H(-z) \cdot G(z)} \quad (2.1.1.9)$$

Evident, o condiție care trebuie impusă filtrelor din structura codorului este ca ecuația :

$$H(z) \cdot G(-z) - H(-z) \cdot G(z) = 0 \quad (2.1.1.10)$$

să nu aibă nici o rădăcină diferită de rădăcinile ecuației :

$$z^{-d} = 0$$

De aceea o condiție potrivită pentru filtrele cu răspunsurile la impuls $h[n]$ și $g[n]$ ar fi :

$$H(z) \cdot G(-z) - H(-z) \cdot G(z) = 2z^{-d} \quad (2.1.1.11)$$

În acest caz relațiile (2.1.1.8) și (2.1.1.9) devin :

$$H_r(z) = G(-z) \quad (2.1.1.12)$$

$$G_r(z) = -H(-z) \quad (2.1.1.13)$$

Deci răspunsurile în frecvență ale filtrelor de reconstrucție depind de răspunsurile în frecvență ale filtrelor din structura codorului conform relațiilor :

$$H_r(\Omega) = G(\Omega + \pi) \quad (2.1.1.14)$$

$$G_r(\Omega) = -H(\Omega + \pi) \quad (2.1.1.15)$$

iar răspunsurile în frecvență ale filtrelor din structura codorului satisfac :

$$H(\Omega) \cdot G(\Omega + \pi) - H(\Omega + \pi) \cdot G(\Omega) = 2 \cdot e^{-j\Omega d} \quad (2.1.1.16)$$

$H_r(z)$ și $G_r(z)$ sunt funcțiile de transfer ale filtrelor introduse de Esteban și Galand, [8], sub numele de "Quadrature Mirror Filters", QMF.

OBSERVAȚIE : Relația corespunzătoare lui (2.1.1.11) în domeniul timp este:

$$\sum_k h[k] \cdot g[n-k] \cdot \{(-1)^{n-k} - (-1)^k\} = 2 \cdot \delta[n-d] \quad (2.1.1.17)$$

Pentru valori pare ale lui n această relație devine :

$$\delta[n-d] = 0$$

rezultând astfel necesitatea ca d să fie un număr natural impar.

S-a demonstrat așadar că în urma folosirii filtrelor QMF se poate realiza o reconstrucție perfectă pentru o codare în două subbenzi, dacă filtrele de reconstrucție îndeplinesc condițiile (2.1.1.12) și (2.1.1.13) iar filtrele de sinteză (cele cu răspunsurile la impuls $h[n]$ și $g[n]$) îndeplinesc condiția (2.1.1.11) în care valoarea lui d trebuie să fie impară. Relația (2.1.1.16) este generală. Ea nu furnizează informații despre modul în care se proiectează filtrele de sinteză. Smith și Barnwell au determinat o clasă de filtre de sinteză, [9]. Este vorba despre clasa filtrelor "conjugate quadratur filters", CQF. Ei au propus următoarea legătură între răspunsurile în frecvență ale filtrelor de sinteză, presupuse ca fiind cu răspunsuri la impuls reale :

$$G(\Omega) = -e^{-j\Omega d} \cdot H^*(\Omega + \pi) \quad (2.1.1.18)$$

Folosind această condiție membrul drept al relației (2.1.1.16) devine, pentru d impar :

$$|H(\Omega)|^2 + |H(\Omega + \pi)|^2 = 2 \quad (2.1.1.19)$$

În acest caz răspunsurile în frecvență ale filtrelor de reconstrucție devin:

$$H_r(\Omega) = e^{-j\Omega d} \cdot H^*(\Omega) \quad (2.1.1.20)$$

$$G_r(\Omega) = -H(\Omega + \pi) \quad (2.1.1.21)$$

COMENTARIU

1. Fie :

$$h'[n] = (-1)^n h[n]$$

Se constată că :

$$h'[n] \leftrightarrow H(\Omega + \pi)$$

Relația corespunzătoare relației (2.1.1.19) în domeniul timp este, conform relației Wiener - Hincin:

$$R_{hh}[n] + R_{h'h'}[n] = 2\delta[n] \quad (2.1.1.22)$$

De aceea se poate afirma că, din punctul de vedere al proiectării filtrelor din structura codorului, respectiv a decodorului, relația (2.1.1.19) este mai avantajoasă decât relația (2.1.1.16).

2. Cunoscându-se avantajele de implementare ale filtrelor RFI în comparație cu filtrele RII, în continuare se vor presupune ca fiind de tip RFI atât filtrele de sinteză cât și cele de analiză. Dacă filtrul cu răspuns la impuls $h[n]$ este causal atunci transformata sa Fourier în timp discret este :

$$H(\Omega) = \sum_{n=0}^{L-1} h[n] \cdot e^{-j\Omega n}$$

iar transformata sa z este :

$$H(z) = \sum_{n=0}^{L-1} h[n] \cdot z^{-n}$$

unde L reprezintă lungimea răspunsului la impuls pentru filtrul considerat.

De aceea, admițând că $h[n]$ sunt numere reale :

$$H^*(\Omega) = \sum_{n=0}^{L-1} h[n] \cdot e^{j\Omega n}$$

și:

$$H^*(\Omega + \pi) = \sum_{n=0}^{L-1} (-1)^n \cdot h[n] \cdot e^{j\Omega n}$$

Conform relației (2.1.1.18) rezultă că răspunsul în frecvență al celui alt filtru de sinteză va fi:

$$G(\Omega) = -e^{j\Omega d} \cdot \sum_{n=0}^{L-1} (-1)^n \cdot h[n] \cdot e^{j\Omega n} = - \sum_{n=0}^{L-1} (-1)^n \cdot h[n] \cdot e^{j\Omega(n-d)}$$

Pentru ca acest răspuns în frecvență să corespundă unui filtru causal este necesar ca pentru orice n cuprins între 0 și $L-1$ (inclusiv capetele) să fie îndeplinită condiția:

$$n - d < 0$$

și deci întârzierea d trebuie să satisfacă condiția:

$$d > L - 1 \quad (2.1.1.22)$$

Dacă se respectă această condiție atunci cele două filtre de sinteză sunt ambele cauzale. Rezultă că valoarea minimă a lui d este:

$$(2.1.1.23)$$

$$d_{\min} = L$$

Pentru a putea reconstrui cu întârziere minimă este deci necesar să se folosească filtre de sinteză de lungime impară.

Pe baza relațiilor (2.1.1.20) și (2.1.1.21) se constată că dacă este respectată condiția (2.1.1.23) atunci și filtrele de reconstrucție sunt cauzale.

3. Toate cele patru filtre (cu răspunsurile în frecvență $H(\Omega)$, $G(\Omega)$, $H_r(\Omega)$ și $G_r(\Omega)$) au aceeași lungime. Cu modificări minore schema poate funcționa cu filtre de analiză de o anumită lungime și cu filtre de sinteză de altă lungime [10], [11].

2.1.1.1.1. Metode de proiectare a filtrelor CQF

Se face notația :

$$F(z) = H(z) \cdot H(z^{-1})$$

sau :

$$F(\Omega) = H(\Omega) \cdot H^*(\Omega) = |H(\Omega)|^2 \quad (2.1.1.25)$$

Condiția (2.1.1.20) devine :

$$F(\Omega) + F(\Omega + \pi) = 2 \quad (2.1.1.26)$$

Se proiectează sistemul cu răspuns în frecvență $F(\Omega)$ pe baza relației (2.1.1.26). Apoi se deduce $H(\Omega)$ pe baza relației (2.1.1.25) și în final se deduc $G(\Omega)$, $H_r(\Omega)$ și $G_r(\Omega)$.

În [9] sunt prezentate mai multe exemple de răspunsuri în frecvență $H(\Omega)$ obținute pe baza metodei de proiectare descrise. Clasa acestor filtre poate fi restrânsă dacă se impun condiții suplimentare. De exemplu se poate impune:

- condiția de fază liniară (simetria răspunsului la impuls),
- condiția de lungime minimă a răspunsului la impuls,
- condiția ca expresiile eșantioanelor răspunsului la impuls să fie cât mai simple.

Toate aceste condiții sunt foarte importante atunci când se pune problema codării în mai multe subbenzi deoarece favorizează stabilitatea numerică a algoritmilor care implementează sistemele din figurile 2.1.1.11 și 2.1.1.16. Această stabilitate este asigurată dacă filtrele îndeplinesc o anumită condiție de regularitate [12], [13]. Condiția de regularitate este partea care leagă teoria sistemelor de codare subbandă de teoria funcțiilor "wavelet".

2.1.2 Legătura dintre sistemele de codare în subbenzi și teoria seriilor de funcții "wavelet"

Teoria seriilor de funcții "wavelet" dezvoltată în [12], [13], [14] are ca scop construcția unor noi baze Riesz ale spațiului $L^2(\mathbb{R})$. Se pornește de la definiția analizei multirezoluție.

DEFINIȚIA 1. Se numește analiză multirezoluție a spațiului $L^2(\mathbb{R})$, mulțimea de subspații Hilbert închise $\{V_m\}_{m \in \mathbb{Z}}$ care satisfac proprietățile :

i). $\dots V_{m+1} \subset V_m \subset V_{m-1} \dots m \in \mathbb{Z}$

ii). $\bigcup_{m \in \mathbb{Z}} V_m = L^2(\mathbb{R}), \bigcap_{m \in \mathbb{Z}} V_m = \{0\}$

iii). $\forall f(x) \in V_m, f(2x) \in V_{m-1}$

iv). $\exists \phi \in V_0$, astfel încât mulțimea $\{\phi_{m,n}(x)\} = \{2^{-m/2} \phi(2^{-m}x - n)\}_{n \in \mathbb{Z}}$ să formeze o bază Riesz a lui V_m pentru orice m .

Sunt prezentate numeroase exemple de analiză multirezoluție în [15] și [16]. Funcția $\varphi(x)$ se numește funcție de scalare. Numeroase exemple de funcții de scalare se găsesc în lucrările deja citate. Conform [8] orice bază Riesz poate fi transformată într-o bază ortonormală.

Se va considera în continuare că mulțimea $\{\varphi(x-k)\}_{k \in \mathbb{Z}}$ este o bază ortonormală a spațiului V_0 . În majoritatea lucrărilor deja citate este demonstrată următoarea teoremă.

TEOREMA 1. În ipoteza că $\{\varphi(x-k)\}_{k \in \mathbb{Z}}$ este o bază ortonormală a spațiului V_0 , mulțimea $\{\varphi_{m,k}(x)\}_{k \in \mathbb{Z}}$ este o bază ortonormală a spațiului V_m .

În continuare se determină proiecțiile unei funcții $f_0(x)$ din V_0 pe spațiile V_1, \dots, V_M , adică funcțiile $f_1(x), \dots, f_M(x)$:

$$f_1(x) = \sum_k \langle f(x), \varphi_{1,n}(x) \rangle \cdot \varphi_{1,n}(x)$$

Aceste funcții reprezintă aproximările de diferite rezoluții ale funcției $f(x)$. În anumite aplicații nu este necesară rezoluția maximă pentru a prelucra această funcție. De exemplu în cazul compresiei este uneori suficientă informația conținută într-una dintre aproximările de rezoluție mai scăzută. Coeficienții acestei dezvoltări în serie Fourier generalizată se notează cu $s_1[n]$ și sunt dați de:

$$\begin{aligned} s_1[n] &= \langle f(x), \varphi_{1,n}(x) \rangle = \langle f(x), \sum_k h^*[2n-k] \cdot \varphi[x-k] \rangle = \\ &= \sum_k h[2n-k] \cdot \langle f(x), \varphi[x-k] \rangle \end{aligned}$$

Folosind notația :

$$\langle f(x), \varphi(x-k) \rangle = s_0[k]$$

se obține :

$$s_1[n] = \sum_k s_0[k] \cdot h[2n-k] \quad (2.1.2.1)$$

relație care exprimă legătura între coeficienții proiecțiilor funcției $f(x)$ pe primele două elemente ale analizei multirezoluție. Coeficienții dezvoltării proiecției pe V_2 se notează cu $s_2[n]$ și sunt dați de :

$$s_2[n] = \langle f(x), \varphi_{2,n}(x) \rangle = \int_{-\infty}^{\infty} f(x) \cdot \varphi_{2,n}^*(x) dx$$

Dar, revenind la definiția 1, pentru $m = 2$, avem :

$$\begin{aligned} \varphi_{2,n}(x) &= 2^{-\frac{2}{2}} \cdot \varphi(2^{-2}x - n) = 2^{-\frac{1}{2}} \cdot 2^{-\frac{1}{2}} \varphi(2^{-1}(2^{-1}x) - n) = \\ &= 2^{-\frac{1}{2}} \cdot \varphi_{1,n}(2^{-1}x) \end{aligned}$$

sau:

$$2^{-\frac{1}{2}} \cdot \varphi_{1,n}(2^{-1}u) = 2^{-\frac{1}{2}} \cdot \sum_k h^*[2n-k] \cdot \varphi(2^{-1}u-k)$$

sau:

$$\varphi_{2,n}(u) = \sum_k h^*[2n-k] \cdot \varphi_{1,k}(u)$$

Procedând analog se poate demonstra că pentru orice m pozitiv este valabilă relația :

$$\varphi_{m,n}(x) = \sum_k h^*[2n-k] \cdot \varphi_{m-1,k}(x) \quad (2.1.2.2)$$

Se mai poate scrie:

$$\begin{aligned} s_2[n] &= \int_{-\infty}^{\infty} f(x) \cdot \left(\sum_k h^*[2n-k] \cdot \varphi_{1,k}(x) \right)^* dx = \\ &= \sum_k h[2n-k] \cdot \langle f(x), \varphi_{1,k}(x) \rangle \end{aligned}$$

adică:

$$s_2[n] = \sum_k s_1[k] \cdot h[2n-k] \quad (2.1.2.3)$$

Se poate demonstra prin recurență că :

$$s_m[n] = \sum_k s_{m-1}[k] \cdot h[2n-k] \quad (2.1.2.4)$$

pentru orice m pozitiv.

Analizând membrul drept al ultimei realții se constată că:

$$s_m[n] = s_{m-1}[p] * h[p] \Big|_{p=2n} \quad (2.1.2.5)$$

Cu alte cuvinte coeficienții dezvoltărilor proiecțiilor semnalului $f_0(t)$ pe două subspații succesive V_{m-1} și V_m , adică $s_{m-1}[n]$ și $s_m[n]$ se pot determina prin filtrare cu filtrul cu răspuns la impuls $h[n]$ și prin decimare.

Făcând notația :

$$s_0[n] = x[n]$$

rezultă că secvențele $s_m[n]$, $m=1 \div M$ pot fi obținute folosind sistemul din figura 2.1.1.11. **Aceasta este legătura dintre sistemele de codare în subbenzi și teoria seriilor de funcții "wavelet" care reprezintă subiectul acestui paragraf.**

OBSERVAȚII.

O1. Se calculează transformata Fourier, notată cu F , a variantelor translate și scalate ale funcțiilor de scalare:

$$F\{\varphi_{1,n}(x)\} = \hat{\varphi}_{1,n}(\omega) = F\left\{2^{-\frac{1}{2}} \cdot \varphi(2^{-1}x - n)\right\}(\omega) = 2^{-\frac{1}{2}} \cdot \int_{-\infty}^{\infty} \varphi(2^{-1}x - n) \cdot e^{-j\omega x} dx$$

Făcând schimbarea de variabilă $2^{-1}x - n = u$ se obține :

$$\begin{aligned} F\{\varphi_{1,n}(x)\} &= 2^{-\frac{1}{2}} \cdot \int_{-\infty}^{\infty} \varphi(u) \cdot e^{-j\omega 2(u+n)} \cdot 2 du = 2^{-\frac{1}{2}} \cdot \int_{-\infty}^{\infty} \varphi(u) \cdot e^{-2j\omega u} \cdot e^{-2j\omega n} du = \\ &= 2^{-\frac{1}{2}} \cdot e^{-2j\omega n} \cdot \int_{-\infty}^{\infty} \varphi(u) \cdot e^{-2j\omega u} du = 2^{-\frac{1}{2}} \cdot e^{-2j\omega n} \cdot \hat{\varphi}(2\omega) \end{aligned}$$

Deci :

$$\hat{\varphi}_{1,n}(\omega) = \sqrt{2} \cdot e^{-2j\omega n} \cdot \hat{\varphi}(2\omega)$$

sau:

$$\sqrt{2} \cdot \hat{\varphi}(2\omega) = \left(\sum_k h^*[2n - k] \cdot e^{-j\omega(k-2n)} \right) \cdot \hat{\varphi}(\omega)$$

adică :

$$\hat{\varphi}(2\omega) = \left(\frac{1}{\sqrt{2}} \cdot \sum_k (h^v)^*[k - 2n] \cdot e^{-j\omega(k-2n)} \right) \cdot \hat{\varphi}(\omega)$$

unde am făcut notația :

$$h^v[m] = h[-m]$$

Se face schimbarea de variabilă $k-2n=p$:

$$\frac{1}{\sqrt{2}} \cdot \sum_k h^v^*[k - 2n] \cdot e^{-j\omega(k-2n)} = \frac{1}{\sqrt{2}} \cdot \sum_p h^v^*[p] \cdot e^{-j\omega p}$$

În continuare, dacă facem notația :

$$m_0(\omega) = \frac{1}{\sqrt{2}} \sum_k h^v^*[k - 2n] \cdot e^{-j\omega(k-2n)}$$

se poate scrie:

$$\hat{\phi}(2\omega) = m_0(\omega) \cdot \hat{\phi}(\omega)$$

Se observă că $m_0(\omega)$ are semnificația de transformată Fourier în timp discret a secvenței $h^*[p]$, de variabilă ω .

În ultima relație se face schimbarea de variabilă $2\omega = u$ și avem :

$$\hat{\phi}(u) = m_0\left(\frac{u}{2}\right) \cdot \hat{\phi}\left(\frac{u}{2}\right)$$

sau:

$$\hat{\phi}(2\omega) = m_0(\omega) \cdot m_0\left(\frac{\omega}{2}\right) \cdot \hat{\phi}\left(\frac{\omega}{2}\right)$$

Procedând iterativ se poate demonstra că :

$$\hat{\phi}(\omega) = \prod_{p=1}^{\infty} m_0\left(\frac{\omega}{2^p}\right) \cdot \hat{\phi}(0)$$

Dar funcția de scalare reprezintă de obicei răspunsul la impuls al unui filtru trece jos. De aceea :

$$\hat{\phi}(0) = 1, \quad m_0(0) = 1, \quad \sum_p h^*[p] = 1$$

și ultima relație devine :

$$\hat{\phi}(\omega) = \prod_{p=1}^{\infty} m_0\left(\frac{\omega}{2^p}\right)$$

În consecință, în ipoteza că produsul din membrul drept converge, rezultă că ultima relație poate fi folosită pentru construcția unei funcții de scalare. Convergența produsului din membrul drept este asigurată de satisfacerea condiției de regularitate amintită anterior.

Deci mecanismul de construcție al unei funcții de scalare este următorul:

- Se alege un răspuns la impuls de filtru trece jos $h[n]$.
- Se construiește secvența $h^*[n]$.
- Se calculează $m_0(\omega)$.
- Se calculează $\phi(\omega)$.

Acest mecanism de construcție este remarcabil prin faptul că folosește metode de prelucrare în timp discret pentru construcția unui semnal în timp continuu.

În legătură cu analiza multirezoluție introdusă prin definiția 1 se poate defini complementul ortogonal al lui V_m în V_{m-1} , W_m :

$$V_{m-1} = V_m \oplus W_m$$

Șirul de subspații $\{W_m\}_{m \in \mathbb{Z}}$ astfel definite reprezintă o descompunere ortogonală a spațiului Hilbert al semnalelor de energie finită $L^2(\mathbb{R})$, [12].

DEFINIȚIA 2 : Șirul de subspații Hilbert închise $\{W_m\}_{m \in \mathbb{Z}}$ este o descompunere ortogonală a lui $L^2(\mathbb{R})$ dacă sunt îndeplinite condițiile:

i). $m \neq p \Rightarrow W_m \perp W_p$

ii). $\bigcup_{m \in \mathbb{Z}} W_m = L^2(\mathbb{R})$

conform [17].

În legătură cu descompunerile ortogonale ale lui $L^2(\mathbb{R})$ se poate demonstra următoarea teoremă.

TEOREMA 2. În W_0 există o funcție $\psi(x)$ astfel încât :

i) mulțimea $\{\psi(x-n)\}_{n \in \mathbb{Z}}$ să fie o bază ortonormală a lui W_0 ;

ii) mulțimea $\{\psi_{m,n}(x) = 2^{-m/2} \psi(2^{-m}x-n)\}_{n \in \mathbb{Z}}$ să fie o bază ortonormală a lui W_m pentru orice m

din \mathbb{Z} .

Funcția generatoare a acestor baze se numește "wavelets mother", iar elementele acestor baze se numesc funcții "wavelet".

OBSERVAȚII :

O1. Pot fi demonstrate și relațiile :

$$\sum_k g^*[2n-k] \cdot g[2m-k] = \delta[m-n]$$

$$\sum_k h^*[2n-k] \cdot h[2m-k] = \delta[m-n]$$

care sunt utile pentru descrierea comportării în domeniul timp a filtrelor cu răspunsurile în frecvență $H(\Omega)$ și $G(\Omega)$.

O2. Se determină caracterizarea în domeniul frecvență a funcțiilor "wavelet". În acest scop se calculează transformatele Fourier ale celor doi membri ai relației lor de definiție:

$$F\{\psi(x-n)\} = \sum_k g^*[2n-k] \cdot F\left\{2^{\frac{1}{2}} \cdot \varphi(2x-k)\right\}$$

sau :

$$e^{-j\omega n} \cdot \hat{\Psi}(\omega) = \sum_k g^*[2n-k] \cdot 2^{\frac{1}{2}} \cdot F\{\varphi(2x-k)\}$$

unde, făcând schimbarea de variabilă $2x-k=u$, se obține :

$$\begin{aligned} F\{\varphi(2x-k)\} &= \int_{-\infty}^{\infty} \varphi(2x-k) \cdot e^{-j\omega x} dx = \\ &= \int_{-\infty}^{\infty} \varphi(u) \cdot e^{-j\omega \frac{u+k}{2}} \cdot \frac{du}{2} = \frac{1}{2} \cdot e^{-j\frac{\omega}{2}k} \cdot \hat{\varphi}\left(\frac{\omega}{2}\right) \end{aligned}$$

și revenind la relația anterioară:

$$e^{-j\omega n} \cdot \hat{\Psi}(\omega) = \left(\sum_k \frac{1}{\sqrt{2}} \cdot g^*[2n - k] \cdot e^{-j\frac{\omega}{2}k} \right) \cdot \hat{\phi}\left(\frac{\omega}{2}\right)$$

de unde rezultă relația pentru transformata Fourier a lui $\psi(x)$:

$$\hat{\Psi}(\omega) = \left(\sum_k \frac{1}{\sqrt{2}} \cdot g^*[2n - k] \cdot e^{-j\frac{\omega}{2}(k-2n)} \right) \cdot \hat{\phi}\left(\frac{\omega}{2}\right)$$

Cu notația:

$$m_1\left(\frac{\omega}{2}\right) = \sum_k \frac{1}{\sqrt{2}} \cdot g^*[p] \cdot e^{-j\frac{\omega}{2}p}$$

ultima relație devine:

$$\hat{\Psi}(\omega) = m_1\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right)$$

Ținând seama de expresia transformatei Fourier a funcției de scalare expresia transformatei Fourier a funcției "wavelets mother" devine:

$$\hat{\Psi}(\omega) = m_1\left(\frac{\omega}{2}\right) \prod_{p=2}^{\infty} m_0\left(\frac{\omega}{2^p}\right)$$

Această relație permite construcția unei undișoare mamă pornind de la un anumit răspuns de tipul $h[n]$. **O2.** Funcția de scalare se construiește cu ajutorul filtrului cu răspunsul la impuls $h[n]$ iar undișoara mamă cu ajutorul filtrului cu răspunsul la impuls $g[n]$.

În continuare se stabilește legătura între proiecțiile unui semnal $f(x)$ din V_0 pe subspații succesive V_{m-1} , W_m . În acest scop se calculează produsul scalar $\langle \psi_{1,n}(x), \varphi(x-1) \rangle$. Se poate scrie:

$$\psi_{1,n}(x) = \sum_k g^*[2n - k] \cdot \varphi_{0,k}(x)$$

sau:

$$\langle \psi_{1,n}(x), \varphi(x-1) \rangle = \left\langle \sum_k g^*[2n - k] \cdot \varphi_{0,k}(x), \varphi_{0,1}(x) \right\rangle = g^*[2n - 1]$$

Vom putea scrie astfel :

$$\psi_{1,n}(x) = \sum_l g^*[2n-l] \cdot \varphi(x-l)$$

și :

$$e_1(x) = \sum_n \langle e_1(x), \psi_{1,n}(x) \rangle \cdot \psi_{1,n}(x)$$

Aplicând teorema proiecției (Riesz), obținem :

$$\begin{aligned} \langle e_1(x), \psi_{1,n}(x) \rangle &= \langle f(x), \psi_{1,n}(x) \rangle = \langle \sum_k s_0[k] \cdot \varphi(x-k), \psi_{1,n}(x) \rangle = \\ &= \sum_k s_0(k) \cdot \langle \psi_{1,n}(x), \varphi(x-k) \rangle^* = \sum_k s_0[k] \cdot g[2n-k] \end{aligned}$$

Deci coeficienții dezvoltării semnalului $e_1(x)$ (care reprezintă eroarea cu care semnalul $f_1(x)$ aproximează semnalul $f_0(x)$) în baza $\{\psi_{1,n}(x)\}_{n \in \mathbb{Z}}$ sunt $\sum_k s_0[k] \cdot g[2n-k]$. Însă pentru acești coeficienți am făcut notația $d_1[n]$ în figura 2.1.1.11, deci:

$$d_1[n] = \sum_k s_0[k] \cdot g[2n-k]$$

Procedând prin recurență se poate demonstra că:

$$d_m[n] = \sum_k s_{m-1}[k] \cdot g[2n-k] = \langle f(x), \psi_{m,n}(x) \rangle$$

Cu alte cuvinte coeficienții dezvoltării proiecției semnalului $f(x)$ din V_0 pe subspațiul W_m , se pot determina prin filtrare cu un filtru cu răspunsul la impuls $g[n]$ și prin decimare pornind de la proiecția sa pe spațiul V_{m-1} . Se constată că pentru $m=1, M$, secvențele $d_m[n]$ pot fi obținute folosind sistemul din figura 2.1.1.11. Cu alte cuvinte, folosind acest sistem, poate fi determinată aproximarea de rezoluție M a unui semnal precum și eroarea produsă de această aproximare.

OBSERVAȚII :

O1. În lucrarea sa, [12], Ingrid Daubechies determină toate răspunsurile la impuls de filtre FIR, $h[n]$ și $g[n]$ care satisfac o anumită condiție de regularitate. Alte condiții de regularitate sunt prezentate în [18]... [22]. Aceste răspunsuri la impuls sunt tabelate și sunt clasificate după lungimea lor. De exemplu filtrul DAU2 este unul cu lungimea răspunsului la impuls egală cu 4. Dezavantajul major al acestor filtre este că nu au caracteristici de fază liniare. Cu cât lungimea filtrelor crește, cu atât erorile de rotunjire ale coeficienților sunt mai însemnate.

O2. Dacă se abandonează ipoteza de ortonormalitate a mulțimii $\{\varphi(x-n)\}_{n \in \mathbb{Z}}$, considerându-se că aceasta este doar o bază Riesz, atunci teoria prezentată în acest paragraf poate fi generalizată. Această generalizare a fost făcută în [22] obținându-se clasa undișoarelor biortogonale cu suport compact. Filtrele corespunzătoare sunt tot de tip FIR dar de această dată filtrele de reconstrucție au lungime diferită de filtrele de sinteză. Ele pot fi filtre cu fază liniară. Numeroase exemple de răspunsuri la impuls de filtre din această clasă sunt prezentate în lucrările citate.

O3. Avantajul pentru compresie al abordării bazate pe utilizarea undișoarelor ortonormale asupra celei bazate pe undișoare biortogonale este prezentat în continuare.

Folosind notațiile utilizate până aici, putem scrie:

$$f_0(x) = f_M(x) + \sum_{m=1}^M e_m(x)$$

Astfel:

$$\begin{aligned} \left\| f_0(x) \right\|^2 &= \left\langle f_M(x) + \sum_{m=1}^M e_m(x), f_M(x) + \sum_{m=1}^M e_m(x) \right\rangle = \\ &= \left\langle f_M(x), f_M(x) + \sum_{m=1}^M e_m(x) \right\rangle + \left\langle \sum_{m=1}^M e_m(x), f_M(x) + \sum_{m=1}^M e_m(x) \right\rangle = \\ &= \left\langle f_M(x), f_M(x) \right\rangle + \left\langle f_M(x), \sum_{m=1}^M e_m(x) \right\rangle + \left\langle f_M(x), \sum_{m=1}^M e_m(x) \right\rangle^* + \left\langle \sum_{k=1}^M e_k(x), \sum_{m=1}^M e_m(x) \right\rangle^* = \\ &= \left\| f_M(x) \right\|^2 + 2\operatorname{Re} \left\{ \left\langle f_M(x), \sum_{m=1}^M e_m(x) \right\rangle \right\} + \left\| \sum_{m=1}^M e_m(x) \right\|^2 \end{aligned}$$

În această relație, deoarece avem:

$$f_M(x) \perp e_m(x), \quad \forall m = 1 \div M$$

va rezulta și:

$$f_M(x) \perp \sum_{m=1}^M e_m(x), \quad \forall m = 1 \div M$$

iar în ceea ce privește norma lui $f_0(x)$ putem scrie:

$$\left\| f_0(x) \right\|^2 = \left\| f_M(x) \right\|^2 + \left\| \sum_{k=1}^M e_k(x) \right\|^2$$

apoi, folosind proprietăți ale produsului scalar:

$$\begin{aligned} \left\| \sum_{k=1}^M e_k(x) \right\|^2 &= \left\langle \sum_{k=1}^M e_k(x), \sum_{l=1}^M e_l(x) \right\rangle = \sum_{k=1}^M \left\langle e_k(x), \sum_{l=1}^M e_l(x) \right\rangle = \\ &= \sum_{k=1}^M \sum_{l=1}^M \left\langle e_l(x), e_k(x) \right\rangle = \sum_{k=1}^M \left\| e_k(x) \right\|^2 \end{aligned}$$

Revenind, se poate spune că am demonstrat că:

$$\left\| f_0(x) \right\|^2 = \left\| f_M(x) \right\|^2 + \sum_{k=1}^M \left\| e_k(x) \right\|^2$$

Pe baza relației lui Parseval se poate scrie relația în timp discret echivalentă relației anterioare. Aceasta este:

$$\left\| s_0[n] \right\|^2 = \left\| s_M[n] \right\|^2 + \sum_{k=1}^M \left\| d_k[n] \right\|^2$$

Ultimele două relații pot fi privite ca forme ale principiului conservării energiei, specific pentru transformări ortogonale. Aceste relații nu sunt valabile în cazul funcțiilor "wavelet" biortogonale. De aceea se poate afirma că este de preferat să se utilizeze undișoarele ortogonale atunci când sunt necesare aproximări de eroare medie pătratică minimă. Se poate afirma și că transformările discrete bazate pe funcții "wavelet" biortogonale nu sunt ortogonale. De aceea ele sunt redundante.

Iată un motiv serios ca astfel de transformări să nu fie folosite la compresia de date.

O4. Teoria expusă poate fi generalizată și pentru codoare în subbenzi cu structură arborescentă simetrică. Această generalizare este făcută în [23].

O5. Teoria expusă poate fi generalizată și pentru codoare în subbenzi care utilizează decimatoare și interpolatoare cu constante M , diferite de 2, [24].

O6. O altă direcție de dezvoltare a codoarelor în subbenzi este cea bazată pe utilizarea filtrelor de analiză și sinteză RII, [25], sau a celor variabile în timp [26].

2.1.3. Transformarea "wavelet" discretă TUD

În paragraful anterior s-a stabilit legătura dintre seriile de funcții "wavelet" și tehnica codării în subbenzi. Cu ajutorul sistemului din figura 2.1.1.11 poate fi introdusă noțiunea de transformare "wavelet" discretă. Acest sistem transformă secvența $x[n]$ în secvențele $s_M[n]$ și $d_1[n], d_2[n], \dots, d_M[n]$. Fie $y[n]$ secvența obținută prin concatenarea acestor secvențe :

$$y[n] = \{s_M[n], d_1[n], \dots, d_M[n]\}$$

Operația :

$$x[n] \rightarrow y[n]$$

poartă numele de transformare "wavelet" discretă DWT sau transformare undișoară discretă (TUD).
Operația:

$$y[n] \rightarrow x[n]$$

care poate fi implementată de sistemul din figura 2.1.1.16 poartă numele de transformare "wavelet" discretă inversă IDWT sau transformare undișoară discretă inversă (TUDI).

Se poate demonstra că TUD este liniară și ortogonală. În continuare se prezintă, pe un exemplu, algoritmul lui Mallat de calcul al TUD, [27]. Fie X vectorul secvenței de intrare:

$$X = S_0 = \begin{bmatrix} s_0[8] \\ s_0[7] \\ \vdots \\ s_0[1] \end{bmatrix}$$

Se consideră că lungimea filtrelor $h[n]$ și $g[n]$ este 4. Primul pas al algoritmului este descris de relația:

$$Y_1 = M_0 X$$

unde matricea M_0 este dată de relația:

$$M_0 = \begin{bmatrix} h[0] & h[1] & h[2] & h[3] & 0 & 0 & 0 & 0 \\ -h[3] & h[2] & -h[1] & h[0] & 0 & 0 & 0 & 0 \\ 0 & 0 & h[0] & h[1] & h[2] & h[3] & 0 & 0 \\ 0 & 0 & -h[3] & h[2] & -h[1] & h[0] & 0 & 0 \\ 0 & 0 & 0 & 0 & h[0] & h[1] & h[2] & h[3] \\ 0 & 0 & 0 & 0 & -h[3] & h[2] & -h[1] & h[0] \\ h[2] & h[3] & 0 & 0 & 0 & 0 & h[0] & h[1] \\ -h[1] & h[0] & 0 & 0 & 0 & 0 & -h[3] & h[2] \end{bmatrix}$$

Se constată că se obține:

$$Y_1^T = [s_1[4] d_1[4] s_1[3] d_1[3] s_1[2] d_1[2] s_1[1] d_1[1]]$$

Prin permutări rezultă:

$$(Y_1^1)^T = [s_1[4] s_1[3] s_1[2] s_1[1] d_1[4] d_1[3] d_1[2] d_1[1]]$$

care este un vector obținut prin concatenarea secvențelor $s_1[n]$ și $d_1[n]$. Separând aceste secvențe se obțin vectorii:

$$(X_1^1)^T = [s_1[4] s_1[3] s_1[2] s_1[1]]$$

$$(X_1^2)^T = [d_1[4] d_1[3] d_1[2] d_1[1]]$$

Fie M_1 matricea obținută prin restrângerea matricii M_0 la sfertul său din stânga sus :

$$M_1 = \begin{bmatrix} h[0] & h[1] & h[2] & h[3] \\ -h[3] & h[2] & -h[1] & h[0] \\ 0 & 0 & h[0] & h[1] \\ 0 & 0 & -h[3] & h[2] \end{bmatrix}$$

Cel de-al doilea pas al algoritmului este descris de relația:

$$Y_2 = M_1 X_1^1$$

și rezultatul este:

$$Y_2^T = [s_2[2] d_2[2] s_2[1] d_2[1]]$$

În mod analog rezultă prin permutări:

$$\left(Y_2^1\right)^T = [s_2[2]s_2[1]d_2[2]d_2[1]]$$

unde dacă separăm secvențele $s_2[n]$ și $d_2[n]$ obținem:

$$\left(X_2^1\right)^T = [s_2[2]s_2[1]] \quad \text{și} \quad \left(X_2^2\right)^T = [d_2[2]d_2[1]]$$

Acum, cu ajutorul vectorilor X_2^1 , X_2^2 și X_1^2 se construiește vectorul Y :

$$Y^T = \left[\left(X_2^1\right)^T \left(X_2^2\right)^T \left(X_1^2\right)^T \right]$$

Această relație reprezintă rezultatul aplicării transformării "wavelet" discrete vectorului X . Analizând numărul de operații efectuate se constată că pentru primul pas al algoritmului au fost necesare 32 de înmulțiri și că pentru al doilea pas al algoritmului au fost necesare 16 înmulțiri, în total 48. Dacă vectorul X ar fi avut N elemente atunci s-ar fi efectuat un număr de înmulțiri de ordinul $4N$ [28]. Dacă s-ar fi folosit filtre de lungime L atunci acest număr ar fi fost LN . Pentru N suficient de mare se constată că numărul de înmulțiri necesare este inferior lui $N \log_2 N$, adică transformarea "wavelet" discretă poate fi efectuată mai rapid decât FFT a aceleiași secvențe. Acesta este motivul pentru care această transformare se mai numește și transformarea "wavelet" rapidă.

Pentru calculul transformării inverse trebuie aplicate operațiile descrise anterior în ordine inversă. Bineînțeles în locul matricilor M_0 , M_1 , ... trebuiesc folosite matricile M_0^T , M_1^T , etc.

Ca orice transformare, care se aplică unei secvențe de durată finită, și această transformare prezintă erori la capetele intervalului de timp considerat. Pentru primele eșantioane ale secvenței $x[n]$, filtrele $h[n]$ și $g[n]$ încă nu sunt în regim permanent iar, la terminarea secvenței $x[n]$, filtrele folosite nu sunt încă relaxate. Pentru diminuarea acestor erori, sunt prezentate diferite metode în [29]. Dacă se dorește realizarea unei TUD pe blocuri atunci, pentru diminuarea erorilor provocate de problemele de la marginile blocurilor, se poate aplica una din metodele denumite "overlap and add" sau "overlap and save" [30].

Transformarea TUD este caracterizată de câțiva parametri. Unul dintre aceștia este expresia răspunsului la impuls $h[n]$, (adică a funcției de scalare și respectiv a funcției "wavelets mother"). Conform [29], acesta trebuie corelat cu forma semnalului $x[n]$. În cazul în care semnalul $x[n]$ variază rapid este preferabil să se utilizeze un filtru cu răspuns la impuls mai scurt. Există aplicații în care este necesar ca răspunsul la impuls $h[n]$ să se modifice pe parcursul calculului transformatei TUD [31]. **Una dintre preocupările cele mai interesante ale autorului acestei teze a fost cea mai bună alegere a răspunsului la impuls $h[n]$, pe baza semnalului de prelucrat, pentru maximizarea factorului de compresie, la o distorsiune de reconstrucție impusă. Această tehnică este prezentată în [43].**

Un alt parametru al transformării este numărul său de iterații M . În exemplul dat pentru descrierea algoritmului de calcul al transformării s-a folosit pentru M valoarea sa maximă posibilă. Nu este însă necesar ca lungimea secvenței $s_M[n]$ din structura vectorului Y să fie minimă (adică 2). Există aplicații în care lungimea secvenței $s_M[n]$ din structura vectorului Y este mai mare.

În sfârșit, un ultim parametru al TUD este lungimea secvenței de intrare, N . Aceasta trebuie să fie o putere a lui 2. Pentru o alegere convenabilă este posibil să avem nevoie de o transformare pe blocuri [32].

Pe lângă utilizarea sa la compresie, transformarea "wavelet" discretă mai are și alte aplicații. Câteva dintre acestea sunt prezentate în: [33]... [39].

2.1.4. Pachete de funcții "wavelet"

În continuare se prezintă o generalizare a noțiunii de analiză multirezoluție care contribuie la dezvoltarea teoriei funcțiilor "wavelet". Este vorba de noțiunea de pachet de funcții "wavelet". Aceasta conduce la o nouă transformare "wavelet" discretă, transformarea cu pachete de funcții "wavelet" discretă (TPWD), care generalizează TUD.

Parametrii unei TUD sunt numărul de iterații și undișoara mamă utilizate. Alegerea acestor parametri în acord cu aplicația considerată este decisivă pentru găsirea celei mai bune rezolvări a problemei respective. Utilizarea TPWD ajută la rezolvarea acestei probleme de alegere.

O generalizare simplă dar foarte utilă a noțiunilor de funcție "wavelet" respectiv de analiză multirezoluție apare în cazul pachetelor de funcții "wavelet". Pentru a introduce această noțiune este utilă folosirea următoarei notații:

$$m_e(\omega) = m_0^e(\omega) m_1^{1-e}(\omega), \quad e = 0, 1$$

Observația fundamentală care stă la baza construcției pachetelor de funcții "wavelet" este așa numitul artificiu de împărțire. Se presupune că mulțimea de funcții $\{f(\tau - k)\}_{k \in Z}$ este o bază ortonormală a

spațiului Hilbert S . Atunci funcțiile $f_k^0(\tau) = \frac{1}{\sqrt{2}} \cdot f^0\left(\frac{\tau}{2} - k\right)$ și $f_k^1(\tau) = \frac{1}{\sqrt{2}} \cdot f^1\left(\frac{\tau}{2} - k\right)$, $k \in Z$

unde $F\{f^e\}(\omega) = m_e\left(\frac{\omega}{2}\right) F\{f\}\left(\frac{\omega}{2}\right)$, constituie de asemenea o bază ortonormală a spațiului S , $\{f_k^0(\tau), f_k^1(\tau)\}_{k \in Z}$.

O analiză multirezoluție clasică este obținută împărțind spațiile V_m , folosind artificiuul descris mai sus în spațiile V_{m-1} și W_{m-1} și apoi făcând la fel, într-un mod recursiv, pentru spațiul V_{m-1} .

Pachetele de funcții "wavelet" sunt funcțiile elemente ale bazelor ortonormale care se obțin dacă se folosește artificiuul de împărțire și pentru spațiile W_m , cu alte cuvinte dacă se utilizează un sistem de codare în subbenzi cu structură simetrică. Pornind de la spațiul V_m , se obțin, după aplicarea de L ori a artificiuului de împărțire, funcțiile (elemente ale unor baze ortonormale):

$$\Psi_{e_1, \dots, e_L; m, k}^L(\tau) = 2^{-\frac{(m-L)}{2}} \Psi_{e_1, \dots, e_L}^L(2^{m-L} \tau - k)$$

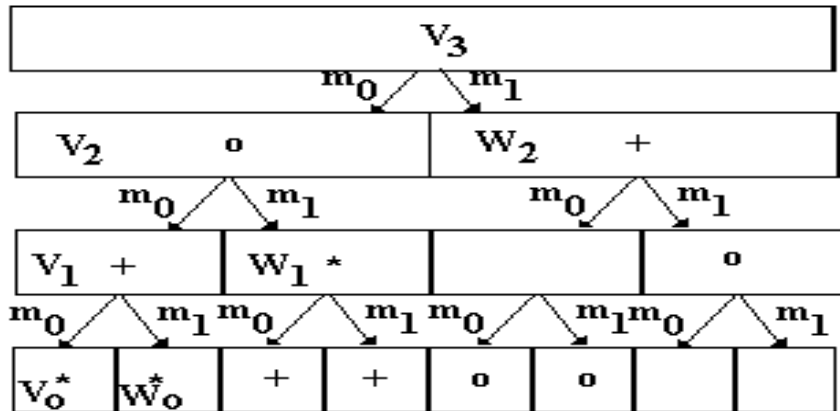


Figura 2.1.4.1 Schemă pentru generarea pachetelor de funcții "wavelet".

cu:

$$F \left\{ \psi_{e_1, \dots, e_L}^L \right\} (\omega) = \prod_{l=1}^L m_{e_l} (2^{-l} \omega) \cdot F\{\phi\} (2^{-L} \omega)$$

Astfel, după \$L\$ împărțiri, se obțin \$2^L\$ funcții de bază și translatatele lor cu întregi multipli de \$2^{L-m}\$ ca și elemente ale bazei ortonormale a spațiului \$V_m\$. Legătura dintre pachetele de funcții "wavelet" și funcțiile de scară respectiv funcțiile "wavelet" corespunzătoare este:

$$\phi(\tau) = \psi_{0, \dots, 0}^L(\tau) \quad \text{și} \quad \psi(\tau) = \psi_{1, 0, \dots, 0}^L(\tau)$$

De fapt nu este necesar să se împartă fiecare subspațiu pentru fiecare valoare a lui \$m\$. În figura 2.1.4.1. se prezintă o modalitate de împărțire a spațiului \$V_3\$ corespunzătoare schemei care generează pachete de funcții "wavelet". În figură sunt notate cu * spațiile care aparțin unei analize multirezoluție:

$$V_3 = V_0 \oplus W_2 \oplus W_1 \oplus W_0$$

Cu \$^\circ\$ s-au notat spațiile care pot participa la construcția unui pachet de funcții "wavelet". Baza ortonormală a lui \$V_0\$, corespunzătoare pachetului de funcții "wavelet" ales în acest exemplu este \$\left\{ \psi_0^1(4\tau - k), \psi_{1,1}^2(2\tau - k), \psi_{1,0,0}^3(\tau - k), \psi_{1,0,1}(\tau - k) \right\}_{k \in \mathbb{Z}}\$. Un alt pachet de funcții "wavelet" poate fi construit dacă se aleg funcțiile notate cu + în figura 2.1.4.1. Acestui nou pachet de funcții "wavelet" îi corespunde următoarea bază ortonormală a lui \$V_3\$, \$\left\{ \psi_1^1(4\tau - k), \psi_{1,0}^2(2\tau - k), \psi_{0,1,0}^3(\tau - k), \psi_{0,1,1}(\tau - k) \right\}_{k \in \mathbb{Z}}\$.

Transformările "wavelet" discrete directă și inversă, corespunzătoare primului exemplu de pachet de funcții "wavelet" dat mai sus sunt prezentate în figura 2.1.4.2.

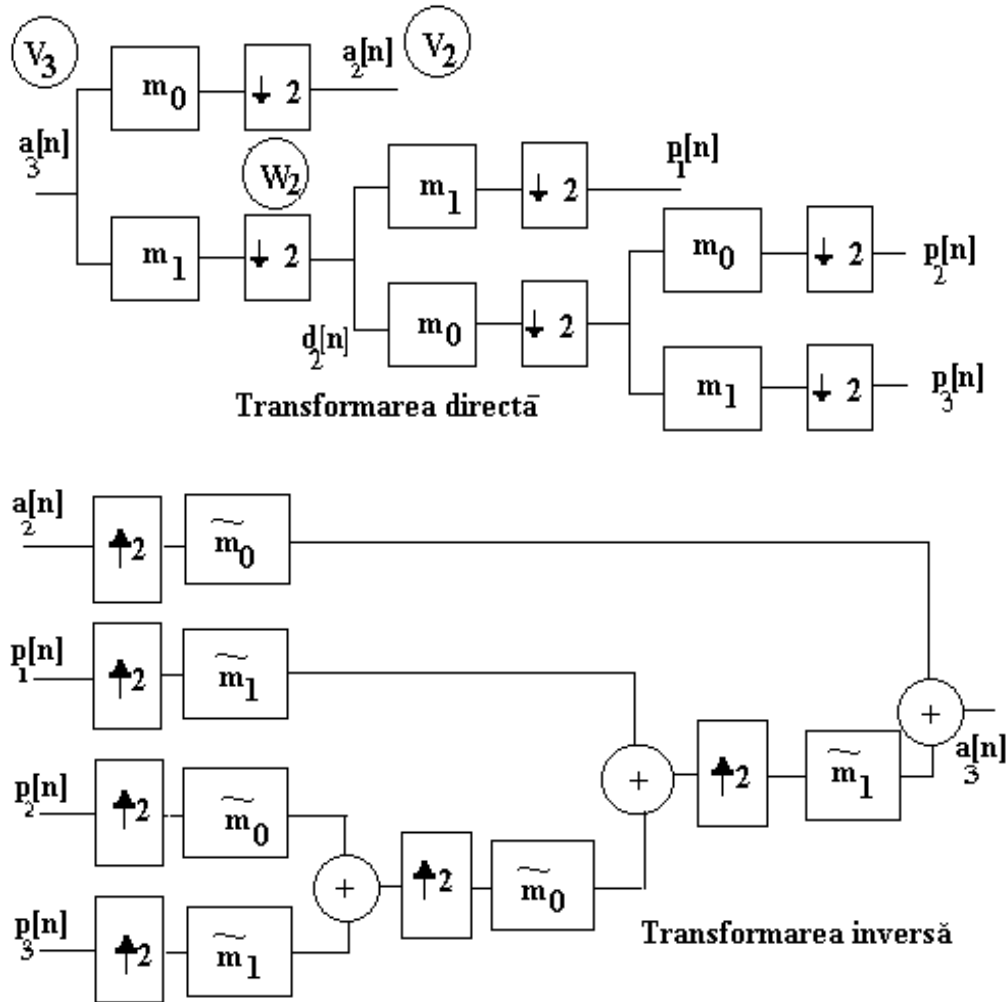


Figura 2.1.4.2. Transformările "wavelet" discrete, directă și inversă corespunzătoare primului exemplu de pachet de funcții "wavelet".

Principalul avantaj al pachetelor de funcții "wavelet" este că avem mult mai multă libertate în alegerea bazei în care să descompunem semnalul pe care dorim să-l analizăm. **Având la dispoziție un număr mai mare de subbenzi, se poate îmbunătăți localizarea frecvențială a componentelor semnalului de analizat.** Există criterii de alegere a bazei în acord cu semnalul de analizat. M. Wikerhauser a propus un astfel de criteriu. Procedura introdusă de el se numește "alegerea celei mai bune baze". O prezentare exhaustivă a acestui concept este făcută în [7]. Teoria pachetelor de funcții "wavelet" este prezentată și în [40].

2.1.4.1. Alegerea celei mai bune baze

Poate fi aleasă în consecință o submulțime cu caracteristici de bază, adaptată la un semnal particular sau la o problemă particulară. Când există posibilitatea alegerii unei baze pentru reprezentarea unui anumit semnal atunci poate fi căutată cea mai bună bază din punctul de vedere al unui anumit criteriu. În acest mod se obține baza adaptată la semnalul considerat. Fie B o mulțime numărabilă de baze ale spațiului Hilbert separabil X . Se prezintă o listă de proprietăți utile pentru mulțimea B .

- Calculul rapid al produselor scalare cu elementele bazelor din B ,

- Localizarea temporală bună a elementelor bazelor din B ,
- Localizarea frecvențială bună a elementelor bazelor din B ,
- Independență, astfel încât să nu existe multe elemente ale unei baze care să se potrivească cu o anumită porțiune din semnalul de analizat.

Pentru a alege cea mai bună bază este necesar să se aprecieze în ce măsură fiecare bază din B are proprietățile enunțate mai sus.

Înainte de a putea defini o reprezentare optimă, este necesar să se poată aprecia care este costul memorării unei anumite reprezentări. Vom numi acest cost, cost de informație. Fiind dată secvența $u[k]$ se poate defini o funcțională de cost de informație corespunzătoare secvenței $u[k]$ prin:

$$M(u) = \sum_{k \in Z} \mu \left(|u[k]| \right), \quad \mu(0) = 0$$

unde μ este o funcție reală definită pe $[0, \infty)$.

Pentru orice element $x \in X$ se definește $u[k] = \langle b_k, x \rangle$, unde b_k este cel de al k -lea element al bazei B din B . Costul de informație al reprezentării lui x în baza B este $M \left(\langle b_k, x \rangle \right)$. S-a definit în acest mod funcționala M_x pe B :

$$M_x : B \rightarrow \mathbb{R}, \quad B \rightarrow M \left(\langle b_k, x \rangle \right)$$

S-a obținut în acest mod costul de informație M al lui x în baza B . Cea mai bună bază din B pentru semnalul x , în raport cu costul de informație M , este acea bază B pentru care $M \left(\langle b_k, x \rangle \right)$

are valoarea minimă. În continuare se prezintă câteva exemple de funcționale de cost de informație.

Exemplul 1. Numărul de eşantioane peste un anumit prag.

Se fixează un prag ε și se numără elementele din secvența $u[n]$ a căror valoare obținută depășește pragul.

$$\mu(w) = \begin{cases} |w|, & |w| \geq \varepsilon \\ 0, & |w| < \varepsilon \end{cases}$$

Exemplul 2. Concentrarea în spațiul l^p , $0 < p < 2$.

$$\mu(w) = |w|^p$$

$$M(u) = \left\| u \right\|_p^p$$

Exemplul 3. Entropia.

Entropia secvenței $u[n]$ este definită cu relația:

$$E(u) = \sum_k p(k) \cdot \log \frac{1}{p(k)}$$

unde:

$$p(k) = \frac{|u[k]|^2}{\|u[k]\|^2}; \quad p \log \frac{1}{p} = 0 \text{ dacă } p = 0$$

Funcționala:

$$l(u) = \sum_k |u[k]|^2 \cdot \log \frac{1}{|u[k]|^2}$$

este o funcțională de cost de informație.

Exemplul 4. Logaritmul energiei.

$$M(u) = \sum_{k=1}^N \log |u[k]|^2$$

Mulțimea B poate fi numită "bibliotecă" de baze. Dacă biblioteca de baze este un arbore de "înălțime" finită L (există L nivele de descompunere), atunci cea mai bună bază pentru un semnal x poate fi determinată prin calculul costului de informație în fiecare "nod" al arborelui și prin compararea nodului "copil" cu nodul "părinte", începând de jos.

În acest mod fiecare nod este examinat de două ori, odată considerându-se că este un nod "copil" și a doua oară considerându-se că este nod "părinte". Acest algoritm de căutare este exemplificat în figurile următoare. În figura 2.1.4.3 au fost plasate numere în interiorul nodurilor arborelui pentru a specifica costurile de informație. Se marchează cu asterisc toate nodurile de pe nivelul de jos. Costul lor total de informație este egal cu 36. Se încearcă reducerea acestei valori. Ori de câte ori un nod "părinte" are un cost de informație inferior costului total de informație al nodurilor sale "copii", acest nod "părinte" se marchează cu un asterisc. Dacă nodul "părinte" are un cost de informație superior celui al nodurilor

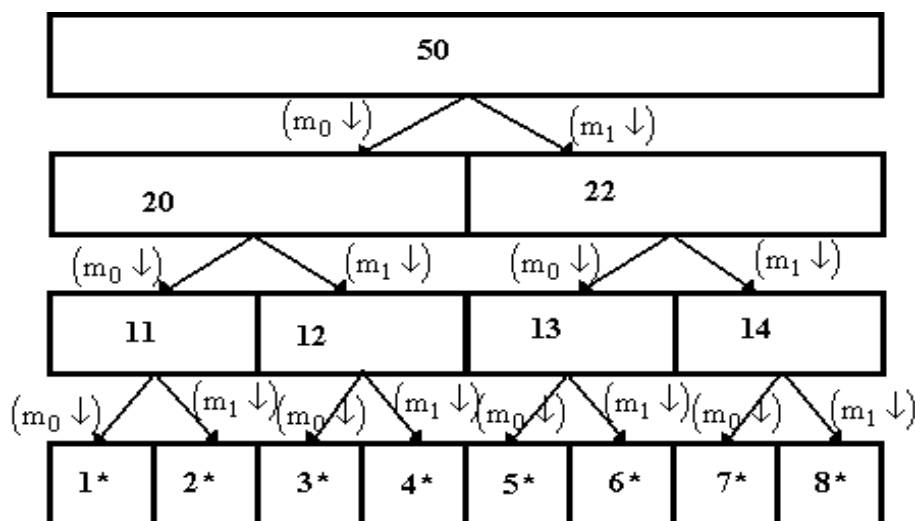


Figura 2.1.4.3. Inițializarea algoritmului de căutare a celei mai bune baze.

"copii" care-i corespund, atunci acest nod "părinte" nu se marchează cu asterisc ci i se alocă costul total de informație al nodurilor "copii" care îi corespund.

În figura 2.1.4.4 aceste costuri de informație transferate sunt prezentate între paranteze. Rezultatul căutării celei mai bune baze este prezentat în figura 2.1.4.5. Se constată că a avut loc o reducere a costului de informație de la valoarea 50 la valoarea 32.

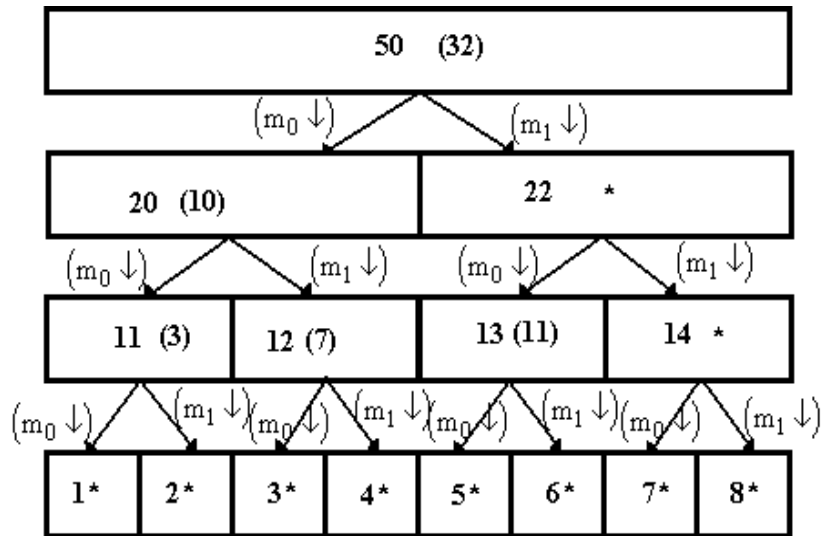


Figura 2.1.4.4 Primul pas al algoritmului de căutare a celei mai bune baze.

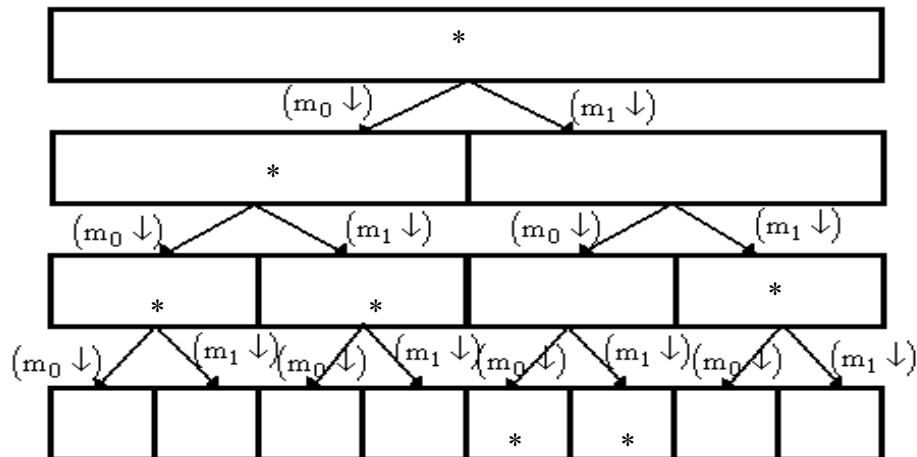


Figura 2.1.4.5 Rezultatul algoritmului de căutare a celei mai bune baze.

Trebuie remarcat că această metodă de căutare se referă doar la bazele generate de o anumită funcție "wavelets mother". Dar pentru o anumită aplicație și pentru un anumit semnal, existând mai multe funcții "wavelets mother" care s-ar putea folosi, există și o cea mai bună funcție "wavelets mother". Pentru aplicațiile de compresie această funcție optimă poate fi determinată folosind metoda propusă de autorul acestei teze în [43].

2.1.4.2. Pachete de funcții "wavelet" de tip Malvar

O descompunere în forme de undă Malvar a semnalului $x(\tau)$ asociată partiției intervalului $[0, T]$:

$$[0, T] = \bigcup_k I_k$$

cu:

$$I_k = [a_k, a_{k+1}]$$

este descrisă de relația:

$$x(\tau) = \sum_{m,k} c_{m,k} \psi_{m,k}(\tau)$$

unde:

$$\psi_{m,k}(\tau) = w_m(\tau) \cdot g_{m,k}(\tau)$$

și:

$$g_{m,k}(\tau) = \frac{\sqrt{2}}{\sqrt{|I_m|}} \cdot \cos \left[\frac{\pi}{|I_m|} \left(k + \frac{1}{2} \right) (\tau - a_m) \right]$$

cu:

$$w_m(\tau) = \begin{cases} b_m(\tau), & \tau \in [a_m - r, a_m + r] \\ 1, & \tau \in [a_m + r, a_{m+1} - r] \\ b_{m+1}(2a_{m+1} - \tau), & \tau \in [a_{m+1} - r, a_{m+1} + r] \end{cases}$$

unde:

$$b_m(\tau) = \sin \left[\frac{\pi}{4} \left(1 + \sin \left(\frac{\pi}{2} \cdot \frac{(\tau - a_m)}{r} \right) \right) \right]$$

Aceste funcții au fost introduse de Malvar, [21], [41], din dorința de a diminua distorsiunile date de problemele de margine, care apar la compresia imaginilor, când se folosește, pe blocuri, transformarea DCT, ca și transformare ortogonală.

Fie U_m un șir de operatori definiți pe $L^2[a_m - r, a_m + r]$, care transformă semnalul $x(\tau)$ în șirul de funcții:

$$U_m \{x(\tau)\} = \begin{cases} b_m(\tau) x(\tau) + b_m(2a_m - \tau) x(2a_m - \tau), & \tau \in (a_m, a_m + r] \\ b_m(2a_m - \tau) x(\tau) - b_m(\tau) x(2a_m - \tau), & \tau \in [a_m - r, a_m) \end{cases}$$

Se poate scrie:

$$U_m \{x(\tau)\} = \sum_k d_{m,k} \cdot \phi_{m,k}(\tau)$$

unde:

$$\phi_{m,k}(\tau) = \chi_{I_m}(\tau) \cdot g_{m,k}(\tau)$$

Meyer, a fost primul care a demonstrat că aceste funcții sunt elemente ale unui pachet de funcții "wavelet", [42]. El le-a numit funcții "wavelet" de tip Malvar.

Cu $\chi_{I_m}(\tau)$ s-a notat funcția caracteristică a intervalului I_m , iar $U_m \{x(\tau)\}$ reprezintă produsul segmentării semnalului $x(\tau)$ corespunzător intervalului I_m . În continuare se consideră că semnalul $x(\tau)$ este definit pe $[0, T]$. Acest interval suferă următoarea împărțire:

$$\begin{aligned}
I_0^0 &= [0, T] \\
I_0^1 &= [0, T/2]; \quad I_1^1 = [T/2, T] \\
I_0^{11} &= [0, T/4]; \quad I_0^{12} = [T/4, T/2] \\
I_1^{11} &= [T/2, 3T/4]; \quad I_1^{12} = [3T/4, T] \\
&\dots
\end{aligned}$$

Wikerhauser a demonstrat că funcțiile "wavelet" de tip Malvar, obținute prin folosirea partiției descrisă în ultima relație se constituie într-un pachet de funcții "wavelet". El a numit pachetele de acest tip cosinusoidale, sau pachete de funcții trigonometrice locale. Căutând cea mai bună bază în acest pachet, folosind, de exemplu, criteriul de minimizare a entropiei, se obține mulțimea $\{U_m \{x(\tau)\}\}_{m \in I}$. Fiecare element al acestei mulțimi reprezintă un segment, corespunzător unui anumit interval al partiției descrisă de ultima relație, al semnalului $x(\tau)$. Fiecare astfel de segment reprezintă un semnal cvasistaționar. Oricare dintre aceste semnale poate fi ulterior supus compresiei.

Pachetele cosinusoidale rezolvă o deficiență cronică a pachetelor de funcții "wavelet", și anume localizarea în timp. Orice pachet de funcții "wavelet" corespunde unei anumite bănci de filtre de analiză, care realizează o anumită codare în subbenzi. Pe tot parcursul calculului TPWD, aceste filtre rămân neschimbate. În consecință, TPWD nu realizează nici o localizare în timp a acestei bănci de filtre. În cazul TPC, în fiecare interval I_m , se lucrează cu o altă funcție "wavelets mother", deci cu o altă bancă de filtre de analiză, făcându-se o localizare temporală a acestor bănci.

2.2. Utilizarea transformărilor ortogonale prezentate la compresia de date

În acest paragraf vom presupune că se utilizează o transformare ortogonală de tip "wavelet" (transformarea "wavelet" discretă, transformarea cu pachete de funcții "wavelet" discretă, sau transformarea cu pachete cosinusoidale discretă). Pentru început se justifică de ce a fost făcută această alegere. În acest scop se demonstrează că aceste transformări converg asimptotic la transformarea Karhunen-Loeve. **În consecință ele realizează o decorelare importantă a semnalului de prelucrat.**

2.2.1. Analiza statistică a TUD

Prima transformare care va fi studiată este TUD. Fie semnalul aleator staționar $x(\tau)$. Proiecțiile acestui semnal pe descompunerea ortogonală a lui $L^2(\mathbb{R})$, $\{W_m\}_{m \in Z}$, generată de funcția "wavelets mother" $\psi(\tau)$ sunt semnale ale căror descompuneri în bazele $\{\psi_{m,n}(\tau)\}_{n \in Z}$ au coeficienții $d_m[n]$:

$$d_m[n] = \langle x(\tau), \psi_{m,n}(\tau) \rangle$$

Se calculează autocorelațiile semnalelor $d_m[n]$:

$$\begin{aligned}
r_{d_m}[k, l] &= E \{ d_m[k], d_m^*[l] \} = E \left\{ \langle x(\tau), \psi_{m,k}(\tau) \rangle, \langle x(\tau), \psi_{m,l}(\tau) \rangle^* \right\} = \\
&= E \left\{ \int_{\mathbb{R}} x(\tau) \cdot \psi_{m,k}^*(\tau) d\tau \int_{\mathbb{R}} x^*(u) \cdot \psi_{m,l}(u) du \right\} = \\
&= E \left\{ \int_{\mathbb{R}^2} x(\tau) \cdot x^*(u) \cdot \psi_{m,k}^*(\tau) \cdot \psi_{m,l}(u) du d\tau \right\} = \\
&= \int_{\mathbb{R}^2} E \{ x(\tau) \cdot x^*(u) \} \cdot \psi_{m,k}^*(\tau) \cdot \psi_{m,l}(u) du d\tau
\end{aligned}$$

Deoarece semnalul $x(\tau)$ este staționar:

$$E\{x(\tau) \cdot x^*(u)\} = r_{xx}(\tau - u)$$

În consecință:

$$r_{d_m}[k, l] = \int_{\mathbb{R}^2} r_{xx}(\tau - u) \cdot \psi_{m,k}^*(\tau) \cdot \psi_{m,l}(u) du d\tau =$$

$$\int_{\mathbb{R}} \psi_{m,k}^*(\tau) \int_{\mathbb{R}} \psi_{m,l}(u) \cdot r_{xx}(\tau - u) du = \int_{\mathbb{R}} (r_{xx}(\tau) * \psi_{m,l}(\tau)) \cdot \psi_{m,k}^*(\tau) d\tau$$

sau pe baza relației lui Parseval:

$$r_{d_m}[k, l] = \frac{1}{2\pi} \int_{\mathbb{R}} F\{r_{xx}(\tau) * \psi_{m,l}(\tau)\}(\omega) \cdot F^*\{\psi_{m,k}(\tau)\}(\omega) d\omega$$

Admițând că se lucrează cu funcții "wavelet" reale, ultima relație devine:

$$r_{d_m}[k, l] = \frac{1}{2\pi} \int_{\mathbb{R}} F\{r_{xx}(\tau)\}(\omega) \cdot F\{\psi_{m,l}(\tau)\}(\omega) \cdot F^*\{\psi_{m,k}(\tau)\}(\omega) d\omega$$

Dar:

$$F\{\psi_{m,l}(\tau)\}(\omega) = 2^{-\frac{m}{2}} \cdot e^{-j\omega 2^{-m}l} \cdot F\{\psi\}(2^{-m}\omega)$$

Deci:

$$r_{d_m}[k, l] = \frac{1}{2\pi} \int_{\mathbb{R}} F\{r_{xx}(\tau)\}(\omega) \cdot 2^{-\frac{m}{2}} \cdot e^{-j\omega 2^{-m}l} \cdot$$

$$F\{\psi\}(2^{-m}\omega) \cdot 2^{-\frac{m}{2}} \cdot e^{j\omega 2^{-m}k} \cdot F^*\{\psi\}(2^{-m}\omega) d\omega$$

adică:

$$r_{d_m}[k, l] = \frac{1}{2\pi} \int_{\mathbb{R}} F\{r_{xx}(\tau)\}(\omega) \cdot 2^{-m} \cdot e^{-j\omega 2^{-m}(l-k)} \cdot \left| F\{\psi\}(2^{-m}\omega) \right|^2 d\omega$$

Integrala din membrul drept se poate descompune într-o serie de integrale calculate pe intervale de lungime $2\pi \cdot 2^m$.

$$r_{d_m}[k, l] = \frac{1}{2\pi} \sum_{p=-\infty}^{\infty} \int_{(2p-1)\pi \cdot 2^m}^{(2p+1)\pi \cdot 2^m} F\{r_{xx}(\tau)\}(\omega) \cdot 2^{-m} \cdot e^{-j\omega 2^{-m}(l-k)} \cdot \left| F\{\psi\}(2^{-m}\omega) \right|^2 d\omega$$

Cu schimbarea de variabilă:

$$v = 2^{-m} \omega$$

ultima relație devine:

$$r_{d_m}[k, l] = \frac{1}{2\pi} \sum_{p=-\infty}^{\infty} \int_{(2p-1)\pi}^{(2p+1)\pi} F \left\{ r_{xx}(\tau) \right\} (2^m v) \cdot e^{-j(l-k)v} \cdot \left| F \left\{ \psi \right\} (v) \right|^2 dv$$

Cu schimbarea de variabilă:

$$w = v - 2p\pi$$

rezultă:

$$r_{d_m}[k, l] = \frac{1}{2\pi} \sum_{p=-\infty}^{\infty} \int_{-\pi}^{\pi} F \left\{ r_{xx}(\tau) \right\} (2^m (w + 2p\pi)) \cdot e^{-jw(l-k)} \cdot \left| F \left\{ \psi \right\} (w + 2p\pi) \right|^2 dw$$

sau:

$$r_{d_m}[k, l] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{p=-\infty}^{\infty} F \left\{ r_{xx}(\tau) \right\} (2^m (w + 2p\pi)) \cdot e^{-jw(l-k)} \cdot \left| F \left\{ \psi \right\} (w + 2p\pi) \right|^2 dw \quad (2.2.1.1)$$

Dacă semnalul aleator $x(\tau)$ este un zgomot alb de valoare medie nulă și dispersie unitară atunci:

$$F \left\{ r_{xx}(\tau) \right\} (2^m (w + 2p\pi)) = 1, \quad (\forall) w \in \mathbb{R}$$

În acest caz, relația (2.2.1.1) devine:

$$r_{d_m}[k, l] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{p=-\infty}^{\infty} \left| F \left\{ \psi \right\} (w + 2p\pi) \right|^2 \cdot e^{-jw(l-k)} dw \quad (2.2.1.2)$$

În continuare se calculează suma din membrul drept al acestei relații. În acest scop se pornește de la proprietatea de ortogonalitate a mulțimii $\{\psi(x-k)\}_{k \in \mathbb{Z}}$:

$$\langle \psi(x), \psi(x-k) \rangle = \delta[k]$$

Pe baza relației lui Parseval se obține :

$$\frac{1}{2\pi} \langle F\{\psi\}(\omega), e^{-j\omega k} F\{\psi\}(\omega) \rangle = \delta[k]$$

adică :

$$\int_{-\infty}^{\infty} \left| F \left\{ \psi \right\} (\omega) \right|^2 e^{j\omega k} d\omega = 2\pi \cdot \delta[k]$$

sau :

$$\sum_p \int_{(2p-1)\pi}^{(2p+1)\pi} \left| F \{ \psi \} (\omega) \right|^2 e^{j\omega k} d\omega = 2\pi \cdot \delta[k]$$

Dacă se face schimbarea de variabilă $\omega - 2p\pi = u$ se obține :

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_p \left| F \{ \psi \} (u + 2p\pi) \right|^2 e^{ju k} du = \delta[k] \quad (2.2.1.3)$$

Membrul stâng al ultimei relații reprezintă cel de-al k-lea coeficient al descompunerii în serie Fourier a unei funcții periodice de perioadă 2π . Relația (2.2.1.3.) arată că această funcție are doar un singur coeficient Fourier nenul și anume acela cu indicele 0 (adică componenta continuă). Rezultă că funcția considerată este constantă. Având în vedere că pentru $k=0$ relația (2.2.1.3) devine :

$$\int_{-\pi}^{\pi} \sum_{p=-\infty}^{\infty} \left| F \{ \psi \} (w + 2p\pi) \right|^2 du = 2\pi$$

rezultă că valoarea acestei constante este 1 . S-a demonstrat în acest mod că :

$$\sum_{p=-\infty}^{\infty} \left| F \{ \psi \} (w + 2p\pi) \right|^2 = 1$$

Relația (2.1.1.2) devine:

$$r_{d_m} [k, l] = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-jw(l-k)} dw$$

În consecință:

$$r_{d_m} [k, l] = \delta[k - l]$$

Cu alte cuvinte, dacă $x(\tau)$ este un zgomot alb atunci toate semnalele $d_m[n]$ sunt zgomote albe în timp discret. Deci prin aplicarea transformatei wavelet discretă unui semnal în timp discret de tip zgomot alb se obțin tot semnale în timp discret de tip zgomot alb. Evident eșantioanele acestor semnale sunt necorelate.

Se trece la limită pentru m tinzând la $-\infty$, în cei doi membri ai relației (2.1.1.1). Se obține:

$$r_{d_{-\infty}} [k, l] = \frac{1}{2\pi} \int_{-\pi}^{\pi} F \{ r_{xx}(\tau) \} (0) \cdot e^{-jw(l-k)} \cdot \sum_{p=-\infty}^{\infty} \left| F \{ \psi \} (w + 2p\pi) \right|^2 dw$$

sau:

$$r_{d_{-\infty}}[k, 1] = \frac{1}{2\pi} F \left\{ r_{xx}(\tau) \right\}(0) \cdot \int_{-\pi}^{\pi} e^{-jw(l-k)} dw = F \left\{ r_{xx}(\tau) \right\}(0) \cdot \delta[k-1] \quad (2.2.1.4)$$

Deci semnalul $d_{-\infty}[n]$ este un zgomot alb în timp discret. Eșantioanele sale sunt în consecință necorelate. Se poate așadar afirma că transformarea "wavelet" discretă converge asimptotic la transformarea Karhunen-Loeve. Demonstrația de mai sus a fost reprodusă din [4]. O analiză asemănătoare pentru pachete de funcții "wavelet" este prezentată în [44]. Iată de ce este recomandată utilizarea transformării "wavelet" discretă la compresia datelor. În figura 2.2.1.1. se ilustrează rezultatul obținut. În partea de sus a acestei imagini se prezintă densitatea spectrală de putere a unui zgomot colorat. Acesta a fost generat filtrând un zgomot alb cu ajutorul unui mediator numeric alunecător având o fereastră de analiză de lungime 20. În partea de jos a aceleiași figuri a fost reprezentată densitatea spectrală de putere a semnalului obținut prin aplicarea TUD acestui zgomot colorat. Se constată că în urma aplicării TUD s-a obținut un zgomot alb.

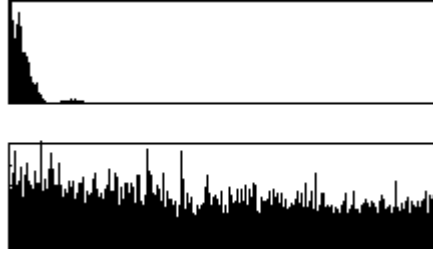


Figura 2.2.1.1. Evidențierea efectului de albire a TUD.

În [39] se demonstrează că oricare ar fi semnalul care trebuie compresat există o anumită funcție "wavelet" care poate conduce la realizarea unei compresii superioare celei care s-ar putea obține folosind transformarea cosinus discretă. Este vorba de obținerea unui factor de compresie superior la erori de aproximare egale. În [29] este prezentată o metodă de compresie cu pierdere de informație controlată. Această metodă are următorii pași:

1. Se calculează $y[n]$, TUD a semnalului care trebuie prelucrat, $x[n]$.
2. Se realizează compresia propriuzisă prin înlăturarea acelor eșantioane ale semnalului obținut la pasul anterior care sunt mai mici decât un prag impus. Se obține semnalul $\hat{y}[n]$.
3. Se calculează TUD inversă obținându-se semnalul $\hat{x}[n]$.

Pragul de la pasul 2 se impune într-un mod adaptiv. El se alege în așa fel încât eroarea medie pătratică de aproximare a semnalului $x[n]$ prin semnalul $\hat{x}[n]$ să nu depășească un procent din energia semnalului $x[n]$.

Sunt valabile relațiile:

$$E_x = \sum_{k=0}^{N-1} x^2[k] = \sum_{k=0}^{N-1} y^2[k] \quad (2.2.1.5)$$

$$E_{\hat{x}} = \sum_{k=0}^{M-1} \hat{x}^2[k] = \sum_{k=0}^{M-1} \hat{y}^2[k] \quad (2.2.1.6)$$

deoarece orice transformare ortogonală conservă energia, [45].

De asemenea se poate scrie:

$$\hat{y}[n] = \begin{cases} y[n], & \text{daca } |y[n]| > P \\ 0, & \text{in rest} \end{cases}$$

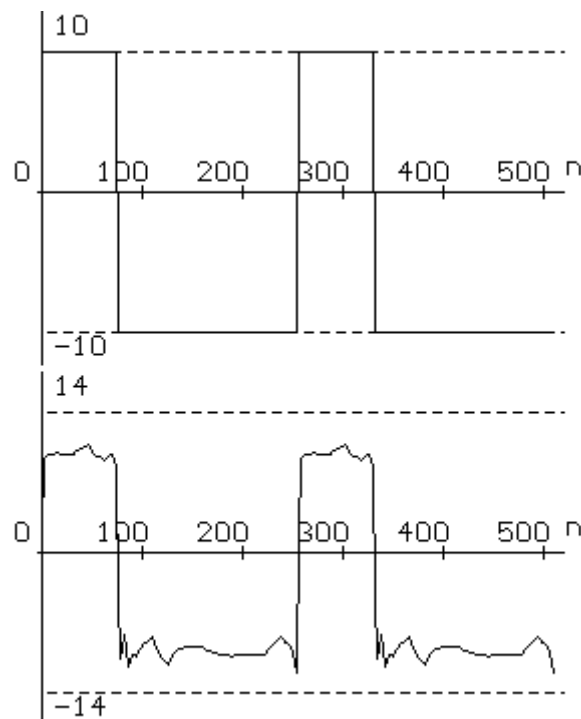
Fie $\circ \hat{y}[n]$ secvența obținută prin ordonarea descrescătoare a eșantioanelor semnalului $\hat{y}[n]$. Eroarea medie pătratică de aproximare a semnalului $x[n]$ prin semnalul $\hat{x}[n]$ este proporțională cu:

$$\varepsilon = \sum_{k=M}^{N-1} \circ \hat{y}^2[k]$$

Valoarea lui M se obține prin rezolvarea ecuației:

$$\max_{M \in Z} \varepsilon = \frac{E_x}{100}$$

În continuare se prezintă un exemplu de aplicare a acestei metode. În figura 2.2.1.2, este prezentat un exemplu de semnal de compresat, $x[n]$ (în partea de sus) și corespunzător $\hat{x}[n]$ (în partea de jos). Este indicat numărul de eșantioane din care a fost reconstruit semnalul $\hat{x}[n]$. Se constată că s-a obținut o valoare mare a factorului de compresie. Semnalul $x[n]$ are 512 eșantioane.



Reconstrucție din 32 de eșantioane.

Figura 2.2.1.2. Compresia unui semnal dreptunghiular. Factorul de compresie este de 8.

Recent a fost construit un nou dicționar timp-frecvență de baze ortonormale asemănător dicționarilor de pachete de funcții "wavelet", [46]. Elementele sale se generează cu ajutorul unor

transformări Karhunen-Loeve localizate. Și în acest dicționar elementele necesare pentru descompunerea unui anumit semnal pot fi găsite folosind algoritmul de căutare al celei mai bune baze deja prezentat. În [91] se demonstrează superioritatea TUD asupra transformării Karhunen-Loeve la compresia semnalelor care pot fi modelate prin procese aleatoare ne-Gaussiene.

O formulă analoagă relației (2.2.1.1) poate fi demonstrată și în cazul semnalelor $s_m[n]$, $m=1 \div M$. Aceasta este :

$$R_{s_m}[k-l] \leftrightarrow \sum_p F \left\{ r_{xx} \right\} \left(2^m (\Omega + 2p\pi) \right) \cdot \left| F \left\{ \varphi \right\} \left(\Omega + 2p\pi \right) \right|^2 \quad (2.2.1.7)$$

Demonstrația este identică cu cea pentru relația (2.2.1.1).

În continuare se calculează mediile și dispersiile semnalelor aleatoare $s_m[m]$ și $d_m[n]$, $m=1 \div M$. Astfel pentru semnalul $d_m[n]$:

$$E \left\{ d_m[n] \right\} = E \left\{ \langle x(t), \psi_{m,k}(t) \rangle \right\} = E \left\{ \int_{-\infty}^{\infty} x(t) \cdot \psi_{m,k}^*(t) dt \right\}$$

sau, aplicând din nou teorema lui Fubini :

$$E \left\{ d_m[k] \right\} = \int_{-\infty}^{\infty} E \left\{ x(t) \right\} \cdot \psi_{m,k}^*(t) dt = \int_{-\infty}^{\infty} M_n \cdot \psi_{m,k}^*(t) dt$$

unde cu M_n s-a notat media semnalului aleator $x(t)$.

Ultima relație se mai poate scrie :

$$E \left\{ d_m[k] \right\} = M_n \cdot F \left\{ \psi_{m,k}^* \right\} (0) \quad (0)$$

Dar :

$$F \left\{ \psi_{m,k} \right\} (\omega) = 2^{\frac{m}{2}} \cdot e^{-j\omega 2^m k} \cdot F \left\{ \psi \right\} (2^m \omega)$$

și, revenind :

$$E \left\{ d_m[k] \right\} = M_n \cdot 2^{\frac{m}{2}} \cdot F \left\{ \psi \right\} (0) \quad (2.2.1.8)$$

Dar :

$$F \left\{ \psi \right\} (0) = m_1(0) \cdot F \left\{ \varphi \right\} (0) = m_1(0)$$

S-a demonstrat în capitolul anterior că:

$$|H(\Omega)|^2 + |G(\Omega)|^2 = 2 \quad (2.2.1.9)$$

unde:

$$h_n \leftrightarrow H(\Omega) \text{ si } g_n \leftrightarrow G(\Omega)$$

Utilizând și relațiile:

$$\begin{aligned} h^{v*}[n] &\leftrightarrow \sqrt{2}m_0(\Omega) \\ g^{v*}[n] &\leftrightarrow \sqrt{2}m_1(\Omega) \end{aligned}$$

vom putea scrie :

$$\begin{aligned} \sqrt{2}m_0(\Omega) &= H^*(\Omega) \\ \sqrt{2}m_1(\Omega) &= G^*(\Omega) \end{aligned}$$

Relația (2.2.1.9) se va scrie, pentru $\Omega=0$:

$$|m_0(0)|^2 + |m_1(0)|^2 = 1 \quad (2.2.1.10)$$

Dar:

$$m_0(0) = 1$$

și pe baza relației (2.2.1.10) vom obține :

$$|m_1(0)| = 0 \quad ; \quad \sum_p g^*[p] = 0 \quad (2.2.1.11)$$

și deci:

$$F\{\psi\}(0) = 0$$

În final, relația (2.2.1.8) se scrie:

$$E\{d_m[k]\} = 0 \quad , \quad m = 1 \div M \quad (2.2.1.12)$$

S-a demonstrat că toate semnalele aleatoare $d_m[n]$ sunt de medie nulă indiferent de m . Acest lucru era de așteptat având în vedere că acești coeficienți sunt obținuți prin folosirea filtrelor cu răspunsurile la impuls $g[n]$ (care sunt filtre trece sus).

În continuare se calculează dispersiile acestor semnale. Având în vedere că media lor este nulă, se obține :

$$E \{ d_m^2[k] \} = R_{d_m}[0]$$

sau, pe baza relației (2.2.1.1):

$$E \{ d_m^2[k] \} = \frac{1}{2\pi} \int_R F \{ r_{xx} \} (2^m u) \cdot \left| F \{ \psi \} (u) \right|^2 du \quad (2.2.1.13)$$

Aceasta este relația care exprimă dispersiile semnalelor $d_m[n]$ pe baza densității spectrale de putere a semnalului aleator $x(t)$.

OBSERVAȚII

O1. Dispersiile semnalelor aleatoare $d_m[n]$ pot fi minimizeze prin alegerea judicioasă a funcției $\psi(t)$ (în acord cu densitatea spectrală de putere a semnalului aleator).

O2. Dacă $x(t)$ este un zgomot alb de medie nulă și dispersie σ^2 atunci:

$$F \{ r_{xx} \} (\omega) = \sigma^2$$

și:

$$E \{ d_m^2[k] \} = \frac{\sigma^2}{2\pi} \cdot \sum_p \int_{(2p-1)\pi}^{(2p+1)\pi} \left| F \{ \psi \} (u) \right|^2 du = \frac{\sigma^2}{2\pi} \cdot \int_{-\pi}^{\pi} \sum_p \left| F \{ \psi \} (u + 2p\pi) \right|^2 du = \sigma^2$$

Deci în cazul în care $x(t)$ este un zgomot alb de medie nulă și dispersie σ^2 atunci semnalele aleatoare $d_m[n]$ sunt tot de tip zgomot alb în timp discret de medie nulă și dispersie σ^2 .

O3. Pentru $m \rightarrow -\infty$ relația (2.1.1.13) devine :

$$E \{ d_{-\infty}^2[k] \} = F \{ r_{xx} \} (0) \quad (2.2.1.14)$$

Această relație descrie comportarea asimptotică a dispersiilor semnalelor aleatoare $d_m[n]$.

În continuare se determină momentele de ordinul I și II ale semnalelor aleatoare $s_m[n]$:

$$E \{ s_m[k] \} = E \{ \langle x(t), \phi_{m,k}(t) \rangle \} = E \left\{ \int_{-\infty}^{\infty} x(t) \cdot \phi_{m,k}^*(t) dt \right\}$$

adică :

$$E \{ s_m[k] \} = \int_{-\infty}^{\infty} E \{ x(t) \} \phi_{m,k}^*(t) dt = \int_{-\infty}^{\infty} M_n \cdot \phi_{m,k}^*(t) dt$$

unde cu M_n s-a notat media semnalului aleator $x(t)$.

Ultima relație se mai poate pune sub forma :

$$E \{ s_m[k] \} = M_n \cdot F \left\{ \phi_{m,k} \right\}^* (0)$$

Deoarece se poate scrie:

$$F \left\{ \varphi_{m,k} \right\} (\omega) = 2^{\frac{m}{2}} \cdot e^{-j\omega 2^m k} \cdot F \left\{ \varphi \right\} (2^m \omega)$$

vom avea :

$$E \left\{ s_m[k] \right\} = M_n \cdot 2^{\frac{m}{2}} \cdot F \left\{ \varphi \right\} (0) = M_n \cdot 2^{\frac{m}{2}} \quad (2.2.1.15)$$

Deci media semnalelor $s_m[k]$ descrește cu creșterea lui m în valori absolute (conform convenției făcute în capitolul anterior m ia valori negative).

Dispersiile acestor semnale sunt :

$$E \left\{ s_m^2[k] \right\} = R_{s_m}[0] - E^2 \left\{ s_m[k] \right\}$$

Valoarea autocorelației în origine este:

$$R_{s_m}[0] = \frac{1}{2\pi} \int_R F \left\{ r_{xx} \right\} (2^{-m} u) \left| F \left\{ \varphi \right\} (u) \right|^2 du$$

și obținem :

$$E \left\{ s_m^2[k] \right\} = \frac{1}{2\pi} \int_R F \left\{ r_{xx} \right\} (2^{-m} u) \left| F \left\{ \varphi \right\} (u) \right|^2 du - 2^m \cdot M_n^2 \quad (2.1.1.16)$$

Dacă semnalul $x(t)$ este de medie nulă atunci:

$$E \left\{ s_m^2[k] \right\} = \frac{1}{2\pi} \int_R F \left\{ r_{xx} \right\} (2^{-m} u) \cdot \left| F \left\{ \varphi \right\} (u) \right|^2 du \quad (2.1.1.17)$$

OBSERVAȚII :

O1. Dispersiile semnalelor aleatoare $s_m[n]$ pot fi minimizeze prin alegerea judicioasă a funcției de scalare $\varphi(t)$ (în acord cu densitatea spectrală de putere $F\{r_{xx}\}(\omega)$).

O2. Dacă $x(t)$ este un zgomot alb de medie nulă și dispersie σ^2 atunci:

$$F\{r_{xx}\}(\omega) = \sigma^2$$

și:

$$E \left\{ s_m^2[k] \right\} = \frac{\sigma^2}{2\pi} \cdot \sum_p \int_{(2p-1)\pi}^{(2p+1)\pi} \left| F \left\{ \varphi \right\} (u) \right|^2 du = \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} \sum_p \left| F \left\{ \varphi \right\} (u + 2p\pi) \right|^2 du = \sigma^2$$

Deci în cazul în care $x(t)$ este un zgomot alb de medie nulă și dispersie σ^2 atunci semnalele aleatoare $s_m[n]$ sunt tot de tip zgomot alb în timp discret de medie nulă și dispersie σ^2 .

O3. Pentru $m \rightarrow \infty$ relația (2.2.1.17) devine:

$$E\{s_m^2[k]\} = F\{r_{xx}\}(0)$$

Această relație descrie comportarea asimptotică a dispersiilor semnalelor aleatoare $s_m[n]$.

O4. Condiția:

$$E\{x(t)\} = 0$$

previne divergența șirurilor $E\{s_m[k]\}$ și $E\{s_m^2[k]\}$ când $m \rightarrow \infty$.

O5. Dacă $x(t)$ este un semnal aleator și staționar de medie nulă atunci secvențele $s_m[n]$ și $d_m[n]$ converg asimptotic (pentru $m \rightarrow \infty$) la semnale aleatoare de tip zgomot alb de medie nulă și dispersie $F\{r_{xx}\}(0)$.

Pe lângă proprietatea de albire TUD are și proprietatea de Gaussianizare. Chiar dacă semnalul a cărui TUD se calculează are o densitate de probabilitate ne-Gaussiană, densitățile de probabilitate ale semnalelor s și d sunt Gaussiene. Această proprietate este ilustrată în figura 2.2.1.3. În partea de sus a acestei figuri este reprezentată densitatea de probabilitate a unui semnal distribuit uniform. În partea de jos a figurii este reprezentată densitatea de probabilitate a semnalului obținut în urma calculării TUD a semnalului distribuit uniform. Se constată că semnalele s și d , din componența transformării calculate sunt distribuite Gaussian.

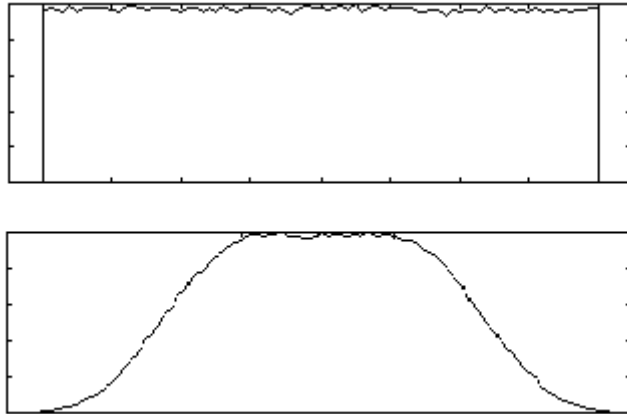


Figura 3.1.1.3. Efectul de Gaussianizare al TUD.

2.2.2. Analiza statistică a TPWD

Spre deosebire de analiza statistică făcută în paragraful anterior, care s-a bazat pe proprietățile matematice specifice teoriei funcțiilor "wavelet", analiza statistică din acest paragraf se bazează pe teoria codării în subbenzi. Se consideră pentru început celulele de bază ale unui sistem de codare în subbenzi, reprezentate în figura 2.2.2.1.

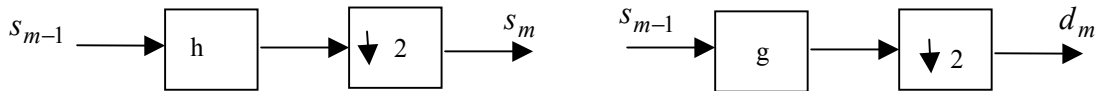


Figura 2.2.2.1. Cele două celule de bază pentru construcția unui codor în subbenzi. Cu ajutorul filtrului h se generează coeficienții de tip s iar cu ajutorul filtrului g coeficienții de tip d.

Relațiile intrare-ieșire pentru cele două sisteme sunt:

$$s_m[n] = \sum_{k=-\infty}^{\infty} h[k] \cdot s_{m-1}[2n-k], \quad d_m[n] = \sum_{k=-\infty}^{\infty} g[k] \cdot s_{m-1}[2n-k] \quad (2.2.2.1)$$

Se calculează autocorelațiile statistice ale semnalelor de la ieșire, considerând semnalele de intrare staționare (proprietate demonstrată în paragraful anterior). Pentru sistemul reprezentat în stânga în figura 2.2.2.1.:

$${}_s r_m[k-1] = E \left\{ s_m[k] \cdot s_m[l] \right\} \stackrel{(2.2.2.1)}{=} E \left\{ \sum_{p=-\infty}^{\infty} s_{m-1}[2k-p] \cdot h[p] \cdot \sum_{q=-\infty}^{\infty} s_{m-1}[2l-q] \cdot h[q] \right\} =$$

$$\sum_{p,q=-\infty}^{\infty} h[p] \cdot h[q] \cdot E \left\{ s_{m-1}[2k-p] \cdot s_{m-1}[2l-q] \right\}$$

Dar valoarea medie din membrul drept reprezintă tocmai autocorelația statistică a semnalului de la intrare, motiv pentru care ultima relație se mai scrie:

$${}_s r_m[k-1] = \sum_{p,q=-\infty}^{\infty} h[p] \cdot h[q] \cdot {}_s r_{m-1} \left[2(k-1) - (p-q) \right] \quad (2.2.2.2)$$

O relație similară poate fi scrisă pentru sistemul din dreapta figurii 2.2.2.1.:

$${}_d r_m[k-1] = \sum_{p,q=-\infty}^{\infty} g[p] \cdot g[q] \cdot {}_s r_{m-1} \left[2(k-1) - (p-q) \right] \quad (2.2.2.3)$$

sau ținând seama de relația anterioară:

$${}_d r_m[k-1] = \sum_{p_1, p_2, q_1, q_2 = -\infty}^{\infty} g[p_1] \cdot g[q_1] \cdot h[p_2] \cdot g[q_2] \cdot {}_s r_{m-2} \left[2^2(k-1) - 2(p_1 - q_1) - (p_2 - q_2) \right] \quad (2.2.2.4)$$

După cum s-a arătat în capitolul anterior calculul TPWD se bazează pe o schemă de forma:

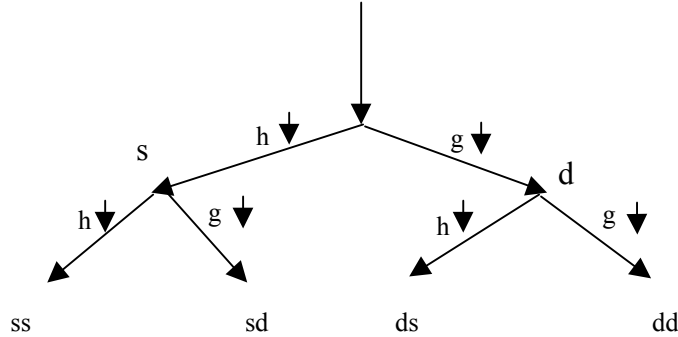


Figura 2.2.2.2. Schema de calcul a TPWD. Primele două iterații.

După modelul relației (2.2.2.4) se poate scrie expresia autocorelației statistice a semnalului obținut după efectuarea a "a" filtrări trece sus și a "b" filtrări trece jos:

$$({}_{dd\dots d})_a ({}_{ss\dots s})_b r[k-1] = \sum_{p_1, q_1, p_2, q_2, \dots, p_a, q_a, p_\alpha, q_\alpha = -\infty}^{\infty} \left\{ \left(\prod_{k=1}^a (g[p_k] \cdot g[q_k]) \cdot h(p_\alpha) \cdot h(q_\alpha) \right) \cdot ({}_{ss\dots s})^{b-1} r \left[2^{a+1}(k-1) - 2^a(p_1 - q_1) - \dots - 2(p_a - q_a) - (p_\alpha - q_\alpha) \right] \right\} \quad (2.2.2.5)$$

S-a obținut dependența autocorelației statistice a semnalului obținut după efectuarea a "a" filtrări trece sus și a "b" filtrări trece jos de autocorelația statistică a semnalului obținut după efectuarea a "b-1" filtrări trece sus. Dar această autocorelație statistică a fost calculată în paragraful anterior, ${}_{s_{b-1}} r_{[k-1]}$.

De aceea se pot utiliza rezultatele analizei asimptotice deja făcute pentru a face analiza asimptotică a TPWD. Pentru TPWD implicată în relația (2.2.2.5) numărul de iterații este: $-m = a + b$. Acest număr tinde la infinit dacă b tinde la infinit. În continuare se calculează limita membrului stâng al relației (2.2.2.5) când b tinde la infinit. Această limită va depinde de limita autocorelației statistice din membrul drept. Dar aceasta poate fi calculată pe baza analizei asimptotice efectuate în paragraful anterior. Se poate scrie:

$$\lim_{b \rightarrow \infty} {}_s r_{b-1} \left[2^{a+1}(k-1) - 2^a(p_1 - q_1) - \dots - 2(p_a - q_a) - (p_\alpha - q_\alpha) \right] = F\{r_{xx}\}(0) \cdot \delta \left[2^{a+1}(k-1) - 2^a(p_1 - q_1) - \dots - 2(p_a - q_a) - (p_\alpha - q_\alpha) \right]$$

sau, ținând seama de staționaritatea semnalului aleator cu funcția de autocorelație statistică ${}_s r_\infty[k-1]$ și de proprietățile impulsului unitar:

$$\lim_{b \rightarrow \infty} (dd\dots d)_a (ss\dots s)_b r[k-1] = \left(\sum_{p_1, p_2, \dots, p_a, p_\alpha = -\infty}^{\infty} \left(\prod_{k=1}^a (g[p_k] \cdot h[p_\alpha]) \right) \right)^2 \cdot F\{r_{xx}\}(0) \cdot \delta[k-1]$$

(2.2.2.6)

Deci semnalul $(dd\dots d)_a (ss\dots s)_b$ obținut în urma a a filtrări trece sus și a b filtrări trece jos devine un zgomot alb în timp discret când b tinde la infinit. Iată de ce se poate afirma că TPWD se comportă asimptotic la fel ca și TUD convergând la transformarea Karhunen-Loeve. Și analiza asimptotică a semnalelor de formele: $(dd\dots d)_b (ss\dots s)_a$, $(ss\dots s)_b (dd\dots d)_a$ sau $(sdsd\dots sd)_b (dsds\dots ds)_a$ conduce la aceleași rezultate.

În continuare se calculează media și dispersia semnalului aleator $(dd\dots d)_a (ss\dots s)_b$.

Media este dată de relația:

$$(dd\dots d)_a (ss\dots s)_b M = E \left\{ \sum_{p_1 = -\infty}^{\infty} g[p_1] \cdot \left\{ (dd\dots d)_{a-1} (ss\dots s)_b \left[2n - p_1 \right] \right\} \right\} =$$

(2.2.2.7)

$$= (dd\dots d)_{a-1} (ss\dots s)_b M \cdot \sum_{p_1 = -\infty}^{\infty} g[p_1] = 0$$

deoarece suma coeficienților răspunsului la impuls al filtrului trece sus g este nulă.

Deci media semnalului aleator considerat este nulă.

În continuare se determină dispersia sa:

$$(dd\dots d)_a (ss\dots s)_b \sigma^2 = (dd\dots d)_a (ss\dots s)_b r[0] =$$

$$= \sum_{p_1, q_1, p_2, q_2, \dots, p_a, q_a, p_\alpha, q_\alpha = -\infty}^{\infty} \left(\prod_{k=1}^a (g[p_k] \cdot g[q_k]) \cdot h(p_\alpha) \cdot h(q_\alpha) \right) \cdot$$

$$\cdot (ss\dots s)_{b-1} r \left[-2^a (p_1 - q_1) - \dots - 2(p_a - q_a) - (p_\alpha - q_\alpha) \right] =$$

$$= \sum_{p_1, p_2, \dots, p_a, p_\alpha = -\infty}^{\infty} \left(\prod_{k=1}^a (g[p_k]) \cdot h(p_\alpha) \right)^2 \cdot \sigma_{b-1}^2 +$$

$$+ \sum_{p_1, q_1, p_1 \neq q_1, p_2, q_2, p_2 \neq q_2, \dots, p_a, q_a, p_a \neq q_a, p_\alpha, q_\alpha, p_\alpha \neq q_\alpha = -\infty}^{\infty} \left(\prod_{k=1}^a (g[p_k] \cdot g[q_k]) \cdot h(p_\alpha) \cdot h(q_\alpha) \right) \cdot$$

$$s_{b-1} \left[-2^a (p_1 - q_1) - \dots - 2(p_a - q_a) - (p_\alpha - q_\alpha) \right]$$

sau:

$$\begin{aligned}
& (dd\dots d)_a (ss\dots s)_b \sigma^2 = {}_s \sigma_{b-1}^2 \left(\sum_{p_1, p_2, \dots, p_a, p_\alpha = -\infty}^{\infty} \left(\prod_{k=1}^a (g[p_k]) \cdot h(p_\alpha) \right) \right)^2 + \\
& + \sum_{p_1, q_1, p_1 \neq q_1, p_2, q_2, p_2 \neq q_2, \dots, p_a, q_a, p_a \neq q_a, p_\alpha, q_\alpha, p_\alpha \neq q_\alpha = -\infty}^{\infty} \left(\prod_{k=1}^a (g[p_k]) \cdot g[q_k] \cdot h(p_\alpha) \cdot h(q_\alpha) \right) \cdot \\
& {}_s \rho_{b-1} \left[-2^a (p_1 - q_1) - \dots - 2(p_a - q_a) - (p_\alpha - q_\alpha) \right]
\end{aligned} \tag{2.2.2.8}$$

unde s-a notat cu ρ coeficientul de autocorelație:

$${}_s \rho_{b-1}(x) = \frac{{}_s r_{b-1}(x)}{{}_s \sigma_{b-1}^2}$$

Formulele (2.2.2.7) și (2.2.2.8) sunt foarte importante deoarece pot fi utilizate, așa după cum s-a arătat în [1], pentru cuantizarea optimală (neuniformă) a semnalului care a suferit transformarea wavelet. Avantajul lor asupra formulelor stabilite în paragraful anterior (pentru cazul particular al TUD) rezidă în forma lor recurentă. Asimptotic:

$$(dd\dots d)_a (ss\dots s)_\infty M = 0$$

și:

$$(dd\dots d)_a (ss\dots s)_\infty \sigma^2 = \sum_{p_1, p_2, \dots, p_a, p_\alpha = -\infty}^{\infty} \left(\prod_{k=1}^a (g[p_k]) \cdot h(p_\alpha) \right)^2 \cdot F\{r_{xx}\}(0)$$

formulă care dă dispersia zgomotului alb la care converge șirul dispersiilor coeficienților TPWD atunci când numărul de iterații al acestuia tinde la infinit. În consecință dacă numărul de iterații al TPWD este suficient de mare atunci la fiecare nouă iterație a acestuia se obțin niște coeficienți care reprezintă eșantioanele unui zgomot alb de medie nulă și dispersie dată de ultima relație. În consecință cuantizarea optimală a acestor coeficienți poate fi realizată folosind această valoare pentru dispersia lor.

În figura 2.2.2.3. se ilustrează proprietatea de albire a TPWD. În partea de sus este prezentată densitatea spectrală de putere a semnalului de intrare. Acesta este un zgomot colorat, obținut prin filtrarea trece bandă a unui zgomot alb. În partea de jos se prezintă densitatea spectrală de putere a semnalului obținut în urma calculului TPWD a semnalului a cărui densitate spectrală de putere este prezentată în partea de sus.

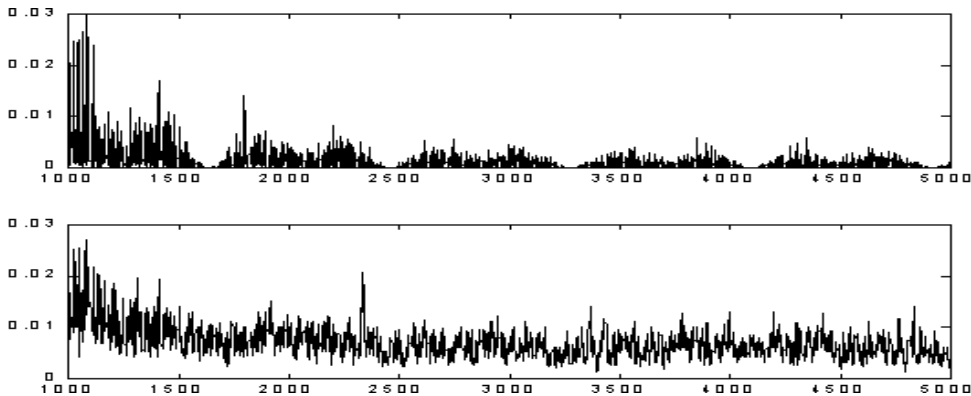


Figura 2.2.2.3. Efectul de albire al TPWD.

Se constată că densitatea spectrală din partea de jos are o anvelopă mult mai asemănătoare cu o constantă, așa cum arată densitatea spectrală de putere a unui zgomot alb.

2.2.3. Analiza statistică a TPC

Transformarea cu pachete cosinusoidale, TPC, este o combinație între transformarea cosinus discretă (TCD) și teoria pachetelor de funcții "wavelet". În cazul acestei transformări se realizează o segmentare a suportului semnalului de analizat (considerat de durată N) în blocuri de lungime dependentă de numărul de ordine al iterației. La a m-a iterație aceste blocuri au lungimea $2^{-m} N$. Semnalul conținut în fiecare dintre aceste blocuri este transformat folosind DCT. Procedura de segmentare și de calcul a DCT este ilustrată în figura 2.2.3.1.

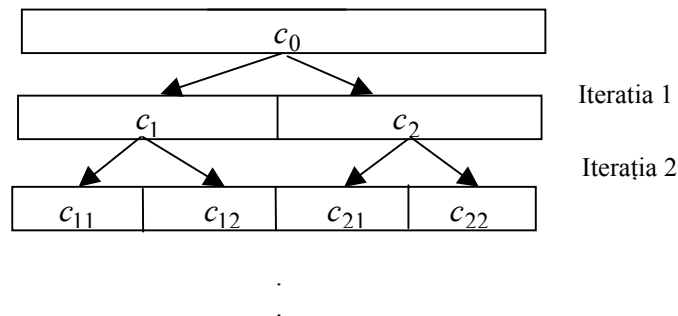


Figura 2.2.3.1. Procedura de calcul a TPC.

Rezultatul TPC constă din alegerea unui număr total de N coeficienți, din unul sau mai multe blocuri de coeficienți specifice uneia sau mai multor iterații. Această alegere se face prin "căutarea celei mai bune baze". De exemplu un rezultat posibil pentru TPC, prezentată în figura 2.2.3.1. este secvența de coeficienți $\{c_1, c_{21}, c_{22}\}$. Un alt rezultat posibil este secvența $\{c_0\}$. Analiza asimptotică a TPC se bazează pe faptul că atunci când N tinde la infinit DCT converge la transformarea Karhunen-Loeve, [46]. Dar dacă N tinde la infinit atunci numărul de eșantioane al fiecărei secvențe de coeficienți DCT corespunzătoare unui anumit segment și unei anumite iterații tinde la infinit și deci coeficienții DCT ai secvenței corespunzătoare converg la transformarea Karhunen-Loeve. Deoarece fiecare secvență de coeficienți din cadrul rezultatului aplicării TPC converge asimptotic la transformarea Karhunen-Loeve se poate afirma că TPC converge asimptotic la transformarea Karhunen-Loeve. Efectul de albire al TPC este ilustrat în figura 2.2.3.2. În partea de sus a acestei figuri este prezentată densitatea spectrală de putere a unui zgomot colorat, obținut prin filtrarea cu un mediator numeric alunecător a unui zgomot alb. În partea de jos a figurii este prezentată densitatea spectrală de putere a semnalului obținut în urma calculului TPC a semnalului cu densitatea spectrală de putere din partea de sus a figurii. Se constată că anvelopa densității spectrale de putere din partea de jos a figurii aproximează bine o constantă. De aceea se poate afirma că semnalul a cărui densitate spectrală de putere este reprezentată în partea de jos a figurii este o bună aproximare pentru un zgomot alb. În consecință a fost pusă în evidență proprietatea de albire a TPC. Făcând o analiză comparativă a figurilor 2.2.1.1., 2.2.2.3 și 2.2.3.2. se constată că, din punct de vedere al vitezei de convergență a transformării "wavelet" spre un zgomot alb, cel mai bine se comportă TPC, urmată de TUD. Cea mai lentă convergență o are TPWD.

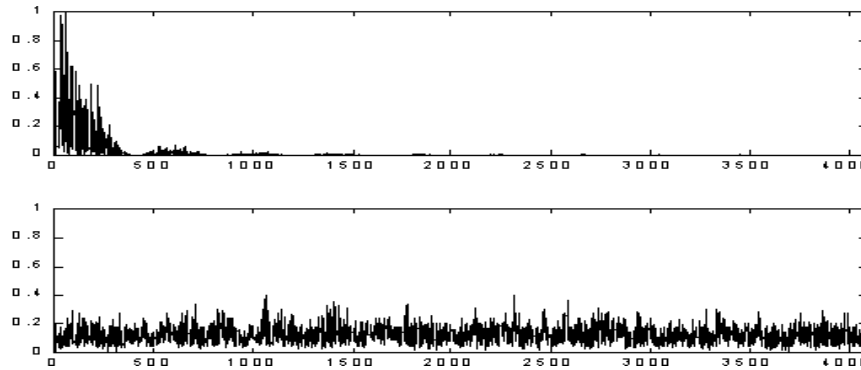


Figura 2.2.3.2. Efectul de albire al TPC.

Deci oricare dintre cele trei transformări bazate pe teoria funcțiilor "wavelet", TUD, TPWD sau TPC, poate fi utilizată drept transformare ortogonală, într-o schemă de compresie de date, deoarece oricare dintre acestea converge asimptotic la transformarea Karhunen-Loeve. Ținând seama de criteriul vitezei de convergență, cea mai potrivită pare a fi TPC. Această transformare are și avantajul localizării temporale a filtrelor de analiză, în intervalele I_m . În plus ea este și cea mai adecvată pentru prelucrarea semnalelor de vorbire, ținând seama de modelul sinusoidal al acestora. Această afirmație se justifică în continuare.

2.3. Alegerea celei mai bune funcții de tipul "wavelets mother"

Unul dintre parametrii unei transformări "wavelet" ortogonală discretă este funcția "wavelets mother" folosită. Celălalt parametru este numărul de iterații. După cum s-a văzut în cazul TUD și TPWD acest număr este bine să fie cât mai mare posibil (pentru a se atinge regimul asimptotic). În cazul TPC numărul de iterații poate fi ales și pe baza altor considerente.

În cazul aplicațiilor de compresie alegerea funcției "wavelets mother" trebuie făcută în funcție de natura semnalului de prelucrat, în așa fel încât să se maximizeze factorul de compresie la un nivel de distorsiuni de reconstrucție impus. Maximizarea factorului de compresie se realizează prin minimizarea numărului de coeficienți de valoare superioară unui anumit prag. Valoarea pragului depinde de nivelul de distorsiuni acceptat la reconstrucție. Cu cât numărul de coeficienți superiori pragului este mai mic cu atât numărul de biți necesar pentru descrierea variantei compresate a semnalului este mai mic și deci factorul de compresie obținut este mai mare. Pentru alegeri diferite ale funcției "wavelets mother" se obțin numere diferite de coeficienți superiori pragului. Această afirmație poate fi ilustrată folosind metoda de compresie descrisă în paragraful 2.2.1. În figura 2.3.1. se prezintă semnalul care urmează să fie compresat. În figura 2.3.2 se prezintă rezultatul compresiei urmată de reconstrucția semnalului din figura 2.3.1. atunci când pentru calculul TUD s-a utilizat funcția "wavelets mother" de tip Dau 4 (a se vedea [12]). În figura 2.3.3 se prezintă rezultatul reconstrucției în urma compresiei pentru cazul utilizării funcției "wavelets mother" de tip Dau 20. Comparând figurile 2.3.2. și 2.3.3. se constată că deși în ambele experimente puterea distorsiunii nu depășește 1% din puterea semnalului din figura 2.3.1. totuși experimentul descris în figura 2.3.2. este superior din punct de vedere al compresiei, deoarece în cazul său au fost utilizați doar 7 coeficienți ai transformării "wavelet" în timp ce în cazul experimentului ilustrat în figura 2.3.3. au fost utilizați 26 de coeficienți ai TUD. În consecință ar fi necesară o procedură de selecție a funcției "wavelets mother" în acord cu forma semnalului de prelucrat în scopul maximizării factorului de compresie în cazul în care se utilizează TUD. La aceeași concluzie se ajunge și atunci când se utilizează TPWD.

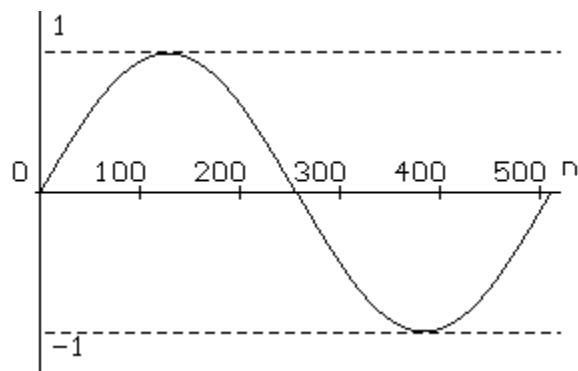


Figura 2.3.1. Semnalul de prelucrat.

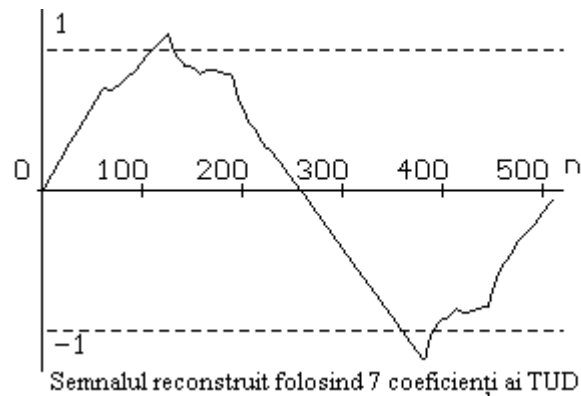


Figura 2.3.2. Rezultatul operațiilor de compresie și reconstrucție când pentru calculul TUD s-a utilizat funcția "wavelets mother" Dau 4.

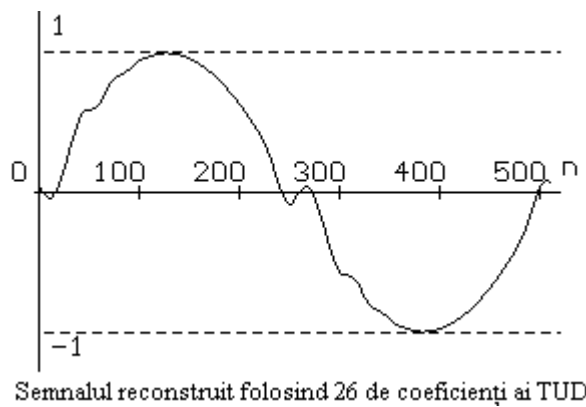


Figura 2.3.3. Rezultatul operațiilor de compresie și reconstrucție când pentru calculul TUD s-a utilizat funcția "wavelets mother" Dau 20.

O astfel de procedură de selecție este descrisă în [43]. Ea se bazează pe segmentarea semnalului de prelucrat. Acesta este aproximat polinomial, pe segmente, prin dezvoltare în serie Taylor. Gradul fiecărui polinom este fixat în așa fel încât eroarea de aproximare pe segmentul corespunzător să nu depășească o valoare impusă. Pentru prelucrarea fiecărui segment se alege o funcție "wavelets mother" cu un număr de momente nule egal cu gradul polinomului de aproximare de pe segmentul respectiv. Primul segment considerat are o lungime egală cu durata semnalului. Dacă nu există nici un polinom aproximant pentru acest segment (adică dacă întreg semnalul nu poate fi aproximat suficient de bine cu un polinom de grad mai mic sau egal cu numărul maxim de momente nule al unei funcții de tip "wavelets mother" disponibilă) atunci segmentul se înjumătățește și se încearcă aproximarea polinomială pe fiecare jumătate. Această procedură se repetă până când se realizează segmentarea întregului semnal. În funcție de gradele polinoamelor aproximante se aleg, pentru fiecare segment, funcțiile "wavelets mother" folosite pentru calculul TUD pe acel segment. În acest fel pe fiecare segment se folosește o altă bancă de filtre pentru calculul TUD. S-a obținut deci o bancă de filtre variabilă în timp, care se adaptează la semnalul de prelucrat, făcând o anumită localizare temporală a parametrilor acestuia. În acest mod poate fi ameliorată și deficiența cronică a transformării cu pachete de funcții "wavelet".

În cazul TPC nu există diferențe majore între diferitele funcții "wavelets mother" care pot fi utilizate. Singurele diferențe provin din deosebirile dintre diferitele ferestre $w(t)$ care pot fi utilizate.

În continuare se prezintă o modalitate de selecție a celei mai bune funcții "wavelets mother" pentru cazul compresiei semnalului de vorbire.

Fiecare propoziție rostită este o secvență de tonuri care au diferite intensități, frecvențe și durate. Fiecare ton este un semnal sinusoidal cu amplitudine, frecvență și durată specifice. Acesta este modelul sinusoidal al vorbirii. O descriere matematică pentru acest model este:

$$x(t) = \sum_{q=1}^{Q(t)} A_q \cos \theta_q(t) \quad (2.3.1)$$

[47], unde componentele sunt numite parțiale. Fiecare termen al acestei sume este un semnal cu dublă modulație. Deci nu este vorba despre semnale staționare. Dar vorbirea este privită frecvent ca și o succesiune de semnale staționare. Împărțind semnalul de vorbire într-o succesiune de segmente, fiecare având o durată mai mică de 25 ms, se obține o secvență de semnale staționare. Pe fiecare segment modelul vorbirii poate fi de forma:

$$x_s(t) = \sum_{q=1}^Q A_q \cos \omega_q t \quad (2.3.2)$$

Această descompunere seamănă mult cu descompunerea semnalului $x_s(t)$ într-un pachet de cosinusuri.

Descompunerea aceluiași semnal, folosind o bază de funcții "wavelet" este de forma:

$$x_s(t) = \sum_{k=1}^K \sum_{l=1}^L \langle x_s(t), \psi_{k,l}(t) \rangle \psi_{k,l}(t) \quad (2.3.3)$$

unde $\psi_{k,l}(t)$ sunt undișoarele generate de funcția "wavelets mother" $\psi(t)$. Factorul de compresie obținut folosind o funcție "wavelets mother" specificată este mai mare dacă numărul de coeficienți:

$$d_{k,l} = \langle x_s(t), \psi_{k,l}(t) \rangle \quad (2.3.4)$$

nenuli, al acestei descompunerii, N_ψ , este mai mic. Dar:

$$d_{k,l} = \int_{-\infty}^{\infty} x_s(t) \cdot \psi_{k,l}^*(t) dt = r_{x_s, \psi_{k,l}}(0) \quad (2.3.5)$$

unde membrul drept reprezintă valoarea intercorelației semnalelor $x_s(t)$ și $\psi_{k,l}(t)$ calculată în origine. Aceasta este valoarea maximă a acestei funcții. Intercorelația măsoară gradul de asemănare al celor două semnale. Deci mărimea coeficientului $d_{k,l}$ este mai mare dacă semnalele $x_s(t)$ și $\psi_{k,l}(t)$ sunt mai asemănătoare. Folosind relația (2.3.2) se poate afirma că funcțiile "wavelets" cele mai asemănătoare cu semnalul $x_s(t)$ sunt elementele unui pachet de cosinusuri. Dar dacă mulțimea $\{\psi_{k,l}(t)\}_{k \in Z, l \in Z}$ este o bază ortonormată atunci energia semnalului $x_s(t)$ se poate calcula cu formula:

$$E_x = \sum_{k=1}^K \sum_{l=1}^L |d_{k,l}|^2 \quad (2.3.6)$$

Deoarece energia semnalului $x_s(t)$ este o constantă independentă de funcția "wavelets mother" selecționată, se poate afirma că numărul N_ψ este mai mic dacă mărimea coeficienților $d_{k,l}$ nenuli este mai mare. **Îată de ce pentru compresia semnalului de vorbire cea mai bună transformare "wavelet" ortogonală este TPC.** Bineînțeles această afirmație este valabilă în măsura în care modelul sinusoidal al vorbirii este respectat. În continuare se analizează modalitățile de optimizare a compresiei vorbirii bazată pe TPC prin alegerea celei mai bune funcționale de cost folosită în algoritmul de alegere al celei mai bune baze. TPC este o transformare adaptivă. Rezultatul utilizării sale într-o anumită aplicație poate fi optimizat folosind procedura de căutare a celei mai bune baze. Aceasta este o procedură foarte eficientă, care poate crește mult calitatea unei anumite metode de prelucrare a semnalelor. Așa după cum s-a arătat deja există mai multe funcționale de cost a căror minimizare poate conduce la găsirea celei mai bune baze. Cea mai utilizată dintre acestea este entropia coeficienților $d_{k,l}$. Dar minimizarea acestei funcționale nu conduce la maximizarea factorului de compresie. **Funcționala de cost, a cărei minimizare conduce la maximizarea factorului de compresie este cea care conduce la minimizarea numărului de coeficienți, N_s , superiori unei anumite valori de prag, t , (care fixează puterea distorsiunii care apare la reconstrucție).** Într-adevăr, folosind această funcțională de cost, pentru alegerea celei mai bune baze (cel mai bun pachet

de cosinusuri) se obține un anumit număr, N_{ψ} , de coeficienți ai TPC nenuli (dintre care un număr de N_s sunt superiori pragului t). La ieșirea detectorului de prag, DP, din schema bloc de compresie, figura 2.1.1, se obțin N_s coeficienți nenuli. Dar acesta este un număr minim, deoarece a fost minimizat prin procedura de alegere a celei mai bune baze. Iată de ce utilizarea acestei funcționale de cost conduce la maximizarea factorului de compresie.

Mărind valoarea pragului t , numărul N_s devine mai mic și valoarea factorului de compresie devine mai mare. Din nefericire valoarea raportului semnal pe zgomot la ieșirea sistemului de reconstrucție (a semnalului compresat), rsz_0 scade cu creșterea lui t . Deci creșterea pragului t trebuie controlată pentru a păstra "transparența" (capacitatea de a păstra imperceptibilă simțurilor umane) compresiei. Acesta este motivul pentru care detectorul de prag DP, trebuie să fie un sistem adaptiv. Un alt parametru al TPC care poate fi considerat pentru optimizarea compresiei este numărul de iterații.

Deci transformarea "wavelet" recomandată pentru compresia vorbirii este TPC (în măsura în care modelul sinusoidal al vorbirii este valabil). Alegerea unui anumit număr de iterații al acesteia poate fi utilizată pentru maximizarea factorului de compresie.

Capitolul 3. Detectorul de prag

Unul dintre cele mai importante blocuri din structura sistemului de compresie din figura 2.1.1. este detectorul de prag. Rolul acestui sistem este de a elimina toți coeficienții din domeniul transformării "wavelet" discretă mai mici decât o anumită valoare. Acesta este de fapt mecanismul prin care se realizează compresia. Și acest bloc are în schema propusă o structură adaptivă.

3.1. Detecția adaptivă de prag

Analizând sistemul din figura 2.1.1 se poate constata că distorsiunea datorată compresiei are valoarea medie pătratică:

$$D = E \left\{ \left(x[n] - \hat{x}[n] \right)^2 \right\} \quad (3.1.1)$$

Pentru că TPC și inversa sa TPCI sunt transformări ortogonale ultima relație devine:

$$D = E \left\{ \left(y[n] - u[n] \right)^2 \right\} \quad (3.1.2)$$

Valoarea pragului, t , trebuie aleasă astfel încât să fie satisfăcută condiția:

$$D < \alpha \cdot E_x \quad \alpha < 1 \quad (3.1.3)$$

unde E_x reprezintă energia semnalului de intrare $x[n]$. Se demonstrează următoarea propoziție:

Propoziția 3.1.1.

O margine superioară a distorsiunii semnalului reconstruit, obținut după compresia adaptivă bazată pe TPC este $N \cdot t^2$ unde N reprezintă numărul de eșantioane al semnalului de prelucrat, iar t valoarea pragului.

Demonstrație.

Eroarea medie pătratică de aproximare a semnalului $y[n]$ prin semnalul $z[n]$ este:

$$\varepsilon_1 = E \left\{ \left(y[n] - z[n] \right)^2 \right\} = \sum_{k=1}^K y^2[n_k] \quad (3.1.4)$$

unde n_k reprezintă pozițiile eșantioanelor semnalului $y[n]$ care au modulul inferior valorii de prag, t . Fie K astfel de eșantioane. Fie $o[n]$ semnalul obținut după ordonarea în ordine crescătoare a eșantioanelor semnalului $y[n]$. Eroarea medie pătratică devine:

$$\varepsilon_1 = \sum_{k=1}^K o^2[k] \leq K \cdot t^2 \quad (3.1.5)$$

Se consideră că blocul Cu din figura 2.1.1 realizează o cuantizare uniformă, cu cuanta de valoare t . Dacă se realizează o cuantizare neuniformă se obțin rezultate mai bune. Eroarea medie pătratică este:

$$\varepsilon_2 = \sum_{n=1}^N \left(z[n] - u[n] \right)^2 \quad (3.1.6)$$

Pentru fiecare eșantion al secvenței $o[k]$, $k = \overline{1, K}$, se introduce un eșantion nul în secvența $u[n]$. Pentru celelalte eșantioane ale semnalului $z[n]$, diferența $z[n] - u[n]$ este inferioară valorii t . De aceea se poate scrie:

$$\varepsilon_2 \leq \sum_{k=K+1}^N t^2 = (N - K) \cdot t^2 \quad (3.1.7)$$

Pentru că distorsiunea, definită în relația (3.1.2) poate fi scrisă în forma:

$$D = \varepsilon_1 + \varepsilon_2 \quad (3.1.8)$$

ținând seama de relațiile (3.1.5) și (3.1.7) se poate afirma că propoziția a fost demonstrată.

Deci pentru a păstra distorsiunea sub valoarea $\alpha \cdot E_x$ este suficient să se aleagă valoarea de prag:

$$t = \sqrt{\frac{\alpha \cdot E_x}{N}} \quad (3.1.9)$$

Constanta α poate fi exprimată cu ajutorul raportului semnal pe zgomot al semnalelor $u[n]$ și $\hat{x}[n]$ (care au aceeași energie), rsz_0 . Se poate scrie:

$$rsz_0 = 10 \cdot \log_{10} \frac{E_x}{D} \geq -10 \cdot \log_{10} \alpha \quad (3.1.10)$$

În acest mod a fost stabilită o margine inferioară a raportului semnal pe zgomot, care depinde de α :

$$\beta = -10 \cdot \log_{10} \alpha \quad (3.1.11)$$

Luând semnul egal în penultima relație se poate obține o margine inferioară pentru constanta α :

$$\alpha_m = 10^{-\frac{\beta}{10}} \quad (3.1.12)$$

Folosind această valoare și relația (3.1.10) se poate obține o margine inferioară pentru pragul t :

$$t_m = \sqrt{10^{-\frac{\beta}{10}} \cdot \frac{E_x}{N}} \quad (3.1.13)$$

Deci alegând pentru prag o valoare t mai mare decât t_m se obține o valoare a raportului semnal pe zgomot la ieșire superioară lui β . Din nefericire valoarea exactă a lui rsz_0 nu va fi cunoscută. De aceea se recomandă un algoritm adaptiv pentru alegerea pragului. Acest algoritm poate utiliza valoarea t_m pentru inițializare. **Valoarea de prag este crescută pornind de la această valoare. La fiecare iterație se calculează rsz_0 . Dacă această valoare este superioară lui β procesul de creștere a valorii de prag este continuat. Algoritmul ia sfârșit când pentru prima dată valoarea rsz_0 devine mai mică decât β .**

Detectorul de prag este un sistem neliniar care este descris prin următoarea relație de legătură intrare-ieșire:

$$z[n] = \begin{cases} y[n], & |y[n]| > t, \\ 0, & |y[n]| \leq t \end{cases} \quad (3.1.14)$$

În continuare se prezintă o analiză statistică pentru acest sistem.

3.2. Analiza statistică a detectorului de prag

Fie X variabila aleatoare de la intrarea în detectorul de prag și Y variabila aleatoare obținută la ieșire. Se va considera că X este o variabilă aleatoare Gaussiană de medie nulă. Se studiază robustețea acestui sistem. În acest scop se calculează dispersia variabilei aleatoare de la ieșire. Dacă aceasta este mai mică decât dispersia variabilei aleatoare de la intrare atunci se poate declara că sistemul este robust, [49]. În consecință întreg sistemul de compresie din figura 2.1.1. ar fi robust. Relația intrare-ieșire din relația (3.1.14) are reprezentarea grafică din figura 3.2.1.

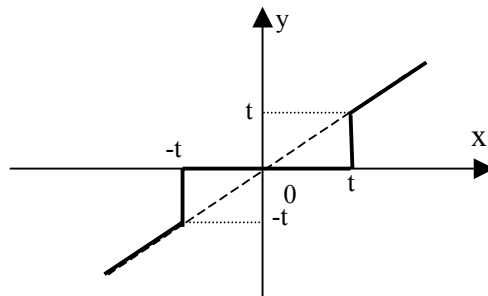


Figura 3.2.1. Relația intrare-ieșire pentru detectorul de prag.

Relația de legătură dintre funcțiile de repartiție a variabilelor aleatoare X și Y este:

$$F_Y(y) = \begin{cases} F_X(y), & y \in (-\infty, -t) \\ F_X(-t), & y \in [-t, 0) \\ F_X(t), & y \in [0, t) \\ F_X(y), & y \in [t, \infty) \end{cases} \quad (3.1.15)$$

Derivând această funcție se obține legătura dintre densitățile de probabilitate ale semnalelor de intrare și ieșire, prezentată în figura 3.2.2.

Pentru media variabilei aleatoare Y se obține:

$$m_Y = 0 \quad (3.1.16)$$

Calculând dispersia lui Y se obține:

$$\sigma_Y^2 = \sigma_X^2 - 2 \int_0^t y^2 p_X(y) dy \quad (3.1.17)$$

unde σ_X^2 reprezintă dispersia lui X. În consecință:

$$\sigma_Y^2 \leq \sigma_X^2 \quad (3.1.18)$$

Și deci detectorul de prag este un sistem robust. De aceea și metoda de compresie propusă în această teză este robustă. Acesta este unul dintre avantajele compresiei bazată pe funcții "wavelet" în comparație cu alte metode de compresie.

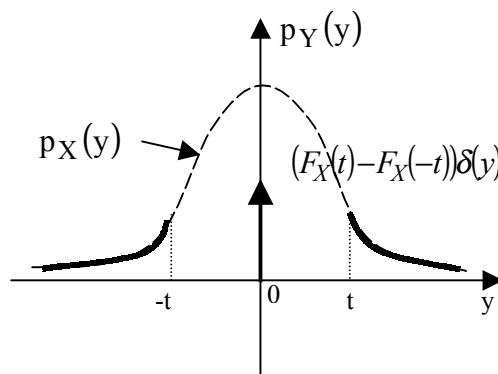


Figura 3.2.2. Densitatea de probabilitate a variabilei aleatoare Y.

Capitolul 4. Sistemul de cuantizare

Un alt bloc foarte important din structura unui sistem de compresie este sistemul de cuantizare. Despre operația de cuantizare s-a scris foarte mult. În [1] s-a făcut o trecere în revistă a principalelor tehnici de cuantizare cunoscute, folosindu-se intensiv lucrările [50]..[54]. Un rezultat remarcabil, obținut în [1], îl constituie generalizarea teoremei de cuantizare a lui Widrow. Acest rezultat a fost publicat de către autorul acestei teze în [55]. Sistemul de cuantizare trebuie proiectat în acord cu aplicația pe care o deservește. Pentru compresia vorbirii este important ca structura sistemului de cuantizare să țină seama de particularitățile acestui tip de semnal. Acestea sunt evidențiate de modelul psiho-acustic al vorbirii. În continuare se face o prezentare succintă a acestui model.

4.1. Proprietățile psiho-acustice ale semnalului de vorbire

Cea mai importantă proprietate psiho-acustică a semnalului de vorbire, din punct de vedere al compresiei, este dată de fenomenul de mascare.

4.1.1. Fenomenul de mascare

Există două tipuri de mascare ale unui sunet de către un altul. Se vorbește despre mascarea frecvențială dacă cele două sunete apar simultan și de mascarea temporală dacă cele două sunete apar succesiv. În continuare se vor face referiri doar la mascarea frecvențială deoarece e dificil să se țină seama în structura unui sistem de compresie de mascarea temporală.

Fie o sinusoidă de frecvență f_1 și de amplitudine A_1 . Urechea nu va sesiza semnalul sonor descris de această sinusoidă, într-o ambianță de liniște perfectă, decât dacă puterea acestei sinusoidă este mai mare decât pragul de audiție absolut, $S_a(f)$. Aria de audiție, este aria unei suprafețe care se întinde deasupra pragului de audiție între 20 Hz și 20 kHz respectiv între 0 și 90 dB.

În continuare se consideră cazul în care s-ar fi emis 2 sinusoidă, prima, cea care maschează având frecvența f_1 și puterea P_1 iar cea de a doua, cea care este mascată, având frecvența f_2 și puterea P_2 . Se măsoară pentru toate valorile posibile ale frecvenței f_2 din banda audio, puterea P_2 pentru care cea de a doua sinusoidă este la limita audibilității. Funcția $P_2(f_2)$ se numește curbă de mascare. Curbele de mascare ale unei sinusoidă de către o altă sinusoidă nu sunt singurele curbe psiho-acustice interesante. **În cadrul operației de cuantizare se va accepta un zgomot de cuantizare într-o anumită bandă de frecvență dacă acesta nu este audibil.** E deci interesant să se examineze cazul în care semnalul mascat este un zgomot de bandă îngustă. Mai mult, deoarece un semnal de vorbire poate fi considerat ca fiind compus dintr-un anumit număr de semnale pure (componente tonale mai numite și parțiale), care pot fi modelate prin sinusoidă și dintr-un anumit număr de semnale care nu sunt pure (componente netonale), care pot fi modelate prin zgomote de bandă îngustă, este necesar să se analizeze următoarele patru cazuri: mascarea unei sinusoidă de către o sinusoidă, mascarea unei sinusoidă de către un zgomot de bandă îngustă, mascarea unui zgomot de bandă îngustă de către o sinusoidă și mascarea unui zgomot de bandă îngustă de către un zgomot de bandă îngustă. Un zgomot de bandă îngustă este caracterizat de trei parametri: frecvența centrală, lărgimea de bandă și puterea. Toate aceste curbe au aceeași alură, o formă triunghiulară. Ele depind de parametrii f_1 și P_1 . Toate curbele de mascare au maximul la frecvența f_1 . Puterea P_2 la frecvența f_1

este puțin mai mică decât puterea P_1 . Diferența $P_2(f_2 = f_1) - P_1$ se numește indice de mascare. Pantele curbelor de mascare sunt mai mari spre frecvențe joase decât spre frecvențe înalte. Aceste pante depind de frecvența f_1 a semnalului mascant. Ele sunt mai mici când această frecvență este mai mare.

Dacă frecvența se măsoară folosind o nouă unitate de măsurare, numită Bark și dacă puterile se măsoară în dB se poate demonstra că aceste curbe de mascare pot fi modelate cu segmente de dreaptă (într-o anumită bandă de frecvență în jurul lui f_1) și că pantele acestor drepte nu mai depind de f_1 . Relația între o frecvență exprimată în Hz, în intervalul [20, 20000] și o frecvență exprimată în Barks, în intervalul [1, 24] este:

$$f_{\text{Bark}} = 13 \cdot \arctg\left(0,76 \cdot \frac{f_{\text{Hertz}}}{1000}\right) + 3,5 \cdot \arctg\left[\left(\frac{f_{\text{Hertz}}}{7500}\right)^2\right] \quad (4.1)$$

Curbele de mascare rămân dependente de P_1 . Panta spre frecvențele înalte este cu atât mai mică cu cât P_1 este mai mare. Funcția de mascare definită de modelul psiho-acustic numărul 1 al standardului MPEG Audio ține seama de influența puterii P_1 a semnalului mascant. Funcțiile de mascare sunt definite de relațiile următoare:

$$v(f_1 - f_2, P_1) = \begin{cases} 17(f_1 - f_2 + 1) - (0,4 \cdot P_1 + 6), & -3 \leq f_1 - f_2 < -1 \\ (0,4 \cdot P_1 + 6)(f_1 - f_2), & -1 \leq f_1 - f_2 < 0 \\ -17 \cdot (f_1 - f_2), & 0 \leq f_1 - f_2 < 1 \\ -(f_1 - f_2 - 1)(17 - 0,15 \cdot P_1) - 17, & 1 \leq f_1 - f_2 < 8 \end{cases} \quad (4.2)$$

unde frecvențele sunt măsurate în Barks iar P_1 și v în dB.

Curba de mascare, exprimată în dB are expresia:

$$P_2(f_2, f_1, P_1) = P_1(f_1) + a(f_1) + v(f_1 - f_2, P_1(f_1)) \quad (4.3)$$

unde indicele de mascare $a(f_1)$ este dat de:

$$a_t(f_1) = -1,525 - 0,275 \cdot f_1 - 4,5 \quad (4.4)$$

dacă sunetul mascant este tonal, sau de:

$$a_n(f_1) = -1,525 - 0,175 \cdot f_1 - 0,5 \quad (4.5)$$

dacă sunetul mascant este netonal (zgomot de bandă îngustă).

Pe baza curbelor de mascare pot fi eliminate din spectrul semnalului de vorbire componentele mascate. În acest mod se realizează o compresie transparentă a semnalului de vorbire.

4.1.2. Benzi critice

Cea de a doua caracteristică a modelului psiho-acustic se bazează pe ipoteza că urechea umană se comportă ca un banc de filtre realizând o partiție neregulată a axei frecvențelor. Banda audibilă este împărțită în 24 de benzi numite BENZI CRITICE.

Fie o sinusoidă de frecvență f_1 . Dacă puterea acesteia, P_1 , verifică:

$$P_1 \geq S_a(f_1) \quad (4.6)$$

această sinusoidă este audibilă. Fie o a doua sinusoidă de frecvență vecină și tot de putere P_1 . Cele două sinusoide sunt audibile dacă:

$$2 \cdot P_1 \geq S_a(f_1) \quad (4.7)$$

Fie N sinusoide de frecvențe $f_1, f_1 + df, \dots, f_1 + (N-1)df$ și de putere P_1 . Mulțimea acestor sinusoide este audibilă dacă:

$$N \cdot P_1 \geq S_a(f_1) \quad (4.8)$$

Această condiție este satisfăcută dacă lărgimea de bandă $\Delta f = N \cdot df$ este mai mică decât un prag, numit lărgime de bandă critică, în vecinătatea lui f_1 . Iată cum pot fi definite benzile critice, subiectul acestui paragraf. Scara frecvențială, măsurată în Barks, corespunde numărului de ordine al fiecărei benzi critice. În tabelul următor se prezintă o alegere posibilă a benzilor critice.

Nr. ord. (1)	Frecv. infer. (2)	Frecv. Sup. (3)	Lărg. bandă (4)	(1)	(2)	(3)	(4)
1	20	100	80	13	1720	2000	280
2	100	200	100	14	2000	2320	320
3	200	300	100	15	2320	2700	380
4	300	400	100	16	2700	3150	450
5	400	510	110	17	3150	3700	550
6	510	630	120	18	3700	4400	700
7	630	770	140	19	4400	5300	900
8	770	920	150	20	5300	6400	1100
9	920	1080	160	21	6400	7700	1300
10	1080	1270	190	22	7700	9500	1800
11	1270	1480	210	23	9500	12000	2500
12	1480	1720	240	24	12000	15500	3500

Tabelul 1. Benzi critice.

În acest tabel benzile critice sunt alese în mod artificial, urechea putând crea o bandă critică în jurul oricărei frecvențe. Dacă se definește o nouă scară frecvențială asociind la fiecare frecvență centrală din acest tabel numărul de ordine al benzii critice corespunzătoare, se regăsește scara frecvențelor măsurate în Barks definită anterior.

4.1.2.1. Pragul de mascare

Cu ajutorul modelului precedent se determină în continuare curba de mascare globală și apoi pragul de mascare provenit din mulțimea curbelor calculate pornind de la toate componentele spectrale ale unui semnal de vorbire. **Orice zgomot de cuantizare a cărui putere este inferioară pragului de mascare va fi inaudibil. Cuantizarea semnalului de vorbire cu această proprietate va fi transparentă.** Principiul este simplu dar implementarea sa este dificilă.

Prima operație constă în estimarea densității spectrale de energie a semnalului de vorbire $x[n]$, $P_x(\Omega)$. Aceasta poate fi estimată cu ajutorul reprezentărilor timp-frecvență sau folosind metoda bazată pe spectrul modelului auto-regresiv asociat. Prin eșantionarea în frecvență a acestei funcții se obține funcția $P_x[k]$. Această funcție se normalizează astfel încât:

$$\max_k \{P_x[k]\} = 96\text{dB} \quad (4.9)$$

Pentru a calcula pragul de mascare trebuie adăugate noi ipoteze modelului psiho-acustic al audiției vorbirii. Se va presupune că fiecare componentă a densității spectrale de putere $P_x[k]$ contribuie independent la pragul de mascare și că este suficient să se adune aceste contribuții. Această ipoteză este rațională în interiorul unei benzi critice dar nimic nu justifică generalizarea sa la întreg spectrul semnalului de vorbire.

Deoarece modelul de mascare este diferit dacă semnalul mascant este tonal sau netonal este necesar să se realizeze această distincție pentru toate componentele $P_x[k]$. Conform modelului psiho-acustic numărul 1 din standardul MPEG Audio o componentă spectrală este considerată tonală dacă verifică următoarele trei condiții:

$$\begin{aligned} P_x[k] &> P_x[k-1] \\ P_x[k] &\geq P_x[k+1] \\ P_x[k] - P_x[k+j] &\geq 7\text{dB} \end{aligned} \quad (4.10)$$

cu j din mulțimile:

$$\begin{aligned} j \in \{-2, 2\} & \quad \text{pentru} \quad 2 < k < 63 \\ j \in \{-3, -2, 2, 3\} & \quad \text{pentru} \quad 63 \leq k < 127 \\ j \in \{-6, \dots, -2, 2, \dots, 6\} & \quad \text{pentru} \quad 127 \leq k \leq 250 \end{aligned} \quad (4.11)$$

Domeniul de variație din ce în ce mai mare pentru j este datorat faptului că rezoluția frecvențială a urechii este mai bună la frecvențe joase.

La puterea fiecărei componente $P_x[k]$ clasificată ca și tonală se adună puterea a 2 componente vecine, folosind formula:

$$P_1\left(f_1 = k \frac{f_e}{N}\right) = 10 \cdot \lg \left(10^{\frac{P_x(k-1)}{10}} + 10^{\frac{P_x(k)}{10}} + 10^{\frac{P_x(k+1)}{10}} \right) \quad (4.12)$$

unde N reprezintă numărul de componente ale lui $P_x [k]$ iar f_c frecvența de eșantionare a semnalului de vorbire.

Pentru componentele clasificate ca și netonale, se adună, în fiecare bandă critică, puterile lor aplicând o formulă echivalentă celei din relația (4.12). Toate componentele, tonale sau nu, care au o putere inferioară pragului de audibilitate absolut pot fi eliminate. Dacă două componente tonale sunt distanțate cu mai puțin de 0,5 Barks atunci cea de putere mai mică poate fi eliminată.

După această prelucrare se obțin N_t componente tonale și N_n componente netonale, cu $N_t \leq 24$ și $N_n \leq 24$. Scopul acestei reduceri a numărului de componente este limitarea complexității prelucrării care urmează. Pragul de mascare $S_m(f_2)$ este calculat adunând contribuția la frecvența f_2 a celor N_t componente tonale și a celor N_n componente netonale:

$$S_m(f_2) = 10 \cdot \lg \left(10^{\frac{S_a(f_2)}{10}} + \sum_{k=1}^{N_t} 10^{\frac{P_2(f_2, f_1, P_1)}{10}} + \sum_{k=1}^{N_n} 10^{\frac{P_2(f_2, f_1, P_1)}{10}} \right) \quad (4.13)$$

unde P_2 se calculează folosind relația (4.3). Se reamintește că în această relație frecvențele sunt calculate în Barks și că influența unui sunet mascant nu se manifestă decât într-o plajă $[f_1 - 3, f_1 + 8]$ Barks. Aceste frecvențe trebuie bineînțeles să fie obținute în urma unei discretizări. Această discretizare este descrisă în cadrul standardului MPEG Audio cu relația următoare:

$$\begin{cases} f_2 \in [0, 3] \text{ kHz} & \Rightarrow i \in [0, 48] \\ f_2 \in [3, 6] \text{ kHz} & \Rightarrow i \in [49, 72] \\ f_2 \in [6, 15,6] \text{ kHz} & \Rightarrow i \in [73, 108] \end{cases} \quad (4.14)$$

Cu ajutorul pragului de mascare se poate calcula în fiecare bandă critică raportul semnal pe mască:

$$\text{RSM} = \frac{P_s}{S_m} \quad (4.15)$$

unde cu P_s s-a notat puterea semnalului din banda respectivă iar cu S_m pragul de mascare din aceeași bandă.

4.1.2.2. Utilizarea pragului de mascare la compresia semnalului de vorbire

Raportul semnal pe mască poate fi utilizat pentru cuantizarea semnalului de vorbire. Deoarece prin cuantizare numărul de biți alocat fiecărui eșantion al acestui semnal scade, rezultă că se realizează o compresie. Cuantizarea trebuie realizată în așa fel încât puterea zgomotului de cuantizare să fie în fiecare bandă critică inferioară pragului de mascare. Dacă se îndeplinește această condiție atunci cuantizarea este transparentă. Presupunând că în fiecare bandă critică a semnalului de vorbire se face câte o cuantizare entropică, puterea zgomotului de cuantizare este conform [1], relația (148):

$$P_z = \frac{1}{12} \cdot 2^{h(S)} \cdot 2^{-2b} \quad (4.16)$$

unde $h(S)$ reprezintă entropia diferențială a semnalului S , de la intrarea în cuantizorul corespunzător benzii critice considerate, iar b reprezintă numărul de biți folosiți pentru descrierea semnalului de la ieșirea din acest cuantizor. Punând condiția ca această putere să fie inferioară pragului de mascare se obține:

$$\frac{1}{12} \cdot 2^{2 \cdot h(S)} \cdot 2^{-2b} < S_m$$

sau:

$$2^{2b} > \frac{2^{2h(S)}}{S_m} \quad (4.17)$$

Dacă în locul cuantizării entropice se realizează o cuantizare uniformă atunci:

$$2^{2 \cdot h(S)} = P_s$$

[2], și relația (4.17) devine:

$$2^{2b} > \frac{1}{12} \cdot \text{RSM} \quad (4.18)$$

Ultima relație (la fel ca și relația (4.17)) permite alegerea numărului de biți b pentru care cuantizarea semnalului de vorbire în banda critică considerată să fie transparentă.

4.2. Cuantizarea adaptivă în domeniul TPC

În paragraful destinat studiului detectorului de prag s-a considerat pentru demonstrarea propoziției care reprezintă rezultatul central al aceluia paragraf că se utilizează o cuantizare uniformă. Așa cum s-a arătat în [1], se pot obține rezultate mai bune dacă se folosește o cuantizare neuniformă. Aceasta se poate baza pe utilizarea modelului psiho-acustic, așa cum s-a arătat în paragraful anterior. Deficiența principală a unei astfel de metode este faptul că ea necesită un volum de calcul important, determinarea pragului de mascare fiind o operație laborioasă. În continuare se prezintă o soluție foarte simplă pentru realizarea unei cuantizări neuniforme. Secvența obținută la ieșirea blocului TO din figura 2.1.1., semnalul $y[n]$, reprezintă, datorită faptului că s-a ales ca și transformare ortogonală TPC, spectrul de amplitudini instantaneu al blocului curent din semnalul de vorbire. De aceea poate fi realizată o cuantizare perceptivă, bazată pe o aproximare a modelului psiho-acustic al semnalului de vorbire. Suportul semnalului $z[n]$, obținut la ieșirea blocului DP, din figura 2.1.1. este împărțit în 32 de intervale, pe care sunt definite blocurile de semnal $z_k[n]$, $k = \overline{1,32}$, de aceeași lungime. Acestea corespund benzilor critice din modelul psiho-acustic al vorbirii. Din fiecare dintre aceste benzi, blocul DP a eliminat toate componentele spectrale, cu modulul mai mic decât un anumit prag, t .

Această valoare aproximează valoarea pragului de mascare. După cum s-a arătat în paragraful dedicat detectorului de prag, valoarea t este aleasă în mod adaptiv, folosind un algoritm de maximizare a raportului semnal pe zgomot pentru semnalul reconstruit. Se realizează cuantizarea uniformă a fiecărui bloc. În acest scop se detectează valorile maxime ale semnalelor $z[n]$ și $z_k[n]$, z_M și z_{kM} . Pentru fiecare bloc este alocat un anumit număr de biți. Această procedură este bazată pe valorile z_{kM} . Pentru valoarea z_M sunt alocați 6 biți (2^6 nivele de cuantizare). Pentru valorile z_{kM} se alocă:

$$\gamma_k = \left\lceil \left\lceil \frac{z_{kM}}{z_M} \cdot 2^6 \right\rceil \right\rceil \quad (4.19)$$

nivele de cuantizare, unde $\lceil \cdot \rceil$ simbolizează funcția parte întreagă. Astfel un număr de $b_k = \lceil \log_2(\gamma_k + 1) \rceil$ biți sunt alocați pentru fiecare eșantion al blocului cu indicele k . Cuantizarea acestui bloc este realizată folosind transformarea:

$$u_k[n] = \left\lceil \left\lceil \frac{z_k[n]}{z_{kM} + 0.01} \cdot \gamma_k \right\rceil \right\rceil \quad (4.20)$$

În acest mod se realizează o normalizare de nivel în fiecare bloc. Denormalizarea corespunzătoare va trebui să fie realizată în faza de reconstrucție înainte de calculul TPCI. Această operație va fi realizată de blocul notat cu D în figura 2.1.1. Marele avantaj al procedurii de cuantizare propusă este datorat proprietății de decorelare a TPC. Datorită acestei proprietăți numeroase valori z_{kM} sunt nule. Și valorile corespunzătoare γ_k și b_k sunt nule. De aceea numărul total de biți alocat eșantioanelor semnalului $u[n]$, N_b , este foarte mic în comparație cu numărul de biți pe care a fost reprezentat semnalul $x[n]$. Această procedură de cuantizare are și un mic dezavantaj. Pentru transmiterea sau memorarea fiecărui bloc al semnalului $u[n]$, $u_k[n]$, trebuie adăugate la "coordonatele" fiecărui eșantion câteva valori suplimentare, valorile z_{kM} . Folosind aceste valori și relația (3.46) numerele γ_k pot fi calculate în faza de reconstrucție. Cu ajutorul parametrilor z_{kM} și γ_k pot fi realizate "operațiile inverse" operațiilor descrise de relația (4.2). Deoarece numărul de biți cerut pentru reprezentarea valorilor z_{kM} este foarte mic în comparație cu N_b factorul de compresie global nu este afectat de necesitatea adăugării valorilor suplimentare pentru fiecare bloc $u_k[n]$. Valoarea factorului de compresie realizat de sistemul din figura 2.1.1. poate fi calculată folosind relația:

$$f_c = \frac{16 \cdot N}{N_c + N_p + B} \quad (4.3)$$

unde s-a presupus că fiecare eșantion al semnalului de intrare este codat pe 16 biți, N_p reprezintă numărul de biți necesar pentru codarea pozițiilor eșantioanelor semnalului $u[n]$ și B reprezintă numărul de biți cerut pentru codarea parametrilor. Numerele N_c și B pot fi calculate folosind relațiile:

$$N_c = \sum_{k=1}^{32} N(k) \cdot b_k \quad (4.4)$$

unde $N(k)$ reprezintă numărul de eşantioane nenule ale blocului cu indicele k și:

$$B = \sum_{k=1}^{32} b_k \quad (4.5)$$

Numărul N_p se calculează cu relația:

$$N_p = \sum_{k=1}^{32} N(k) \cdot \zeta_k \quad (4.6)$$

unde ζ_k reprezintă numărul de biți necesar pentru reprezentarea poziției fiecărui eşantion de valoare nenulă din blocul cu indicele k . Deoarece în fiecare bloc există un număr maxim de 32 de astfel de poziții valorile ζ_k sunt mai mici decât 5. De aceea o margine superioară pentru N_p este:

$$5 \cdot \sum_{k=1}^{32} N(k) = 32 \cdot N_n \quad (4.7)$$

4.3. Celelalte blocuri ale sistemului de compresie

Utilizarea codorului Co din figura 2.1.1 conduce la creșterea factorului de compresie fără a afecta nivelul distorsiunii de reconstrucție, deoarece acest sistem realizează o compresie fără pierderi. Implementarea acestui bloc face apel la una dintre tehnicile clasice de codare ca de exemplu codarea Huffman sau codarea aritmetică, [52], [56]. Utilizarea unor astfel de blocuri pentru compresia semnalelor audio este experimentată în [57]. Rezultate experimentale precum și concluzii referitoare la utilizarea unor astfel de blocuri în scheme de compresie sunt prezentate în aceeași lucrare.

Semnalul $v[n]$ de la ieșirea codorului reprezintă rezultatul procedurii de compresie. Este semnalul care se memorează sau se transmite. Celelalte două blocuri din cadrul sistemului din figura 2.1.1. sunt utilizate în faza de reconstrucție. Blocul D realizează decodarea semnalului $v[n]$. La ieșirea sa se obțin semnalele $u_k[n]$ și secvența z_{kM} , $k = \overline{1,32}$. Folosind relația (3.46) se calculează valorile γ_k . Apoi se realizează denormalizarea:

$$w_k[n] = \frac{u_k[n]}{\gamma_k + 0.01} \cdot z_{kM} \quad (4.8)$$

Prin concatenarea componentelor $w_k[n]$ se obține semnalul $w[n]$. Ultimul bloc din figura 2.1.1. calculează TPCI. Rezultatul este semnalul $\hat{x}[n]$. Acesta reprezintă rezultatul procedurii de reconstrucție. Folosind acest semnal poate fi calculată distorsiunea D. Toate operațiile descrise deja sunt repetate, pentru diferite valori ale pragului, t , în scopul maximizării valorii factorului de compresie global f_c , cu constrângerea ca rsz_0 să fie mai mare decât β .

Capitolul 5. Simulări ale metodei de compresie

În acest capitol se vor prezenta câteva experimente în care se implementează sistemul de compresie a vorbirii, propus în capitolul 2, în diferite variante. Complexitatea sistemului crește de la experiment la experiment. Pentru toate experimentele este comprimat același semnal de vorbire, o propoziție în limba engleză; "Huston we have a problem". Fișierul sursă este în format "wave". La crearea sa a fost folosită o frecvență de eșantionare de aproximativ 44 KHz, fiecare eșantion fiind codat pe 16 biți. Graficul acestui semnal este prezentat în figura 5.1.

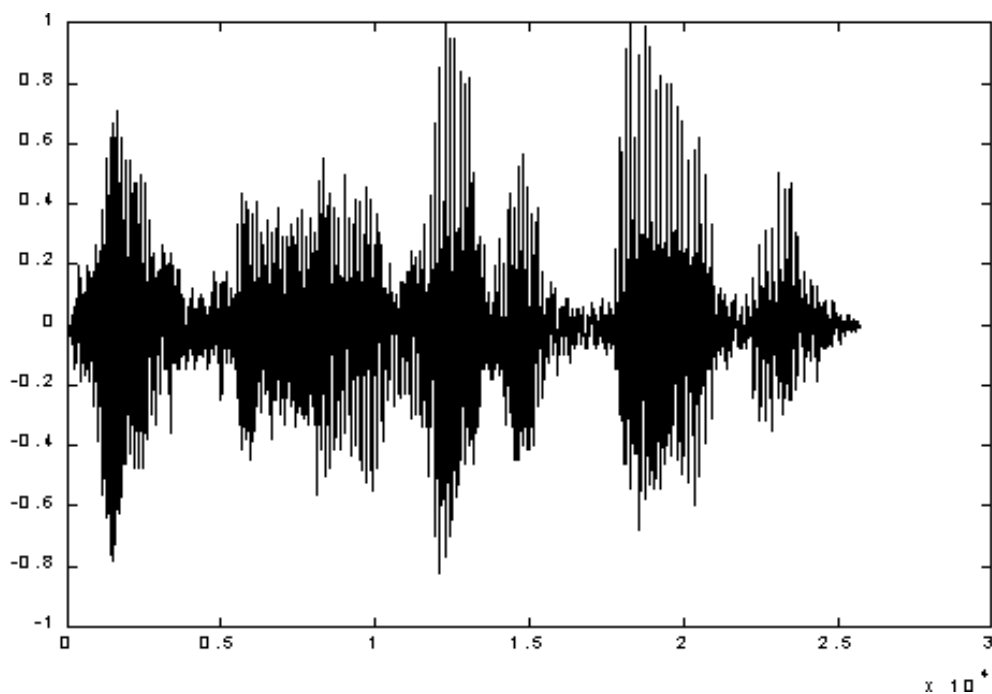


Figura 5.1. Forma de undă a semnalului de prelucrat.

Semnalul este împărțit în blocuri de câte 1024 de eșantioane (durata fiecărui bloc astfel obținut fiind inferioară valorii de 25 ms). Aceste simulări au fost efectuate în Matlab folosind "toolbox"-ul Wavelab. Acesta este prezentat în [58]. Pentru fiecare experiment se prezintă programul folosit, rezultatele obținute și se fac comentarii.

5.1. Primul experiment

Acesta este cel mai simplu experiment. Sistemul de compresie a vorbirii nu este adaptiv. Se utilizează TCP cu o singură iterație (identică cu DCT). Se utilizează o valoare de prag pentru blocul DP egală cu t_m . Se efectuează o cuantizare uniformă pe 6 biți. Nu se utilizează blocurile Co și D. Efectul utilizării acestora este doar estimat. În continuare se prezintă programul care implementează primul experiment.

Programul 1.

Șirul de instrucțiuni

```

[s,fs,wmode,fix]=readwav('problem.wav','y',-1,-1);
x=s(1025:2048);

i=x.*x; Pi=sum(i);
[n,D]=dyadlength(x);
cp=cpanalysis(x,0,'Sine');

t=0.1.*sqrt(Pi./1024);
eps=t; stree=calcestattree(cp,'N(eps)',eps);
[btree,vtree]=bestbasis(stree,0);
Coef=fpt_cp(btree,x,0,'Sine');
y=hardthresh(coef,t);

Ampl=max(abs(y));
N=nnz(y);
b=round(y./ampl.*64);

Nc=N.*16;

yf=ampl.*b./64; z=ipt_cp(btree,yf,0,'Sine');
e=-x+z; o=e.*e; Po=sum(o);
rsb=10.*log10((Pi)/(Po));
fc=(1024.*16)./(Nc);
Figure(1);
Subplot(121); plot(x); title('s de intr');
Subplot(122); plot(z); title('s. rec');
Figure(2); plot(abs(y)); title('spectru');
Figure(3); a=1/512:1/512:1; v=fc.*a./a;
w=rsb.*a./a;
Subplot(121); plot(a,v); title('fc'); subplot(122);
plot(w); title('RSZ0');
save rez z -ascii

```

Comentariu

Achiziția semnalului de vorbire

Segmentarea semnalului de vorbire (Primul segment)

Calculul puterii semnalului de intrare

Analiza cu pachete cosinusoidale

Deoarece se utilizează o singură iterație TPC este identică cu DCT

Inițializarea pragului

Alegerea celei mai bune baze

Calculul transformării cu pachete cosinusoidale

Compararea cu pragul

Cuantizarea scalară uniformă a coeficienților rămași

Determinarea coeficientului de valoare maximă

Determinarea numărului de coeficienți nenuli

Cuantizarea propriuzisă

Determinarea nr. de biți folosiți pentru cuantizare

Se folosesc 6 biți pentru codarea valorii eșantionului curent și 10 pentru codarea poziției sale

Calculul transformării inverse

Calculul raportului semnal pe zgomot la ieșire

Calculul factorului de compresie pe segment

Reprezentări grafice

Semnalul de intrare

Semnalul reconstruit

Analiza spectrala a segmentului

Factorul de compresie și raportul semnal pe zgomot

Salvarea segmentului reconstruit

În urma rulării acestui program pentru cel de al doilea segment s-au obținut rezultatele prezentate în figurile următoare.

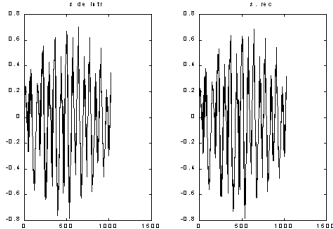


Figura 5.1.1. Semnalele de intrare și reconstruit.

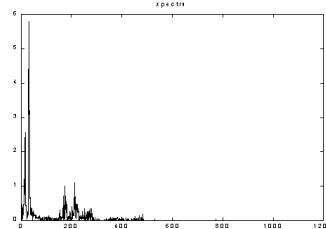


Figura 5.1.2. Spectrul semnalului reconstruit.

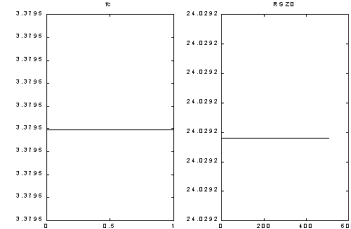


Figura 5.1.3. $f_c = 3,38$,
 $rsb = 24dB$

Rulând pentru fiecare segment al semnalului de prelucrat acest program se obțin rezultatele prezentate în tabelul următor.

Nr. de ordine	f_c	rsz_0 [dB]	Observații
1	1,9248	23,531	zgomot preponderent. (figura 5.1.4)
2	3,38	24	sinus preponderent
3	2,9854	23,8294	sinus preponderent
4	1,4692	27,1865	Preponderent zgomot
5	1,3581	25,7967	Preponderent zgomot
6	2,0317	22,1999	sinus preponderent
7	2,5793	23,3007	sinus preponderent
8	2,6056	22,7519	sinus preponderent
9	3,3907	25,3689	sinus preponderent
10	2,4323	23,9739	sinus preponderent
11	1,9284	22,9190	sinus preponderent
12	1,9140	25,1832	Preponderent zgomot
13	2,5600	25,7500	sinus preponderent
14	2,7234	25,5362	Preponderent zgomot
15	3,9690	26,0653	sinus preponderent (figura 5.1.5)
16	1,7124	25,1859	Preponderent zgomot
17	1,2800	24,0950	Preponderent zgomot (figura 5.1.6)
18	3,4020	27,5556	sinus preponderent
19	2,6528	27,1692	sinus preponderent
20	2,5793	27,3962	Preponderent zgomot

Tabelul 5.1.1. Rezultatele experimentului 1.

Pe ultima coloană a tabelului sunt prezentate observații legate de calitatea segmentului prelucrat. Dacă forma sa de undă este "curată" și dacă indicele ultimei componente spectrale semnificative este inferior lui 512 atunci segmentul se consideră sinusoidal. În caz contrar segmentul se consideră zgomotos. Un exemplu de segment preponderent sinusoidal este segmentul 2 descris de figurile 5.1.1. și 5.1.2. Un exemplu de segment preponderent zgomotos este cel al segmentului 1. Spectrul semnalului obținut după detecția de prag, pentru segmentul 1, este prezentat în figura 5.1.4. Se observă că acesta conține o componentă de zgomot alb.

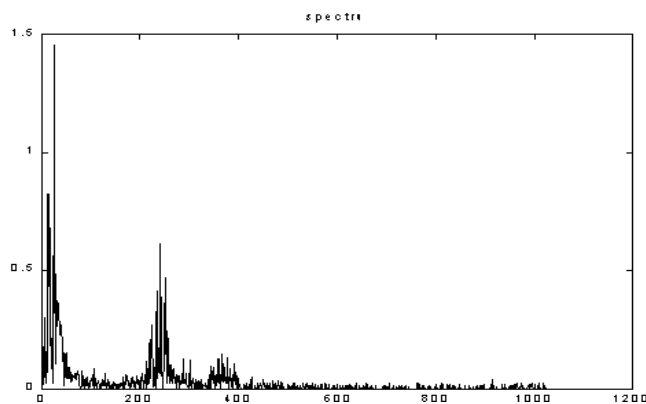


Figura 5.1.4. Spectrul unui segment considerat preponderent zgomotos.

Acest zgomot ar putea fi înlăturat dacă s-ar utiliza o valoare superioară pentru prag, t . De aceea se poate afirma că procedura de alegere adaptivă a pragului este utilă pentru compresie.

Conform tabelului 5.1.1., segmentul cu cel mai mare factor de compresie este segmentul 15 iar cel cu cel mai mic factor de compresie este segmentul 17. În figura 5.1.5 este prezentat segmentul 15 iar în figura 5.1.6 segmentul 17.

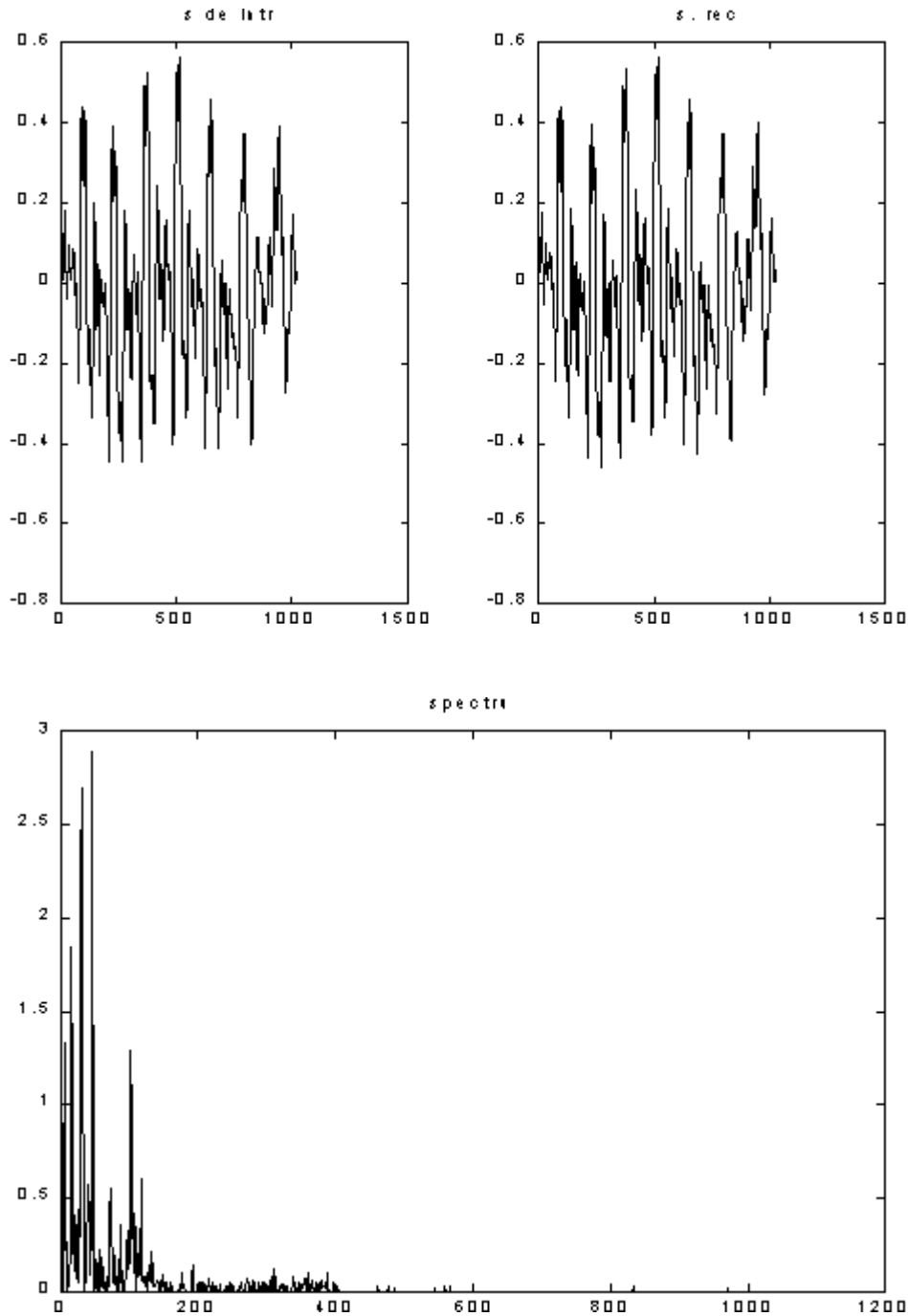


Figura 5.1.5. Formele de undă de la intrarea și ieșirea sistemului de compresie (sus) pentru segmentul 15 și spectrul semnalului reconstruit (jos).

Se constată că segmentul cu cel mai mare factor de compresie este unul preponderent sinusoidal. Există foarte puțini coeficienți nenuli cu indice superior lui 400. În consecință se poate afirma că metoda de compresie funcționează mai bine dacă se respectă modelul sinusoidal al vorbirii. Și pe baza figurii 5.15. se constată că valoarea pragului, t , ar putea fi crescută.

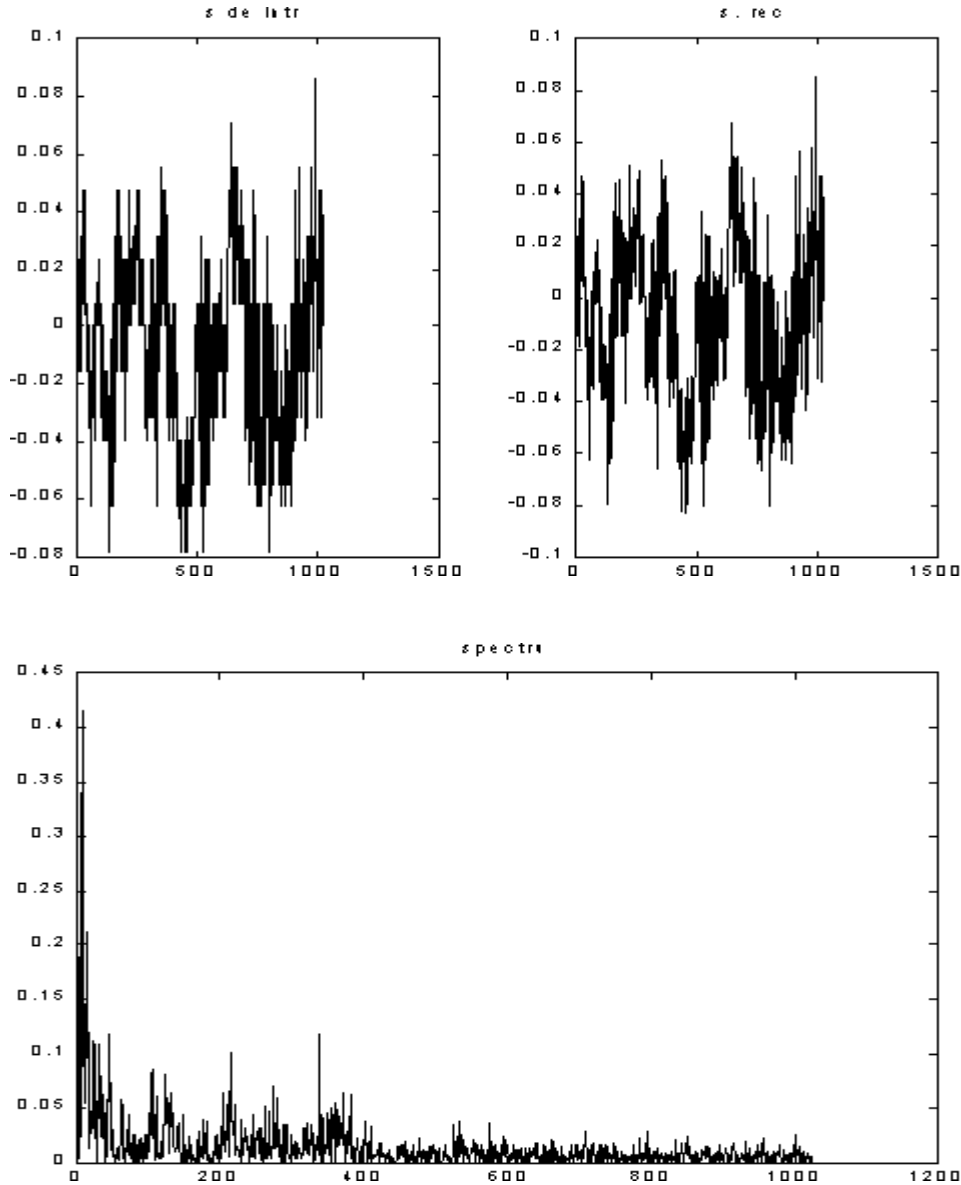


Figura 5.1.6. Formele de undă de la intrarea și ieșirea sistemului de compresie (sus) pentru segmentul 17 și spectrul semnalului reconstruit (jos).

Conform figurii 5.1.6. se constată că în cazul segmentului 17 modelul sinusoidal al vorbirii este respectat în mai mică măsură, motiv pentru care acest segment a fost considerat preponderent zgomotos. Analizând tabelul 5.1.1. se constată că valorile obținute pentru factorul de compresie sunt mici, fiind cuprinse între 1,28 (segmentul 17) și 3,97 (segmentul 15). În schimb valorile raportului semnal pe zgomot la ieșire sunt mari, fiind cuprinse între 22,75 (segmentul 8, preponderent sinusoidal) și 27,55 (segmentul 18, preponderent zgomotos), deși s-a realizat o cuantizare pe un număr de doar 6 biți. Se constată că există rezerve pentru reducerea raportului semnal pe zgomot la ieșire (prin

creșterea valorii de prag, t) în special la segmentele preponderent zgomotoase. **Însă principalul motiv pentru valoarea mică a factorilor de compresie înregistrați este faptul că în calculul acestor factori s-a considerat că pozițiile eşantioanelor semnalului comprimat sunt codate pe 10 biți ($2^{10} = 1024$).** Această problemă poate fi rezolvată dacă se apelează la metoda de cuantizare adaptivă (bazată pe utilizarea benzilor critice) prezentată în capitolul anterior.

5.2. Al doilea experiment

Scopul acestui experiment este de a evidenția importanța numărului de iterații al TPC asupra factorului de compresie. În cazul experimentului anterior nu se făcea de fapt alegerea unei cele mai bune baze deoarece, nefăcându-se nici o iterație, TPC se reducea la DCT. În acest experiment numărul total de iterații va fi de 10 (arborele TPC fiind impus de algoritmul de alegere a celei mai bune baze, care minimizează numărul de coeficienți mai mari decât pragul t). Programul care stă la baza acestui experiment este următorul.

Programul 2

```
[s,fs,wmode,fidx]=readwav('problem.wav','y',-1,-1); x=s(17409:18432);
%calculul puterii semnalului de intrare; i=x.*x; Pi=sum(i);
% analiza cu pachete cosinusoidale; [n,D]=dyadlength(x); cp=cpanalysis(x,D,'Sine');
%inițializarea pragului; t=0.1.*sqrt(Pi./1024);
%alegerea celei mai bune baze; eps=t; stree=calcstattree(cp,'N(eps)',eps);
[btree,vtree]=bestbasis(stree,D);
%calculul transformării cu pachete cosinusoidale; coef=fpt_cp(btree,x,D,'Sine');
% compararea cu pragul; y=hardthresh(coef,t);
% cuantizarea scalară uniformă a coeficienților rămași;
% determinarea coeficientului de valoare maximă; ampl=max(abs(y)); N=nnz(y);
b=round(y./ampl.*64);
%determinarea numărului de biți folosiți pentru cuantizare; Nc=N.*16;
% calculul transformării inverse; yf=ampl.*b./64; z=ipt_cp(btree,yf,D,'Sine');
% calculul raportului semnal pe zgomot la ieșire;
e=-x+z'; o=e.*e; Po=sum(o); rsb=10.*log10((Pi)/(Po));
% calculul factorului de compresie pe segment; fc=(1024.*16)/(Nc);
a=1/512:1/512:1;v=fc.*a./a; w=rsb.*a./a;
% reprezentări grafice; figure(1); subplot(121); plot(x); title('s de intr'); subplot(122);plot(z); title('s.
rec');
figure (2); plotbasistree(btree,D,stree,'Sine'); title('arbore');
```

Au fost îngroșate caracterele instrucțiunilor care sunt diferite în programul 2 față de programul 1. În tabelul următor se prezintă rezultatele rulării acestui program pentru segmentele 1-10, 15 și 17.

În cazul segmentelor pentru care numărul de iterații rezultat în urma aplicării algoritmului de căutare a celei mai bune baze este 0 rezultatele efectuării experimentelor 1 și 2 sunt identice. Este vorba despre segmentele 2,3,7 și 15. În cazul celorlalte segmente se obțin factori de compresie ceva mai mari (ceea ce dovedește superioritatea aplicării TPC asupra transformării DCT) și rapoarte semnal pe zgomot la ieșire ceva mai mici. Cel mai mic raport semnal pe zgomot de ieșire obținut este mai mare de 23 dB. Aceasta este încă o valoare prea mare. În consecință se poate trage aceeași concluzie ca și la experimentul 1 și anume că valoarea pragului, t , poate fi crescută. În figurile 4.2.1. și 4.2.2 sunt prezentați arborii obținuți, în urma aplicării algoritmului de căutare a celei mai bune baze pe segmentele 4 și 5.

Nr. de ordine	f_c	rsz_0 [dB]	Observații
1	1.9922	23.3	1 iterație
2	3.38	24	0 iterații
3	2.9854	23,8294	0 iterații
4	1.4819	26.3483	10 iterații (fig 5.2.1)
5	1.4124	26.0918	10 iterații (fig 5.2.2)
6	2.1695	26.1181	10 iterații
7	2,5793	23,3007	0 iterații
8	3.5189	23.2196	1 iterație
9	3,3907	25,3689	0 iterații
10	2.645	23.4842	1 iterație
15	3,9690	26,0653	0 iterații
17	1.4322	26.5688	10 iterații

Tabelul 4.2.1. Câteva rezultate pentru experimentul 2.

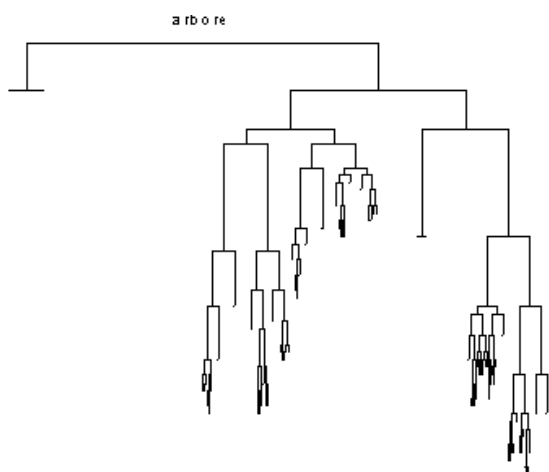


Figura 5.2.1. Arborele celei mai bune baze pentru segmentul 4.

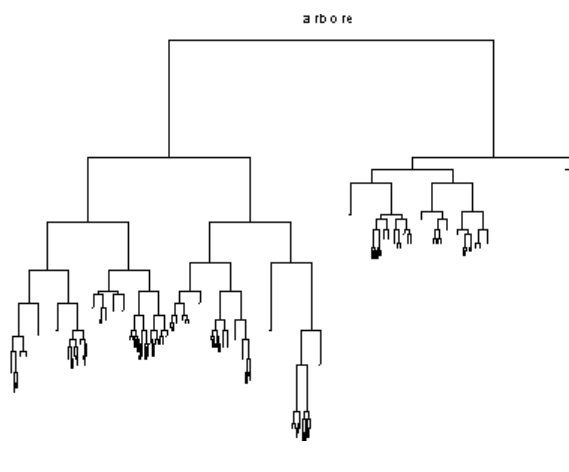


Figura 5.2.2. Arborele celei mai bune baze pentru segmentul 5.

Evident pentru a fi posibilă aplicarea TPCI în scopul reconstrucției este necesară și cunoașterea acestor arbori. De aceea pe lângă coeficienții transformării (valoare și poziție) este necesară și memorarea sau transmiterea acestui arbore în scopul reconstrucției. De aici rezultă necesitatea codării arborelui celei mai bune baze. Pentru această operație sunt necesari biți suplimentari. De numărul lor nu s-a ținut seama în calculul factorilor de compresie prezentați în tabelul 5.2.1. De aceea factorii de compresie care se pot obține, când se utilizează TPC sunt de fapt ceva mai mici.

O posibilitate de codare a arborelui celei mai bune baze este descrisă de următoarele reguli:

- la fiecare iterație (nivel de divizare a suportului semnalului de prelucrat) se asociază un 1 dacă se face divizarea și un 0 dacă nu se face divizarea;
- fiecare nivel al arborelui este parcurs de la stânga la dreapta.

Un exemplu de aplicare a acestei metode de codare este prezentat în figura următoare:

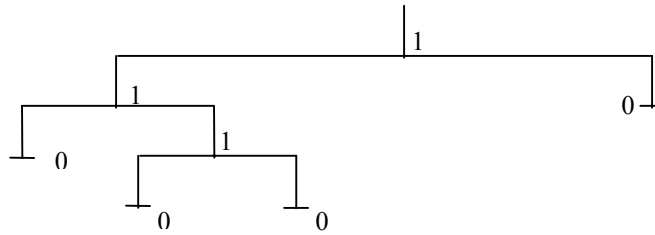


Figura 5.2.3. Un exemplu de codare al unui arbore de cea mai bună bază.

S-a obținut codul: 110010000000000. Acesta se exprimă pe 15 biți ($15=2^4-1$). Deci numărul de biți necesar pentru codarea, folosind metoda descrisă, a unui arbore care are 10 nivele de divizare, este de $2^{11} - 1$. E clar că aceasta este o valoare inadmisibil de mare. De aceea o soluție mai bună pare a fi folosirea unui algoritm de căutare a celei mai bune baze care să utilizeze un număr mai mic de nivele de divizare. Dacă se utilizează doar trei nivele de divizare, ca în exemplul din figura 5.2.3., atunci pentru codarea arborelui celei mai bune baze sunt necesari doar 16 biți. Acest număr nu afectează valoarea factorului de compresie. E clar că și în acest caz utilizarea TPC conduce la factori de compresie mai mari decât utilizarea DCT. În continuare se repetă experimentul 2 limitând numărul de divizări din algoritmul de căutare al celei mai bune baze la maximum 3. În acest mod se obține experimentul 2.1. Programul aferent acestui experiment are două instrucțiuni diferite față de programul 2. Acestea sunt: **[btree,vtree]=bestbasis(stree,3)** și **fc=(1024.*16)/(Nc+16)**.

Rezultatele obținute în urma efectuării experimentului 2.1. sunt prezentate în tabelul următor.

Nr. de ordine	f_c	rsz_0 [dB]	Observații
1	1.9922	23.3	1 iterație
2	3.38	24	0 iterații
3	2.9854	23,8294	0 iterații
4	1.4713	26.4203	1 iterație
5	1.3763	26.2515	1 iterație
6	2.1070	21.5087	1 iterație
7	2,5793	23,3007	0 iterații
8	3.5189	23.2196	1 iterație
9	3,3907	25,3689	0 iterații
10	2.645	23.4842	1 iterație
15	3,9690	26,0653	0 iterații
17	1.2800	24.0950	0 iterații

Tabelul 4.2.2. Câteva rezultate ale experimentului 2.1.

În urma efectuării experimentului 2.1. s-au obținut modificări, față de experimentul 2, pe segmentele 4, 5, 6 și 17. Valorile obținute pentru factorii de compresie corespunzători sunt puțin mai mici (dar încă mai mari decât cele din cazul experimentului 1) iar valorile rapoartelor semnal pe zgomot la ieșire puțin mai mari. De aceea aceasta va fi modalitatea de aplicare a algoritmului de căutare a celei mai bune baze care se va folosi în continuare. Blocul de date corespunzător fiecărui segment va trebui să aibă un antet, de 16 biți, care va conține codul arborelui celei mai bune baze specific aceluia segment. Deoarece valorile factorilor de compresie obținuți pe fiecare segment sunt încă prea mici, în continuare se descrie un nou experiment, în care cuantizarea scalară uniformă, pe 6 biți, avută în vedere până acum, se înlocuiește cu o cuantizare scalară neuniformă (descrisă în capitoul anterior).

5.3. Al treilea experiment

Secvența coeficienților TPC (calculată pe baza alegerii celei mai bune baze folosind ca și funcție de cost numărul de coeficienți superiori valorii de prag, t) este prelucrată de detectorul de prag. Semnalul obținut la ieșirea acestuia, este cuantizat neuniform. În acest scop el este împărțit în 32 de benzi și în fiecare dintre acestea se realizează o normare și o cuantizare uniformă pe 6 biți. Programul folosit pentru efectuarea acestui experiment este prezentat în continuare.

Programul 3.

```
% achiziția semnalului de prelucrat;
[s,fs,wmode,fdx]=readwav('problem.wav','y',-1,-1);
% segmentarea;
x=s(16385:17408);
%calculul puterii semnalului de intrare;
i=x.*x; Pi=sum(i);
% analiza cu pachete cosinusoidale;
[n,D]=dyadlength(x); cp=cpanalysis(x,D,'Sine');
%inițializarea pragului;
t=0.1.*sqrt(Pi./1024);
%alegerea celei mai bune baze;
eps=t; stree=calctattree(cp,'N(eps)',eps); [btree,vtree]=bestbasis(stree,3);
%calculul transformării cu pachete cosinusoidale;
coef=fpt_cp(btree,x,D,'Sine');
% compararea cu pragul;
y=hardthresh(coef,t);
% împărțirea în 32 benzi;determinarea numărului de nivele de cuantizare în fiecare bandă
b(k);determinarea nr. de biți în fiecare bandă nb(k); cuantizarea uniformă pe 6 biți în fiecare
bandă; determinarea numărului de coeficienți nenuli în fiecare bandă N(k);
% inițializarea;
yk=zeros(32,32); ykc=zeros(32,32); ykf=zeros(32,32); yf=zeros(1,1024);
% împărțirea în 32 de benzi:
for k=1:1:32; yk(k,:)=y(32.*(k-1)+1:32.*k);
% normarea în fiecare bandă;
% determinarea valorii maxime în fiecare bandă;
z(k)=max(abs(yk(k,:)));
% cuantizarea uniformă pe 6 biți a valorii maxime din fiecare bandă;
b(k)=round(z(k)./(ampl).*64);
% determinarea numărului de biți necesar pentru cuantizarea benzii k;
nb(k)=round(log2(b(k)+1));
% cuantizarea uniformă pe 6 biți în fiecare bandă;
ykc(k,:)=round((yk(k,:)./(z(k)+0.01)).*b(k));
% determinarea numărului de coeficienți nenuli din fiecare bandă;
N(k)=nnz(ykc(k,:));
% denormarea în fiecare bandă;
ykf(k,:)=(ykc(k,:)./(b(k)+0.01)).*z(k);
% reasamblarea secvenței de 1024 de eşantioane;
yf=[ykf(1,:) ykf(2,:) ykf(3,:) ykf(4,:) ykf(5,:) ykf(6,:) ykf(7,:) ykf(8,:) ykf(9,:) ykf(10,:) ykf(11,:)
ykf(12,:) ykf(13,:) ykf(14,:) ykf(15,:) ykf(16,:) ykf(17,:) ykf(18,:) ykf(19,:) ykf(20,:) ykf(21,);
```

```

ykf(22,:) ykf(23,:) ykf(24,:) ykf(25,:) ykf(26,:) ykf(27,:) ykf(28,:) ykf(29,:) ykf(30,:) ykf(31,:)
ykf(32,:)]];
% determinarea numărului de biți folosiți pentru cuantizare;
Nc=N*nb'; end;
% determinarea numărului de biți necesar ptr. codarea valorilor maxime din fiecare bandă;
B=sum(nb);
% determinarea numărului de coeficienți nenuli;
Nn=sum(N);
% calculul transformării inverse;
z=ipt_cp(btree,yf,D,'Sine');
% calculul raportului semnal pe zgomot la ieșire;
e=-x+z'; o=e.*e; Po=sum(o); rsb=10.*log10((Pi)/(Po));
% calculul factorului de compresie pe segment; numărul de biți necesari pentru codarea
pozițiilor este 5Nn deoarece pentru codarea a 32 de poziții (câte există într-o bandă sunt necesari
5 biți), 16 biți se folosesc pentru codarea arborelui celei mai bune baze;
fc=(1024.*16)./(Nc+Nn.*5+B+16);
% reprezentări grafice;
figure(1); subplot(121); plot(x); title('s de intr'); subplot(122); plot(z); title('s. rec'); figure (2);
plotbasistree(btree,D,stree,'Sine'); title('arbore');
%save rez z -ascii;
Instrucțiunile care fac diferența dintre programele 2.1. și 3 au fost scrise cu litere îngroșate. În tabelul
următor se prezintă rezultatele rulării programului 3.

```

Nr. de ordine	f_c	rsz_0 [dB]	Observații
1	3.4183	26.0701	1 iterație
2	4.7850	28.8661	0 iterații
3	4.4692	27.8201	0 iterații
4	2.5564	25.4162	1 iterație
5	2.4813	23.1911	1 iterație
6	3.2946	26.8812	1 iterație
7	4.0716	27.3254	0 iterații
8	4.0059	26.9479	0 iterații
9	5.1506	27.7076	1 iterație
10	3.8406	27.4574	0 iterații
15	5.6672	28.3064	0 iterații
17	2.9008	21.5571	0 iterații

Tabelul 5.3.1. Rezultatele experimentului 3.

Se constată că toți factorii de compresie au crescut. Valorile rapoartelor semnal pe zgomot la ieșire sunt încă prea mari (se poate considera că s-a realizat o compresie transparentă dacă raportul semnal pe zgomot la ieșire este mai mare decât 20 dB). Factorii de compresie ai fiecărui segment pot fi creșcuți în continuare dacă se face o alegere adaptivă a valorii de prag, t , (astfel încât raportul semnal pe zgomot la ieșire să fie în jur de 20 dB pentru fiecare segment) și dacă se utilizează metode de codare a datelor performante. Este vorba de trei categorii de date: valorile coeficienților nenuli, pozițiile coeficienților nenuli și arborele celei mai bune baze. Pentru fiecare categorie de date pot fi realizate codări (compresii fără pierderi) performante, capabile să scadă numărul de biți necesari pentru reprezentarea categoriei respective. De exemplu, în cazul arborelui celei mai bune baze poate fi utilizată o codare run-length, ținând seama de faptul că acest șir de date conține secvențe lungi de biți de valoare 0. Pentru codarea valorilor coeficienților nenuli s-ar putea utiliza o codare Huffman sau o codare aritmetică. La fel pentru codarea pozițiilor coeficienților nenuli. Structura blocului de date

corespunzător unui segment, inspirată de metoda de compresie folosită în experimentul 3 este prezentată în figura următoare.

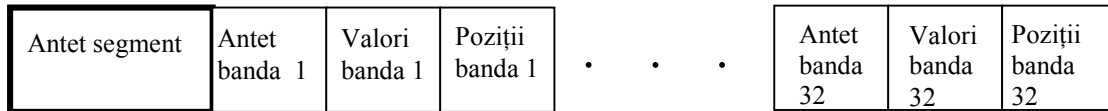


Figura 5.3.1. Structura blocului de date corespunzător unui segment, inspirată de experimentul 3.

În antetul segmentului este codat arborele celei mai bune baze și apoi valoarea maximă a coeficienților de pe acel segment. De aceea acesta va fi un cuvânt de 22 de biți (16 pentru arbore și 6 pentru valoarea maximă). În antetele benzilor sunt codate valorile maxime ale coeficienților din benzile respective precum și numărul de coeficienți nenuli din banda respectivă. De aceea ele au o lungime de 11 biți. În blocurile "Valori banda k" sunt codate valorile coeficienților nenuli din banda k, $k = 1, 32$. Lungimea unui astfel de bloc este un multiplu de 6. În blocurile "Poziții banda k" sunt codate pozițiile coeficienților nenuli din banda k. Fiecare astfel de poziție este codată pe 5 biți. De aceea lungimea acestui bloc trebuie să fie un multiplu de 5. Trebuie remarcat că există multe benzi care conțin doar coeficienți nuli. Blocul corespunzător unei astfel de benzi conține doar 14 biți, toți de valoare 0. Datorită prezenței grupului de 14 zerouri într-un astfel de antet se știe că este vorba despre o bandă care conține doar coeficienți de valoare nulă. De aceea următorul grup de biți este considerat un nou antet de bandă (dacă antetul curent nu era al 32-lea) sau un nou antet de segment.

În continuare se prezintă ultimul experiment din acest capitol, care se referă la alegerea adaptivă a valorii pragului, t.

5.4. Al patrulea experiment

Se utilizează în acest scop algoritmul adaptiv de alegere a pragului, prezentat în capitolul anterior. În continuare se prezintă programul utilizat pentru realizarea acestui experiment.

Programul 4.

```
[s,fs,wmode,fidx]=readwav('problem.wav','y',-1,-1); x=s(1:1024);
%calculul puterii semnalului de intrare; i=x.*x; Pi=sum(i);
% analiza cu pachete cosinusoidale; [n,D]=dyadlength(x); cp=cpanalysis(x,D,'Sine');
%inițializarea pragului; t=0.1.*sqrt(Pi./1024);
%alegerea celei mai bune baze; eps=t; stree=calcstattree(cp,'N(eps)',eps);
[btree,vtree]=bestbasis(stree,3);
%calculul transformării cu pachete cosinusoidale; coef=fpt_cp(btree,x,D,'Sine');
% compararea cu pragul; y=hardthresh(coef,t);
% determinarea raportului semnal pe zgomot după hardthresh; e=-coef+y; o=e.*e; Po=sum(o);
% alegerea pragului optim; rsz=10.*log10((Pi)./(Po)); while (rsz>=20), t=t+t./10; eps=t;
stree=calcstattree(cp,'N(eps)',eps); [btree,vtree]=bestbasis(stree,3);
coef=fpt_cp(btree,x,D,'Sine');
y=hardthresh(coef,t); e=-coef+y; o=e.*e; Po=sum(o); rsz=10.*log10((Pi)./(Po)); end;
ampl=max(y);
% împărțirea în 32 benzi;determinarea numărului de nivele de cuantizare în fiecare bandă
b(k);determinarea nr. de biți în fiecare bandă nb(k); cuantizarea în fiecare bandă; determinarea
numărului de coeficienți nenuli în fiecare bandă N(k); inițializare; yk=zeros(32,32);
ykc=zeros(32,32); ykf=zeros(32,32); yf=zeros(1,1024);
```

```

% împărțirea în 32 de benzi: for k=1:1:32; yk(k,:)=y(32.*(k-1)+1:32.*k);
% normarea în fiecare bandă; determinarea valorii maxime în fiecare bandă; z(k)=max(abs(yk(k,:)));
%cuantizarea uniformă pe 6 biți a valorii maxime din fiecare bandă;
b(k)=round(z(k)./(ampl).*64);
% determinarea numărului de biți necesar pentru cuantizarea benzii k; nb(k)=round(log2(b(k)+1));
% cuantizarea uniformă pe 6 biți în fiecare bandă; ykc(k,:)=round((yk(k,:)/(z(k)+0.01)).*b(k));
% determinarea numărului de coeficienți nenuli din fiecare bandă; N(k)=nnz(ykc(k,:));
% denormarea în fiecare bandă;
ykf(k,:)=(ykc(k,:)/(b(k)+0.01)).*z(k);
% reasamblarea secvenței de 1024 de eșantioane;
yf=[ykf(1,:) ykf(2,:) ykf(3,:) ykf(4,:) ykf(5,:) ykf(6,:) ykf(7,:) ykf(8,:) ykf(9,:) ykf(10,:) ykf(11,:)
ykf(12,:) ykf(13,:) ykf(14,:) ykf(15,:) ykf(16,:) ykf(17,:) ykf(18,:) ykf(19,:) ykf(20,:) ykf(21,:)
ykf(22,:) ykf(23,:) ykf(24,:) ykf(25,:) ykf(26,:) ykf(27,:) ykf(28,:) ykf(29,:) ykf(30,:) ykf(31,:)
ykf(32,:)];
% determinarea numărului de biți folosiți pentru cuantizare; Nc=N*nb'; end;
% determinarea numărului de biți necesar ptr. codarea valorilor maxime din fiecare bandă;
B=sum(nb);
% determinarea numărului de coeficienți nenuli; Nn=sum(N);
% calculul transformării inverse; z=ipt_cp(btrees,yf,D,'Sine');
% calculul raportului semnal pe zgomot pe segment;
e=-x+z'; o=e.*e; Po=sum(o); rsb=10.*log10((Pi)/(Po));
% calculul factorului de compresie pe segment; numărul de biți necesari pentru codarea pozițiilor este
5Nn deoarece pentru codarea a 32 de poziții (câte există într-o bandă) sunt necesari 5 biți, 16 biți se
folosesc pentru codarea arborelui celei mai bune baze; fc=(1024.*16)/(Nc+Nn.*5+B+16);
% reprezentări grafice; figure(1); subplot(121); plot(x); title('s de intr'); subplot(122); plot(z); title('s.
rec'); figure (2); plotbasistree(btrees,D,stree,'Sine'); title('arbore');
%salvarea segmentului reconstruit; Acesta se salvează în format ascii în directorul
d\Matlab5\toolbox\Wavelab. El poate fi reîncărcat dacă se transferă în d\Matlab5\lucru și apoi se
folosește load rez1 ASCII.
%save rez1 z -ascii;

```

Caracterele instrucțiunilor noi, în comparație cu programul 3, au fost îngroșate. În tabelul următor se prezintă rezultatele obținute în urma rulării acestui program. Cele 25 de segmente obținute în urma reconstrucției (efectuată pe fiecare segment) au fost concatenate obținându-se semnalul reconstruit în urma compresiei. Forma sa de undă este prezentată în figura următoare și poate fi comparată cu forma de undă a semnalului inițial, prezentată în figura 5.1. Cele două semnale, inițial și reconstruit pot fi și ascultate. În acest scop au fost create două fișiere de tip wav, init.wav și reconstr.wav. În urma ascultării semnalului reconstruit se remarcă că metoda de compresie descrisă este transparentă, propoziția rostită putând fi înțeleasă cu ușurință. Se remarcă și faptul că metoda de compresie îmbunătățește chiar calitatea semnalului inițial. Acesta este însoțit de un zgomot de fond, care este perceput mai ales la sfârșitul fragmentului. Acest zgomot de fond nu mai poate fi perceput în semnalul reconstruit. Acesta este însă puțin distorsionat, percepându-se un efect de cadare, care este datorat erorilor la marginile segmentelor care apar în calculul TPC. Efectul de cadare poate fi înlăturat prin aplicarea unei metode de tip overlap and save, așa cum se arată în [32]. Analizând tabelul 5.4.1. se constată faptul că s-au obținut valori acceptabile pentru factorii de compresie și pentru rapoartele semnal pe zgomot la ieșire, pentru fiecare segment. Cel mai mic factor de compresie, de 4,05, a fost obținut pe segmentul 17 iar cel mai mare factor de compresie, de 18,3, pe segmentul 15. La valorile factorilor de compresie pe segment deranjează în special dispersia mare obținută. Cel mai mic factor de compresie este de 4 ori mai mic decât cel mai mare factor de compresie obținut.

Nr. de ordine	f_c	rsz_0 [dB]	Observații
1	7.3968	19.4129	zgomot preponderent. (3 iterații)
2	15.9844	18.9756	sinus preponderent. (0 iterații)
3	12.6615	18.9907	sinus preponderent. (3 iterații)
4	4.4317	19.1134	preponderent zgomot. (2 iterații)
5	4.1145	18.8286	preponderent zgomot. (0 iterații)
6	7.8694	19.5280	sinus preponderent. (2 iterații)
7	13.3856	19.4330	sinus preponderent. (0 iterații)
8	12.7900	19.3277	sinus preponderent. (0 iterații)
9	16.4333	19.6479	sinus preponderent. (2 iterații)
10	11.3384	19.3460	sinus preponderent. (2 iterații)
11	8.1189	19.5016	sinus preponderent. (2 iterații)
12	6.1826	19.2819	preponderent zgomot. (3 iterații)
13	7.8168	19.5904	sinus preponderent. (2 iterații)
14	12.6811	19.7608	preponderent zgomot. (0 iterații)
15	18.3061	19.7918	sinus preponderent. (0 iterații)
16	7.4847	18.5319	preponderent zgomot. (1 iterație)
17	4.0534	16.3397	preponderent zgomot. (0 iterații)
18	18.0441	19.3680	sinus preponderent. (3 iterații)
19	12.1094	19.8325	sinus preponderent. (0 iterații)
20	8.3464	19.1047	preponderent zgomot. (0 iterații)
21	9.3303	19.7703	2 iterații
22	14.1853	19.0936	1 iterație
23	14.2099	18.1952	1 iterație
24	14.8271	19.2795	1 iterație
25	8.5289	17.4512	1 iterație

Tabelul 5.4.1. Rezultatele experimentului 4.

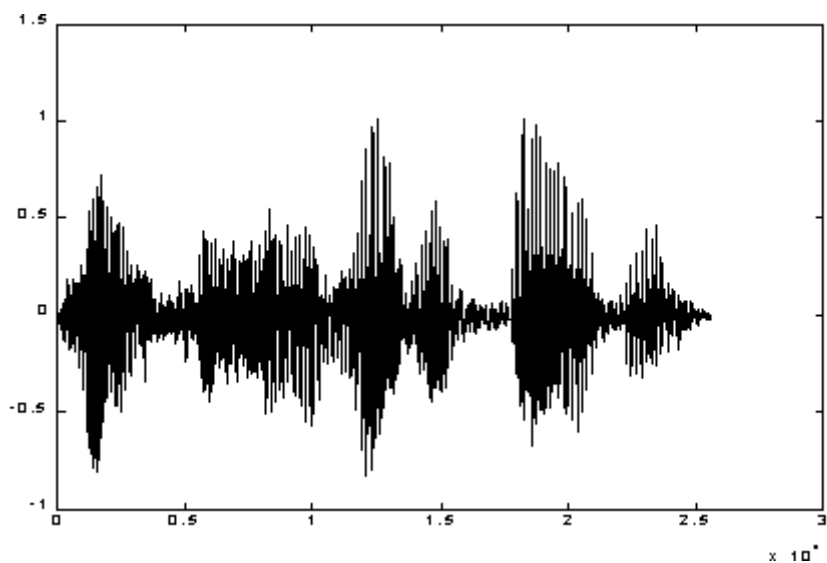


Figura 5.4.1. Forma de undă a semnalului reconstruit.

Pe baza analizei tabelului se constată că valori mai mici ale factorului de compresie, cuprinse între 4 și 8 au fost obținute pe segmentele clasificate ca și zgomotoase.

Repartiția rapoartelor semnal pe zgomot la ieșire este mult mai omogenă. Cea mai mică valoare, 16,33 dB, s-a înregistrat pe segmentul 17 iar cea mai mare valoare, 19,83, pe segmentul 19. Toate aceste valori sunt suficient de mari pentru a certifica o reconstrucție de calitate. Valoarea medie a factorului de compresie obținut este de 10,82. Aceasta este o valoare destul de ridicată ținând seama de faptul că nu s-au utilizat metode de compresie fără pierderi pentru codarea arborelui celei mai bune baze respectiv pentru valorile și pozițiile coeficienților. Este de presupus că dacă s-ar fi utilizat și astfel de tehnici de codare valoarea medie a factorului de compresie ar fi fost de cel puțin 1,5 ori mai mare, adică de 16,23.

Capitolul 6. Concluzii

Problema centrală a acestei teze, prezentată în capitoul 1, este compresia semnalului de vorbire, realizată cu ajutorul funcțiilor "wavelet". Schema sistemului de compresie propus este cea din figura 2.1.1. și reprezintă firul roșu al acestei lucrări. Este vorba despre o schemă de compresie cu pierderi de informație controlate. O astfel de metodă de compresie se pretează în cazul semnalului de vorbire deoarece acesta este foarte redundant. Schema propusă este specifică pentru un sistem de compresie bazat pe utilizarea unei transformări ortogonale. O astfel de transformare este utilă deoarece realizează decorelarea semnalului de prelucrat, făcând posibilă eliminarea unor eșantioane din domeniul transformatei, fără a afecta semnificativ conținutul informațional al semnalului de prelucrat. Există și alte transformări care pot realiza decorelarea, neortogonale, dar acestea sunt mai redundante decât transformările ortogonale, motiv pentru care au fost evitate în această lucrare. Există mai multe transformări ortogonale care ar putea fi folosite pentru compresia semnalului de vorbire. Cea care realizează decorelarea maximă este transformarea Karhunen-Loeve. Din păcate nu există algoritmi rapizi pentru implementarea acestei transformări. Ea este o transformare dependentă de semnalul prelucrat, bazată pe inversarea unei matrici, operație consumatoare de timp și de volum de calcul. Există semnale pentru care matricea nu este nici măcar inversabilă. În lucrarea de față se propune utilizarea transformărilor ortogonale, bazate pe teoria funcțiilor "wavelet". Acestea reprezintă descompuneri ale semnalului de prelucrat în baze ortonormale ale căror elemente se obțin prin translatarea și scalarea unei funcții unice, numită funcție "wavelets mother". Avantajul major al acestor funcții este că ele posedă simultan o bună localizare temporală și frecvențială. De aceea o astfel de descompunere, numită serie de funcții "wavelet", are un număr mic de coeficienți de valori semnificative. Compresia se realizează prin transmisia sau memorarea acestor coeficienți, în locul eșantioanelor semnalului de prelucrat.

Cunoștințele necesare din cadrul teoriei funcțiilor "wavelet" au fost prezentate în capitoul 2, dintr-o perspectivă originală, cea a teoriei codării în subbenzi. Au fost prezentate schemele de codare și decodare în subbenzi. De asemenea s-au prezentat principalele concepte ale teoriei funcțiilor "wavelet", analiza multirezoluție și descompunerea ortogonală și s-a arătat că acestea pot fi descrise prin același formalism matematic ca și codarea în subbenzi. Au fost analizate trei tipuri de transformări ortogonale, bazate pe teoria funcțiilor "wavelet", transformarea "wavelet" discretă, TUD, transformarea cu pachete de funcții "wavelet" discretă, TPWD și transformarea cu pachete sinusoidale discretă, TPC. A fost evidențiat efectul de decorelare al acestor transformări, demonstrându-se că toate trei converg asimptotic spre transformarea Karhunen-Loeve. Aceste demonstrații au fost întărite prin exemple. Pe baza acestor exemple s-a constatat că, din punct de vedere al vitezei de convergență spre transformarea Karhunen-Loeve, cel mai bine se comportă TPC.

Una dintre dificultățile ridicate de folosirea TUD sau TPWD este necesitatea alegerii funcției de tip "wavelets mother" care se utilizează pentru calculul acestor transformări. Această alegere ar putea fi făcută în acord cu forma semnalului de prelucrat. Un criteriu de alegere, util în aplicațiile de compresie, a fost elaborat și publicat de către autorul acestei teze, [43].

Utilizarea TPC are și avantajul că permite alegerea celei mai bune baze, folosind un criteriu util pentru compresie și anume minimizarea numărului de coeficienți ai transformatei mai mari decât un anumit prag. Tot în capitoul 2, pe baza modelului sinusoidal al vorbirii, se sugerează că cea mai potrivită transformare ortogonală, dintre cele trei amintite mai sus, pentru compresia vorbirii, este TPC. Demonstrația acestei propoziții ar merita să se aprofundeze.

În capitoul 3 s-a prezentat detectorul de prag din schema din figura 2.1.1. Se remarcă caracterul adaptiv al acestui sistem. Se propune un algoritm simplu și destul de rapid pentru implementarea acestui bloc. Se prezintă o modalitate de inițializare a pragului, utilă pentru creșterea vitezei algoritmului de detecție de prag. Algoritmul adaptiv propus se bazează pe minimizarea erorii medii pătratice de aproximare a semnalului de prelucrat prin semnalul reconstruit în urma compresiei.

Ținând seama de caracterul ortogonal al transformărilor utilizate, nu este necesară, pentru calculul erorii de aproximare, reconstrucția semnalului supus compresiei. Eroarea medie pătratică de aproximare poate fi calculată în domeniul transformatei, folosind semnalul de la ieșirea blocului de calcul al transformării ortogonale și semnalul de la ieșirea blocului de cuantizare.

Deși criteriul erorii medii pătratice nu este cel mai potrivit pentru aprecierea unei metode de prelucrare a vorbirii, simplitatea sa îl face foarte folositor. Poate că ar merita reluarea acestei probleme și rezolvarea sa prin minimizarea unui alt criteriu, bazat pe calculul unei distanțe, construită folosind modelul psiho-acustic al vorbirii, așa cum se propune în [59].

Capitolul 4 prezintă sistemul de cuantizare din schema din figura 2.1.1. Acesta este un sistem adaptiv, care funcționează pe baza unei variante simplificate a modelului psiho-acustic al audiției vorbirii. Construcția sa exploatează și utilizarea TPC în schema de compresie aleasă. Utilizarea acestui sistem are o contribuție importantă la creșterea factorului de compresie al sistemului propus. Sistemul de cuantizare este foarte simplu, în fiecare bandă critică a semnalului de vorbire efectuându-se o cuantizare uniformă pe 6 biți. Cuanta este diferită de la bandă critică la bandă critică, deoarece în fiecare astfel de bandă se realizează normarea la valoarea maximă. De aceea se poate considera că pe ansamblu sistemul realizează o cuantizare neuniformă. Aceasta este și adaptivă deoarece se ține seama de forma semnalului de prelucrat prin intermediul valorilor maxime din fiecare bandă critică. Sistemul de cuantizare poate fi privit și ca și un sistem de compresie a dinamicii semnalului de prelucrat (cum sunt sistemele Dolby sau DNL) realizând aceeași excursie pentru fiecare dintre semnalele din fiecare bandă critică. Un demers de același tip este prezentat în [61].

Capitolul 5 este dedicat simulării sistemului de compresie descris în capitolele anterioare. Se fac mai multe experimente, mergând de la simplu la complex. În primul experiment, cel mai simplu, blocurile din structura sistemului de compresie nu sunt adaptive. În urma efectuării acestui experiment se constată existența a două tipuri de segmente de vorbire: preponderent sinusoidale și preponderent zgomotoase. Pentru primele, modelul sinusoidal al vorbirii este respectat în mare măsură, în timp ce pentru celelalte modelul sinusoidal este respectat în mică măsură. Se remarcă că pentru segmentele preponderent zgomotoase se obțin factori de compresie mai mici. Cel de al doilea experiment se referă la utilizarea unui prim bloc adaptiv. Este vorba de blocul pentru calculul transformării ortogonale. În primul experiment se folosea transformarea cosinus discretă, DCT (dacă nu se efectuează nici o iterație în calculul TPC atunci se calculează de fapt DCT). În cel de al doilea experiment se calculează TPC, folosindu-se algoritmul de căutare a celei mai bune baze care minimizează funcționala de cost dată de numărul de coeficienți ai TPC mai mari decât un anumit prag. Valoarea acestui prag este cea care corespunde formulei de inițializare propusă în capitolul 3. Este vorba deci despre o transformare ortogonală adaptivă. Superioritatea TPC asupra DCT este evidențiată pe baza creșterii valorilor factorilor de compresie obținuți în cel de al doilea experiment față de cei obținuți în primul experiment. Pe baza celui de al doilea experiment se demonstrează și necesitatea limitării numărului de iterații al TPC. Cel de al treilea experiment este destinat studierii efectului utilizării unui bloc de cuantizare adaptiv. Se dovedește că folosind un astfel de bloc valorile factorilor de compresie obținuți pe fiecare segment cresc substanțial. În sfârșit cel de al patrulea experiment are ca scop evidențierea îmbunătățirilor care se pot obține dacă se utilizează și un detector de prag adaptiv. Se constată că și utilizarea unui astfel de bloc crește substanțial valorile factorilor de compresie de pe fiecare segment. Pentru simularea fiecărui experiment s-a utilizat câte un program Matlab. Aceste programe au fost construite recurent, pentru fiecare nou experiment completându-se cu câteva instrucțiuni noi programul corespunzător experimentului anterior. Astfel programul obținut pentru ultimul experiment reprezintă programul final, cel care simulează funcționarea întregului sistem adaptiv de compresie. De aceea concluziile care se pot trage pentru experimentul 4 sunt valabile pentru metoda de compresie propusă. Acestea sunt:

- Folosind metoda de compresie descrisă se obține un factor de compresie mediu de 10,82, pentru semnalul de prelucrat considerat, superior factorului de compresie realizat de codorul GSM care este de 8 (conform calculului efectuat în [2], la începutul capitolului 6).

- Factorul de compresie al metodei propuse este și mai mare dacă pentru codarea datelor rezultante, conținute în arborele celei mai bune baze și în valorile și pozițiile coeficienților nenuli obținuți, se utilizează metode de codare performante, ca de exemplu codarea run-length sau codarea aritmetică. În toate calculele de factor de compresie efectuate în această lucrare nu s-a ținut seama de faptul că pentru transmiterea semnalului comprimat, pe canale de telecomunicații ar trebui făcută și codarea canalului. Aceasta este o operație redundantă care scade valoarea factorului de compresie global. S-a procedat în acest fel din două motive:
 - nici în cazul standardelor de compresie a semnalului de vorbire deja publicate, cum ar fi GSM sau MPEG, nu se ține seama de codarea canalului la calculul factorului de compresie;
 - există aplicații ale metodei de compresie a vorbirii propusă în această teză, la care nu este necesară codarea canalului, de exemplu memorarea unei convorbiri sau telefonie pe INTERNET (în acest caz se poate considera că se folosește un canal fără zgomot).Oricum pe baza analizei statistice efectuate în paragraful 3.2 se constată că metoda de compresie propusă în această lucrare este robustă.
- Metoda propusă asigură o calitate bună a reconstrucției (pe fiecare segment se obține un raport semnal pe zgomot la ieșire superior valorii de 16 dB, iar valoarea medie a raportului semnal pe zgomot, pe ansamblul segmentelor depășește 20 dB), putând considera că s-a realizat o compresie transparentă, în timp ce codorul GSM, evocat mai sus, nici măcar nu estimează calitatea reconstrucției. De altfel ascultând semnalul reconstruit se constată că zgomotul care perturbă componenta utilă a semnalului de prelucrat a fost în mare măsură eliminat. Este remarcabil faptul că valorile coeficienților nenuli obținuți sunt cuantizate doar pe 6 biți;
- Metoda de compresie folosită este destul de rapidă, ținând seama de faptul că numărul de înmulțiri necesar nu este exagerat. De exemplu calculul TPC al unei secvențe se face la fel de repede ca și calculul FFT-ului aceleși secvențe. De aceea ar merita să se încerce implementarea acestei metode de compresie a vorbirii pe un procesor de semnal;
- Programul 4 ar trebui modificat, în așa fel încât pe fiecare segment să se rezolve problemele la margini, care dau caracterul sacadat al semnalului reconstruit;
- Poate că ar fi util ca înaintea aplicării metodei de compresie să fie crescut raportul semnal pe zgomot al semnalului de prelucrat, folosindu-se în acest scop metoda propusă în [6];
- Metoda de compresie propusă, poate beneficia, la fel ca și metoda de compresie bazată pe predicție liniară, folosită în standardul GSM, de blocuri de preprocesare, cum ar fi de exemplu blocul de identificare a intervalelor de liniște dintre cuvinte, respectiv propoziții sau fraze. În acest mod ar putea fi crescut factorul de compresie global;
- Pentru aprecierea obiectivă a metodelor de compresie a vorbirii ar putea fi utilizate măsurile propuse în [59], bazate pe utilizarea modelului psiho-acustic al vorbirii. De altfel așa cum s-a arătat în paragraful destinat prezentării metodei de cuantizare folosite, se pare că există niște legături subtile între metoda de compresie propusă și modelul psiho-acustic al vorbirii care ar merita să fie investigate mai în detaliu. Ambele se bazează pe analiză spectrală (de fapt coeficienții TPC pot fi priviți ca și componente spectrale ale unor segmente ale blocului de vorbire supus compresiei), ambele fac apel la benzi critice și ambele îndepărtează anumiți coeficienți prin detecție de prag. Un astfel de demers este făcut în [61].
- O posibilitate de creștere a factorului de compresie ar fi și combinarea metodei de compresie bazată pe teoria funcțiilor "wavelet", care face obiectul tezei de față cu metoda de compresie, bazată pe predicția liniară, care stă la baza standardului GSM. S-ar putea ca segmentele preponderent sinusoidale să fie prelucrate cu metoda bazată pe teoria funcțiilor "wavelet" iar segmentele preponderent zgomotoase să fie prelucrate cu metoda bazată pe predicția liniară.

- Pentru metoda de compresie a vorbirii prezentată se au în vedere mai multe aplicații. Ea ar putea fi folosită în aplicațiile de telefonie numerică fixă sau mobilă, încadrându-se în categoria metodelor cu factor de compresie ridicat și cu calitate a reconstrucției controlată. De fapt metoda propusă ar putea fi numită compresie cu pierderi (de informație) controlate, lărgindu-se pe baza sa clasificarea metodelor de compresie care până acum conținea doar două tipuri de tehnici de compresie, cele fără pierderi și cele cu pierderi.

Tot pe baza metodei de compresie propuse ar putea fi concepute noi echipamente de redare a vorbirii, de tipul echipamentelor bazate pe tehnica de compresie MP3, folosite pentru redarea muzicii. Cu alte cuvinte această metodă poate fi folosită pentru memorarea de înaltă fidelitate a convorbirilor telefonice. Ar putea fi concepute noi tipuri de roboți telefonici, pe baza acestei tehnici de compresie.

De asemenea această tehnică de compresie ar putea fi folosită și în aplicațiile de secretizare a convorbirilor telefonice. În acest sens se întrevăd două aplicații. În sistemele de protecție a poștei electronice, cum este de exemplu sistemul PGP, [62], înainte de criptarea efectivă a mesajului se face o compresia a acestuia. Operația de compresie crește gradul de securitate al metodei de criptare folosite. Același lucru ar putea fi făcut în sistemele de criptare a vorbirii, folosindu-se metoda de compresie propusă în această lucrare. Cea de a doua aplicație posibilă este legată de "balizarea" semnalului de vorbire, [63]. În astfel de aplicații compresia este privită ca și un atac involuntar. De aceea cunoașterea metodei de compresie poate conduce la elaborarea unei metode de "balizare" robuste la atacul de compresie.

Metoda de compresie a vorbirii propusă în această lucrare ar putea fi folosită și în cadrul sistemelor de telefonie pe INTERNET, respectiv la construcția unor modemuri de mare viteză, [64].

Capitolul 7. Contribuții originale

În continuare se prezintă contribuțiile originale raportate în această teză. Acestea pot fi grupate în trei categorii:

- Contribuții de importanță majoră,
- Contribuții de importanță limitată,
- Contribuții de detaliu.

Oricare dintre acestea poate fi de natură teoretică sau de natură practică.

Pentru început se prezintă principalele contribuții din prima categorie.

1. Utilizarea TPC la compresia vorbirii.
2. Demonstrația convergenței TUD, TPWD și TPC spre transformarea Karhunen-Loeve. Aceste demonstrații sunt prezentate în paragrafele 2.2.1, 2.2.2 și 2.2.3.
3. Algoritmul de selecție adaptivă a pragului pentru blocul DP, prezentat în paragraful 3.1.
4. Algoritmul de cuantizare adaptivă, propus în paragraful 4.2.
5. Alegerea colecției de rutine Wavelab ale mediului Matlab pentru simularea metodei de compresie a vorbirii, făcută în capitolul 5.

În continuare se prezintă contribuțiile din cea de a doua categorie.

1. Alegerea criteriului de selecție a celei mai bune baze, pentru TPC, bazat pe minimizarea numărului de coeficienți de valoare superioară unui prag impus, prezentat în paragraful 2.3.
2. Limitarea numărului de iterații al TPC, la maximum trei, pentru simplificarea codării arborelui celei mai bune baze, făcută în capitolul 5.
3. Exemplificarea convergenței TUD, TPWD și TPC spre transformarea Karhunen-Loeve. Aceste exemple apar în paragrafele 2.2.1, 2.2.2 și 2.2.3.
4. Analiza statistică a detectorului de prag, făcută în paragraful 3.2.
5. Alegerea TPC ca transformare ortogonală folosită pe baza asemănării dintre combinațiile liniare ale elementelor unui pachet cosinusoidal și modelul sinusoidal al semnalului de vorbire, prezentată în paragraful 2.3.
6. Elaborarea sistemului de cuantizare adaptivă în acord cu modelul psiho-acustic al audiției semnalului de vorbire. Acest sistem este prezentat în paragraful 4.2.
7. Structura adaptivă a sistemului de compresie prezentat în figura 2.1.1. Acest sistem folosește blocuri adaptive pentru calculul transformării ortogonale directă și inversă, pentru detecția de prag și pentru cuantizare. Aceste blocuri sunt prezentate, pe rând, în capitolele acestei teze. În acest sistem ar putea fi incluse și blocuri de codare pentru arborele celei mai bune baze și pentru valoarea și poziția coeficienților TPC superiori pragului t , tot de natură adaptivă. În sistemul de compresie propus se remarcă două categorii de adaptări: locale și globale. Adaptările locale se bazează pe criterii diferite. De exemplu alegerea celei mai bune baze pentru pachetul cosinusoidal urmărește minimizarea numărului de coeficienți ai transformării TPC mai mari decât pragul t . Și alegerea numărului de iterații al TPC se face pe baza aceluiași criteriu. Alegerea valorii de prag t , se face în mod adaptiv, urmărindu-se obținerea unei distorsiuni de reconstrucție inferioare unui prag impus. Cuantizarea adaptivă urmărește minimizarea numărului de biți folosiți pentru descrierea semnalului comprimat. Trebuie remarcat faptul că există "un consens" între naturile convergențelor acestor algoritmi, obținându-se o metodă adaptivă globală convergentă. Criteriul global, pe baza căruia se realizează adaptarea sistemului din figura 2.1.1. este cel al păstrării erorii medii pătratice de reconstrucție sub o valoare impusă. Deci criteriul dominant pentru întreaga schemă de compresie este cel folosit pentru alegerea valorii de prag, t .

8. Concepția programelor de simulare a metodelor de compresie a vorbirii, descrise în cele patru experimente, descrise în capitolul 5, cu o structură cu evoluție gradată, de la simplu la complex.

În continuare se prezintă contribuțiile de detaliu.

1. Prezentarea teoriei funcțiilor "wavelet", făcută în capitolul 2, prin prisma teoriei sistemelor de codare în subbenzi.
2. Evitarea utilizării transformărilor "wavelet" redundante, cum ar fi transformările biortogonale, la compresia semnalului de vorbire, pe baza observației O3, din paragraful 2.1.2.
3. Una dintre preocupările cele mai interesante ale autorului acestei teze a fost cea mai bună alegere a răspunsului la impuls $h[n]$, pentru filtrele trece-jos din structura TUD, pe baza semnalului de prelucrat, pentru maximizarea factorului de compresie, la o distorsiune de reconstrucție impusă. Această tehnică este prezentată în [43].
4. Observația din paragraful 2.1.4. conform căreia principalul avantaj al pachetelor de funcții "wavelet" este că acestea oferă mult mai multă libertate în alegerea bazei în care se descompune semnalul de analizat. Având la dispoziție un număr mai mare de subbenzi, se poate îmbunătăți localizarea frecvențială a componentelor semnalului de analizat. Se poate chiar implementa o bancă de filtre de analiză care să aibă exact partiția în benzi critice recomandată de modelul psiho-acustic de audiere a vorbirii.
5. Observația de la sfârșitul paragrafului 2.1.4 conform căreia: Pachetele cosinusoidale rezolvă o deficiență cronică a pachetelor de funcții "wavelet", și anume localizarea în timp. Orice pachet de funcții "wavelet" corespunde unei anumite bănci de filtre de analiză, care realizează o anumită codare în subbenzi. Pe tot parcursul calculului TPWD, aceste filtre rămân neschimbate. În consecință, TPWD nu realizează nici o localizare în timp a acestei bănci de filtre. În cazul TPC, în fiecare interval I_m , se lucrează cu o altă funcție "wavelets mother", deci cu o altă bancă de filtre de analiză, făcându-se o localizare temporală a acestor bănci.
6. Observațiile din paragraful 2.2.1.
7. Ilustrarea proprietății de Gaussianizare a TUD, din figura 3.1.1.3.
8. Evidențierea importanței alegerii funcției de tip "wavelets mother" pentru compresia bazată pe folosirea TUD, cu ajutorul figurilor 2.3.1-2.3.3.
9. Utilizarea modelului matematic al vorbirii din relația 2.3.2.
10. Propoziția 3.1.1. din paragraful 3.1, pe baza căreia se poate inițializa algoritmul de detecție de prag.
11. Aproximarea pragului de mascare din modelul psiho-acustic al audierii vorbirii prin pragul folosit de sistemul DP din schema din figura 2.1.1., prezentată în paragraful 4.2.
12. Formula de calcul al factorului de compresie, (4.3).
13. Operațiile de normalizare și denormalizare, introduse în legătură cu cuantizarea adaptivă, propusă în paragraful 4.2.
14. Programele 1, 2, 3 și 4. Acestea pot fi folosite, cu modificări minimale, la simularea metodei de compresie propusă și în cazul altor semnale de vorbire. Ele au următoarele caracteristici comune:
 - sunt implementate în MATLAB,
 - citesc datele de intrare dintr-un fișier în format WAVE, care conține semnalul de vorbire codat PCM pe 16 biți,
 - fixează valoarea minimă a raportului semnal pe zgomot al semnalului reconstituit la 16 dB, dar pe majoritatea segmentelor se obțin valori mai mari decât 18 dB.

- Lucrează pe segmente ale semnalului de vorbire conținând câte 512 eşantioane.
 - Numărul de iterații al TPC este limitat la maximum 3.
 - Utilizează criteriul de alegere a celei mai bune baze prin minimizarea numărului de coeficienți ai TPC, superiori unei valori de prag impusă.
 - Face o alegere adaptivă a valorii de prag, t .
15. Clasificarea segmentelor semnalului de prelucrat în preponderent sinusoidale și preponderent zgomotoase.
 16. Concepția celor patru experimente descrise în capitolul 5, precum și rezultatele acestora.
 17. Exemplul de codare a arborelui celei mai bune baze din figura 5.2.3.
 18. Structura blocului de date corespunzător unui segment, inspirată de experimentul 3, propusă în figura 5.3.1.
 19. Evidențierea importanței utilizării unui anumit tip de sistem de cuantizare, prin experimentul 3.
 20. Evidențierea importanței utilizării unui sistem de detecție de prag, adaptiv, prin experimentul 4.
 21. Concluziile obținute pe baza efectuării experimentului 4, raportate în capitolul 5.
 22. Concluziile, direcțiile viitoare de cercetare precum și aplicațiile potențiale ale metodei de compresie propuse, prezentate în capitolul 6.
 23. Evidențierea contribuțiilor originale realizată în acest capitol.
 24. Bibliografia este împărțită în două părți. În prima sunt prezentate lucrările care sunt citate în textul tezei. În cea de a doua sunt grupate, pe subiecte, lucrări de interes pentru domeniul tezei, care au fost consultate, la întocmirea celor trei referate aferente: [1], [3] și [5].

Bibliografie

Partea I. Referințele din teză

- [1] A. Cubițchi. "Stadiul actual al tehnicilor de compresie a semnalului vocal în telefonie numerică", Referat nr.1, în cadrul pregătirii pentru doctorat, *Departamentul de comunicații, Facultatea de Electronică și Telecomunicații, Timișoara*, 1999, Conducător științific Prof. dr. ing. Ioan Naforniță
- [2] N. Moreau. "Techniques de compression des signaux", *Masson*, Paris 1995
- [3] A. Cubițchi. "Metode de compresie a semnalului vocal cu eficiență ridicată", Referat nr.2, în cadrul pregătirii pentru doctorat, *Departamentul de comunicații, Facultatea de Electronică și Telecomunicații, Timișoara*, 2000, Conducător științific Prof. dr. ing. Ioan Naforniță
- [4] A. Isar, I. Naforniță, "Reprezentări timp-frecvență", *Ed. Politehnica*, Timișoara, 1998.
- [5] A. Cubițchi. "Metodă de compresie a semnalului vocal bazată pe utilizarea funcțiilor "wavelet"", Referat nr.3, în cadrul pregătirii pentru doctorat, *Departamentul de comunicații, Facultatea de Electronică și Telecomunicații, Timișoara*, 2001, Conducător științific Prof. dr. ing. Ioan Naforniță
- [6] D. Isar "Îmbunătățirea raportului semnal pe zgomot în sistemele de telecomunicații ", teză de doctorat realizată sub conducerea științifică a d-lui Profesor Ioan Naforniță, *Universitatea "Politehnica" Timișoara*, 1998.
- [7] M. V. Wickerhauser, "Adapted Wavelet Analysis. From theory to software", *A. K. Peters Ltd, Massachusetts*, 1994.
- [8] G. Malgouyres. "Introduction a la théorie des ondelettes". *Curs de vară*, Timișoara 1994.
- [9] M. J. T. Smith, T. P. Barnwell III. "Exact Reconstruction Techniques for Tree-Structured Subband Coders". *IEEE Trans. On ASSP*, vol. 34, pp.434-441, 1986.
- [10] A. Cohen. "Ondelettes et traitement numérique du signal". *Masson*, 1992.
- [11] O. Rioul. "A Discrete Time Multiresolution Theory". *IEEE Trans. on SP*, vol. 41, no. 8, pp. 2591-2606, August 1993.
- [12] I. Daubechies. "Orthonormal Bases of Compactly Supported Wavelets". *Comm. Pure Appl. Math.*, No. 41, pp.909-996, 1988.
- [13] I. Daubechies. "Ten Lectures on Wavelets". *SIAM*, Philadelphia 1992.
- [14] Y. Meyer. "Ondelettes, filtres miroirs en quadrature et traitement numérique de l'image". În *Les ondelettes en 1989*. P. G. Lemarié (editor), *Springer-Verlag*, 1990.
- [15] A. N. Akansu, R. A. Hadad. "Multiresolution Signal Decomposition". *Academic Press*, New York, 1992.
- [16] S. Mallat. "Multifrequency Channel Decomposition". *IEEE Trans. on ASSP*, vol. 37, No.12, pp. 2091-2110, Octobre 1989.
- [17] R. Cristescu. "Analiză funcțională". *Editura Didactică și Pedagogică*, București 1965.
- [18] D. L. Donoho, I. M. Johnstone. "Unconditional Bases are Optimal Bases for Data Compression and for Statistical Estimation". *Technical Report 410, Stanford University*, November 1992.
- [19] J. Froment. "Traitement d'images et applications de la transformée en ondelettes". *Teza de doctorat, Universitatea Paris IX*, 1990.
- [20] J. Kovacevic, M. Vetterli. "Nonseparable Two-and Three-Dimensional Wavelets". *Proceedings of ISCAS'93*, Chicago, 1993.
- [21] H. S. Malvar. "Lapped Transforms for Efficient Transform/Subband Coding". *IEEE Trans. on ASSP*, vol. 38, pp.969-978, June 1990.

- [22] A. Cohen, I. Daubechies, J.C. Feauveau, "Bi-orthogonal bases of compactly supported Wavelets", *Comm. in Pure and Applied Math.*, vol. XLV, pp.485-560, 1992.
- [23] R.R. Coifman, M.V. Wickerhauser "Wavelets and adapted waveform analysis" in *Proceedings of symposia in applied mathematics*, SIAM vol. 47, 1993, editor Ingrid Daubechies.
- [24] A. De Sabata, C. Iung, J. F. Aubry. "A Variable Scale DWT". *Proceedings of the International Symposium Etc '94*, vol. III, pp.43-48, Timișoara Sept. 1994.
- [25] M. Vetterli, C. Herley, "Wavelets and filter banks: Theory and design", *IEEE Transactions on signal processing* 40 (9) pp.2207-2232, september 1992.
- [26] T.P. Barnwell III, I. Sodagar, K. Nayebi "Time-varying filter banks and wavelets", *IEEE Transactions on signal processing*, vol.42, no. 11, november 1994.
- [27] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery. "Numerical Recipes in C". *Cambridge University Press*, 1995.
- [28] O. Rioul, M. Vetterli. "Wavelets and signal processing", *IEEE SP Magazine*, 8 (4) pp.14-38, october 1991.
- [29] T. Asztalos, A. Isar. "An Adaptive Data Compression Method Based on the Fast Wavelet Transform". *Proceedings of the International Symposium Etc '94*, Timișoara, Vol III, pp 37-42, Sept. 1994.
- [30] A. Oppenheim, R. W. Schaefer. "Digital Signal Processing". *Prentice Hall*, 1986.
- [31] J.J. Shynk. Frequency-domain and multirate adaptive filtering, *Signal Processing Magazine*, january 1992.
- [32] T. Asztalos. "An Algorithm for the DWT on Block Computation". *Proceedings of the International Symposium Etc '96*, Timișoara, vol II, pp.128-133, Sept. 1996.
- [33] H. Krim, D. H. Brooks. "Feature-Based Segmentation of ECG Signals". *Proceedings of IEEE Conference, TFTS'96*, pp. 97-100, Paris, July 1996.
- [34] O. Rioul. "Ondelettes régulières: Applications à la compression d'images fixes". *These de doctorat, ENST Paris*, Mars 1993.
- [35] D. Donoho. Smooth Wavelet Decomposition with Blocky Coefficient Kernels, in *Recent Advances in Wavelet Analysis*, L. Schumaker and G. Webb (editors), pp. 1-43, 1993.
- [36] W. Sweldens. "Wavelet Sampling Techniques", *Proceedings of the Joint Statistical Meetings, San Francisco*, 1993.
- [37] E.D. Kolaczyk. "Non-parametric Estimation of Gamma-ray Burst Intensities using Wavelets", în curs de publicare în revista *The Astrophysical Journal*.
- [38] T. Asztalos. "Tomography Imaging. Radon Transform Inversion Procedures". *Raport de stagiul*, Universitatea Paris-Sud, Iulie 1997.
- [39] A. Cohen, J. P. d'Ales. "Nonlinear Approximation of Stochastic Processes". În *Wavelets and Statistics*. A. Antoniadis și G. Oppenheim (editori), *Springer Verlag*, pp.129-132, 1995.
- [40] A. Mateescu, M. Răducanu, L. Stanciu. "Best Basis with Wavelet Packets for a Signal". *Proceedings of International Symposium Etc '96*, Timișoara, vol. II, pp.106-111, September 1996.
- [41] M. V. Wickerhauser. "Best-adapted Wavelet Packet Bases", *Proceedings of Symposia in Applied Mathematics*, vol. 47, 1993.
- [42] Y. Meyer. "Ondelettes et algorithmes concurrents". *Herman*, Paris, 1993.
- [43] A. Isar, A. Cubițchi. "A new best wavelet basis searching method for the compression of smooth signals", *Proceedings of IEEE International Conference, ICT 2001*, Bucarest, Romania, 4-7 June, 2001.
- [44] D. Pastor, R. Gay. "Décomposition d'un processus stationnaire du seconde ordre. Propriétés statistiques d'ordre 2 des coefficients d'ondelettes et localisation fréquentielle des paquets d'ondelettes". *Traitement du signal*, vol. 12, no. 5, pp. 393-420, 1995.
- [45] A. Spataru. "Fondements de la théorie de la transmission de l'information". *Presses Polytechniques Romandes*, Lausanne, 1987
- [46] V. E. Neagoe. "Introducing a new orthogonal spatial transform for significant data selection", *Revista Academiei*, nr. 13, *Editura Academiei, București* 1983, pag. 163-180.
- [47] R. Boite, M. Kunt, "Traitement de la parole", *Presses Polytechniques Romandes*, Lausanne, 1987.
- [48] T. Asztalos, Dorina Isar, A. Isar, "Adaptive Capturing Transient Signals using Wavelets", *8-th*

European Conference on POWER ELECTRONICS AND APPLICATIONS, EpE'99, 7-9 September, 1999, Lausanne, Switzerland.

- [49] Rodica Stoian. Compresia de date. Algoritmi de predicție, *Ed. științifică și enciclopedică*, București, 1988.
- [50] E. Pop, I. Naforniță, V. Tiponuț, A. Mihăescu, L. Toma, "Metode în prelucrarea numerică a semnalelor", vol. 1, *Ed. Facla*, Timișoara, 1986.
- [51] E. Pop, I. Naforniță, V. Tiponuț, A. Mihăescu, L. Toma, "Metode în prelucrarea numerică a semnalelor", vol. 2, *Ed. Facla*, Timișoara, 1989.
- [52] Monica Elena Borda, "Teoria transmiterii informației", *Editura Dacia*, Cluj-Napoca, 1999.
- [53] E. Pop, V. Stoica, "Principii și metode de măsurare numerică", *Ed. Facla*, Timișoara, 1977.
- [54] E. Pop, V. Stoica, I. Naforniță, E. Petriu, "Tehnici moderne de măsurare", *Ed. Facla*, Timișoara, 1983.
- [55] Andrei Cubițchi, Alexandru Isar, "A Statistical Characterization of the uniform Quantization Process", *Proceedings of the International Conference of Romanian Military Technical and Technological Research Agency*, April 12-13, 2001, Bucarest, Romania.
- [56] D. Salomon. "Data compression", *The Complete Reference*, Springer Verlag, New-York, 1998.
- [57] T. Asztalos. "Contribuții la compresia, în domeniul transformatelor "wavelet", a semnalelor audio", Teză de doctorat, Conducător științific Prof. dr. ing. Miranda Naforniță, *Facultatea de Electronică și Telecomunicații Timișoara*, 2001.
- [58] J. B. Buckheit, D. L. Donoho, "WaveLab and Reproducible Research", în *Wavelets and Statistics*, editori: A. Antoniadis și G. Oppenheim, pp. 55-83, Springer-Verlag, 1995.
- [59] Irina Coțanis, "Impacting factors on the Objective Measurement Algorithms for Speech Quality Assessment on Mobile Networks", *Proceedings of IEEE International Conference, ICT 2001*, Bucarest, Romania, 4-7 June, 2001.
- [60] M. R. Zurera, F. L. Ferreras, F. C. Roldan, R. J. Martinez, "Transparent audio coding using orthonormal wavelets with any compact support", *International Conference on Signal Circuits and Systems, SCS'99*, 5-7 July, 1999, Iași, pp 203-206.
- [61] T. Asztalos, A. Isar, "Wavelets and Audio Data Compression", *International Conference on Signal Circuits and Systems, SCS'99*, 5-7 July, 1999, Iași, pp 199-202.
- [62] V. V. Patriciu, M. Pietroșanu-Ene, I. Bica, C. Cristea, "Securitatea informatică în UNIX și INTERNET", *Ed. Tehnică, București*, 1998.
- [63] Titu I. Băjenescu, Monica E. Borda, "Securitatea în informatică și telecomunicații", *Editura Dacia*, Cluj-Napoca, 2001.
- [64] M. Naforniță, "Méthodes modernes de traitement du signal pour la compression de données dans les modems haut débit. Premier rapport: Techniques de compression". Raport de cercetare AUPELF-UREF, finanțat de FICU, *Departamentul de Comunicații al Facultății de Electronică și Telecomunicații din Timișoara*, 2000.

Partea a doua

Lucrări despre compresie

- [65] V. E. Neagoe. "Using Legendre Polynomials to Introduce a New Orthogonal Transform for Significant Feature Selection" *Proceedings of Pattern Recognition and Image Processing Conference*, pp.177-182, Las Vegas, June 1982.
- [66] D. Stanomir, C. Negrescu, L. Jalbă. Algoritmi pentru prelucrarea semnalului vocal. Teorie și aplicații în comunicații GSM, *Ed. Athena*, 1998, București.
- [67] R. R. Coifman, N. Saito. The Local Karhunen-Loeve Bases. *Proceedings of the IEEE Conference "TFTS' 96"*, pp.129-132, Paris, July 1996.
- [68] European Technical Standard (ETS-300-960) GSM-Digital Cellular Telecommunications (Phase 2+) Full Rate Speech: Processing Functions (GSM 06.01), 1997.
- [69] European Technical Standard (ETS-300-961) GSM-Digital Cellular Telecommunications (Phase 2+) Transcoding (GSM 06.10), 1997.

- [70] European Technical Standard (ETS-300-965) GSM-Digital Cellular Telecommunications (Phase 2+) Voice Activity Detector (VAD) for Full Rate Speech Traffic Channels (GSM 06.32), 1997.
- [71] European Technical Standard (ETS-300-964) GSM-Digital Cellular Telecommunications (Phase 2+) Discontinuous Transmission (DTX) for Full Rate Speech Traffic Channels (GSM 06.01), 1997.
- [72] European Technical Standard (ETS-300-963) GSM-Digital Cellular Telecommunications (Phase 2) Confort Noise Aspects for Full Rate Speech Traffic Channels (GSM 06.12), 1997.
- [73] European Technical Standard (ETS-300-962) GSM-Digital Cellular Telecommunications (Phase 2) Substitution and Muting of Lost Frames for Full Rate Speech Traffic Channels (GSM 06.11), 1997
- [74] J.V. Macres. Theory and Implementation of the Digital Cellular Standard Voice Coder: VSELP on TMS320C5x, Application Report SPRA 136, Texas Instruments, October 1994.
- [75] ETSI TS 126090 Universal Mobile Telecommunications System (UMTS), Mandatory Speech Codec speech processing functions AMR speech codec, Transcoding functions, (3G TS 26090) 1999.

Lucrări despre cuantizare

- [76] I. Naforniță, A. Câmpeanu, A. Isar, "Semnale Circuite și Sisteme", curs, *Litografia Universității "Politehnica" Timișoara*, 1995.
- [77] [Bon.'62] Bonnet, "Sur la statistique du second ordre des signaux aleatoires quantifiés", *Comptes Rendus de l'Academie de sciences francaise*, 30 Juillet 1962
- [78] P. Fiche, V. Ricordel, S. Labit, "Etude d'algorithmes de quantification vectorielle arborescente pour la compression d'images fixes", *IRISA, Rennes*, 1994.
- [79] I. Daubechies, R. DeVore, C.S. Gunturk, V. Vaisharnpayan, "Exponential Precision in A/D Conversion with an Imperfect Quantizer, Raport de cercetare, universitatea Stanford, 2001.
- [80] P. A. Chou, M. Effros, R. M. Gray, "A Vector Quantization Approach to Universal Noiseless Coding and Quantization". Acceptată pentru publicare în *IEEE Transactions on Information Theory*, pe data de 16 februarie 1996.
- [81] R. M. Gray, R. A. Olshen, "Vector Quantization and Density Estimation", Raport de cercetare, universitatea Stanford, 1990, <http://www-isl.stanford.edu/~gray/compression.html>

Lucrări despre analiză statistică cu funcții "wavelet"

- [82] U. Amato, D. Vuza. Wavelet Regularization for Smoothing Data, Preprint *Instituto per Applicazioni della Matematica CNR* 1994.
- [83] U. Amato, D. Vuza. Besov Regularization, Thresholding and Wavelets for Smoothing Data, Preprint *Instituto per Applicazioni della Matematica CNR*, 1997.
- [84] U. Amato, D. Vuza. An Alternate Proof of a Result of Johnstone and Silverman Concerning Wavelet Threshold Estimators for Data with Correlated Noise, Preprint *Instituto per Applicazioni della Matematica CNR*, 1997.
- [85] U. Amato, D. Vuza. Wavelet Approximation of a Function from Samples Affected by Noise, propusă la *Revista Academiei Române*.
- [86] U. Amato, D. Vuza. A Collection of Routines for the Wavelet Transform of Daubechies Type, Preprint *Instituto per Applicazioni della Matematica CNR*, 1997.
- [87] A. Antoniadis, G. Gregoire, G. Nason. Density and Hazard Rate Estimation for Right Censored Data Using Wavelet Methods, Preprint *laboratoire LMC-IMAG Grenoble*, 1995
- [88] J. Benedetto, A. Teolis. A Wavelet Auditory Model and Data Compression. În *Applied and Computational Harmonic Analysis*. No.1, pp.3-28, February 1993.
- [89] J. B. Buckheit, D. Donoho, Time-Frequency Tillings which Best Expose the Non-Gaussian Behaviour of a Stochastic Process. *Proceedings of the IEEE Conference "TFTS'96"*, pp.1-4, Paris, July 1996.
- [90] S.Cambanis, E.Masry. Wavelet Approximation of Deterministic and Random Signals: Convergence Properties and Rates, *IEEE Transactions on information theory*, vol.40,no.4, July 1994.
- [91] B. S. Chen, C. W. Lin. Multiscale Wiener Filter for the Restoration of Fractal Signals: Wavelets Filter Bank Approach. *IEEE Transactions on Signal Processing*, vol. 42, No. 11, pp.2972-2982, November

- 1994.
- [92] D. L. Donoho. Wavelet Shrinkage and W.V.D.-A Ten Minute Tour. *Technical Report 416, Stanford University*, January 1993.
 - [93] D. Donoho. CART and Best-Ortho-Basis: A Connection, Preprint *Stanford University*, 1995.
 - [94] H. Y. Gao. Wavelets Shrinkage Estimate for Heteroscedatic Regression Models. Preprint *MathSoft*, 1997.
 - [95] H.-Y. Gao. Threshold Selection in WaveShrink, Preprint *MathSoft*, 1997.
 - [96] H.-Y. Gao. Wavelet Shrinkage Denoising Using the Non-negative Garrote, Preprint *MathSoft*, 1997.
 - [97] D. Isar. De-noising adaptatif. *Seizieme Colloque GRETSI*, pp.1249-1252, Grenoble, 15-19 Septembre 1997.
 - [98] S. Mallat, F. Falzon. Understanding Image Transform Codes. *Proceedings of the SPIE Aerospace Conference*, Orlando, April 1997
 - [99] P. Moulin. Wavelet Thresholding Techniques for Power Spectrum Estimation. *IEEE Trans. on S.P.*, vol. 42, No.11, pp. 3126-3136, November 1994.
 - [100] G. P. Nason. Wavelet Regression by Cross-Validation. *Preprint University of Bristol*, March 1994.
 - [101] G. P. Nason, T. Sapantias, A. Sawezenko. Statistical Modeling of Time Series using Non-decimated Wavelet Representations, Preprint *University of Bristol*, 1997.
 - [102] P. Srinivasan, L. M. Jamieson. Techniques for Variable Rate Speech Coding using Wavelet Representations. *Proceedings of the IEEE Conference "TFTS'96*, pp.109-112, Paris, July 1996.
 - [103] C. Taswell. Speech Compression with Cosine and Wavelet Packet Near-Best Bases. *Preprint, Stanford University*, 1995.
 - [104] E. Wesfreid, M. V. Wickerhauser. Etudes des signaux vocaux par ondelettes de Malvar. *Quatorzieme Colloque GRETSI, Juan-Les-Pins*, Septembre 1993.
 - [105] W. Wijmans, P. Armbruster. Data Compression Techniques for Space Applications. Review of Current ESA/ESTEC Development, *Proceedings of DASIA '96*, Rome, May, 1996.

Lucrări despre soft

- [106] J. Buckheit, D. Donoho. WaveLab Architecture. *Preprint, Stanford University*, November 1995.
- [107] J. Buckheit, S. Chen, D. Donoho, I. M. Johnstone, J. Scargle. WaveLab Reference Manual. *Preprint, Stanford University*, December 1995.
- [108] J. Froment, S. Parrino. MegaWave 2 User's Modules Library. vol. I, vol. III, *Preprint CEREMADE, Univ. Paris Dauphine*, November 1994.
- [109] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery. Numerical Recipes in C. *Cambridge University Press*, 1995.

Noi lucrări ale autorului

- [110] A. Cubițchi. A New Speech Compression Algorithm. *International Workshop "Trends and Recent Achievements in Information Technology"* 16-18 May 2002, Cluj-Napoca Romania
- [111] A. Cubițchi. Une méthode nouvelle pour la compression de la parole, lucrare propusă la Buletinul Științific al Universității "Politehnica" din Timișoara, Seria ELECTRONICĂ și TELECOMUNICAȚII.