

CUANTIZAREA VECTORIALĂ A IMAGINILOR STATICE DESCOMPUSE IERARHIC PE SUBBENZI

Dr.ing. Laurian Mărgărit

Universitatea Politehnică București

Ing. Gabriel Nicula

Institutul de Cercetări în Informatică

Rezumat: Se prezintă o metodă nouă de codare a imaginilor statice, bazată pe cuantizarea vectorială a grupurilor de subbenzi de frecvență. Metoda folosește un predictor pentru zonele de detalii ale imaginii, folosind informația din detaliile de la o rezoluție mai scăzută.

Cuvinte cheie: cuantizare vectorială, subbenzi de frecvență, detalii, aproximare, raport semnal/zgomot.

1. Introducere

Imaginea este cel mai sugestiv mod de comunicare interumană. Cantitatea mare de informație conținută de imagini, pe de o parte, și banda relativ îngustă disponibilă pentru transmisiune, pe de altă parte, au determinat, în acest domeniu, un interes permanent pentru metodele de compresie. O dată cu progresele făcute în domeniul comunicațiilor digitale de mare viteză, atenția s-a îndreptat din ce în ce mai mult asupra serviciilor video (interactive sau de distribuție). Pentru unele din servicii, lățimea de bandă este relativ îngustă din considerente de cost (videotelefonul), pentru altele au sporit pretențiile în privința calității imaginii (televiziune de înaltă definiție), și, drept urmare, compresia de date rămâne în continuare de actualitate. În ultima perioadă, se înregistrează un interes crescut pentru metodele de compresie bazate pe codarea subbenzilor de frecvență ale imaginii, datorită faptului că pot fi ușor adaptate caracteristicilor de percepție vizuală și, în același timp, statisticii semnalului. Codarea pe subbenzi a semnalului a fost propusă pe la jumătatea anilor '70, pentru compresia semnalului vocal și abia după un deceniu și pentru imagini [1]. În momentul de față se experimentează algoritmi de compresie din această categorie în televiziunea de înaltă definiție și în transmisiunea pachetelor video.

Lucrarea de față tratează aspecte legate de codarea subbenzilor pentru imagini statice. Aceste metode de codare prezintă o serie de avantaje:

- posibilitatea înglobării modelelor de percepție vizuală umană;

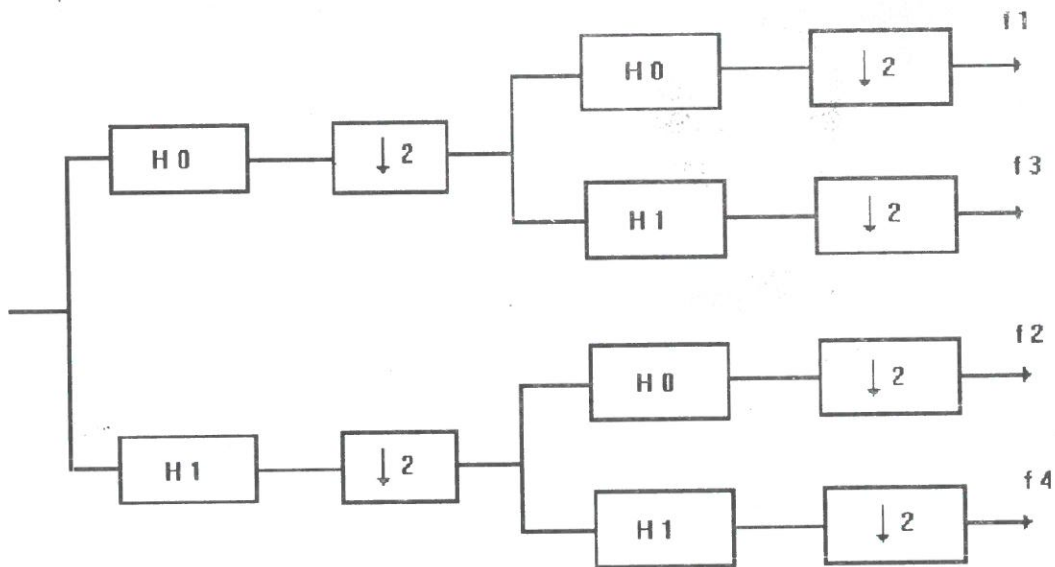
- decorelarea mai ușoară, atât la nivelul fiecărui subsemnal, cât și între subsemnale, prin codarea separată a subbenzilor;
- posibilitatea prelucrării în paralel a semnalelor din fiecare subbandă și la o frecvență mai mică.

Lucrarea este structurată astfel: secțiunea a doua trece în revistă metodele de reprezentare pe subbenzi ale imaginilor; în partea a treia se prezintă câteva considerații legate de codarea reprezentărilor pe subbenzi ale imaginilor, bazate pe cuantizarea vectorială; în secțiunea a patra este descrisă o metodă originală de codare a reprezentărilor ierarhice pe subbenzi ale imaginilor statice și sunt date rezultatele aplicării ei; metoda asigură factori de compresie (0.4361 – 0.5148 biți/pixel) ridicați la o calitate bună a imaginii reconstituite.

2. Analiza pe subbenzi a imaginilor

Un prim aspect de care trebuie ținut seama la analiza semnalului bidimensional discret este tipul rețelei pe care se face reprezentarea (ortogonală sau nu). În cazul utilizării unei rețele ortogonale, descompunerea semnalului poate fi separabilă sau neseparabilă. În cazul în care semnalul bidimensional este definit pe o rețea rectangulară și filtrele sunt separabile, metoda de descompunere este foarte simplă: filtrarea se realizează separat, adică mai întâi are loc descompunerea pe subbenzi pe o dimensiune, fiecare subsemnal obținut fiind descompus, apoi pe cea de-a doua dimensiune (figura 1). Aceasta este metoda folosită și în lucrarea de față.

Există mai multe posibilități de împărțire pe subbenzi a domeniului de frecvență bidimensional. Numărul de subbenzi în care se împarte spectrul depinde de aplicație și rezultă dintr-un compromis între eficiența compresiei și complexitatea schemei de codare, ambele crescând cu mărirea numărului subbenzilor. O primă posibilitate este de a împărți domeniul frecvență în subbenzi de arie egală. Acest mod este cunoscut sub numele de reprezentare uniformă (flat) [2]. O a doua posibilitate este de a face o împărțire ierarhică pe subbenzi prin iterarea descompunerii subbenzilor de joasă frecvență (în câte 4 benzi de arie egală) ca în figura 2. La fiecare iterație are loc împărțirea subbenzii B_{j_i} , care corespunde frecvențelor joase pe ambele dimensiuni, în 4 subbenzi. După decimare cu un factor de 2 pe fiecare dimensiune, subbanda de joasă frecvență $B_{j_{i+1}}$, corespunde unei imagini de rezoluție de două ori mai mică față de imaginea din banda B_{j_i} . Din acest punct de vedere,



Linii

Coloane

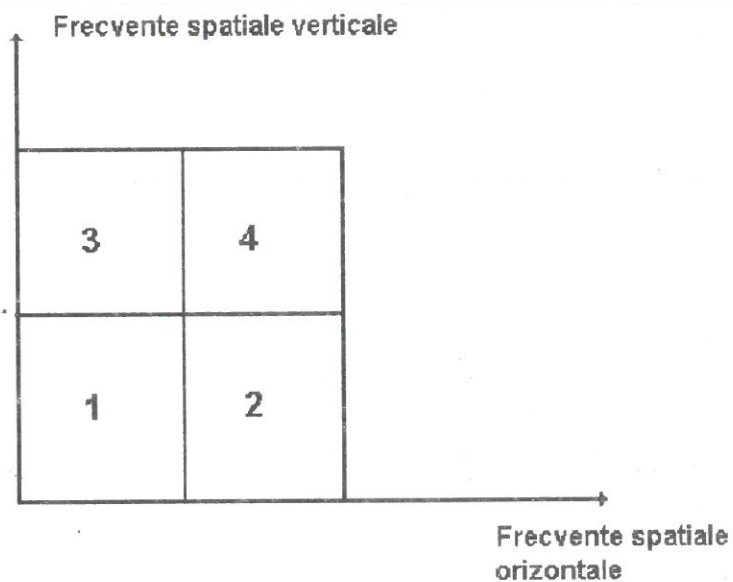


Figura 1.

reprezentările ierarhice pe subbenzi sunt echivalente cu descompunerile piramidale multirezoluție. Imaginea din banda B_j se numește **aproximația imaginii la rezoluție 2^{-j}** (cu convenția ca rezoluția imaginii original este 1). Celelalte subbenzi ale iterației i (D_x , D_y , D_{xy}) corespund semnalului detaliu al imaginii la aceeași rezoluție. D_x , de exemplu, conține frecvențele înalte pe orizontală și joase pe verticală.

1 - aproximare de ordin 3 (frecvențe joase pe ambele direcții)

2, 3, 4 - detalii de ordin 3 (frecvențe înalte pe cel puțin una din direcții)

2 - detalii pe orizontală D_x

3 - detalii pe verticală D_y

4 - detalii pe diagonală D_{xy}

3. Codarea pe subbenzi a imaginilor

Principiul metodelor de codare pe subbenzi este prezentat în figura 4. Semnalul de intrare este descompus pe subbenzi cu ajutorul unui banc de filtre. După codare, transmisie și decodare, procedeul de filtrare se realizează invers și originalul este reconstituit prin însumarea semnalelor din fiecare subbandă.

Prin aceasta se urmărește descompunerea imaginii în componente decorelate și obținerea unei compactări energetice. Codarea subbenzilor e necesară cel puțin din două motive. Primul e legat de repartizarea neuniformă a informației pe subbenzi: benzile de joasă frecvență conțin multă informație și necesită pentru transmisie mulți

biți/pixel, în timp ce informația conținută de benzile mai înalte se poate transmite cu mai puțin de un bit/pixel. Al doilea motiv este că există dependență statistică atât în cadrul subbenzii, cât și între subbenzi, cu toată reducerea datorată descompunerii. Codarea subbenzilor se poate face separat sau pe grupuri. Pentru exemplificare s-a calculat entropia subbenzilor pentru câteva imagini. Descompunerea este ierarhică pe 7 subbenzi și este realizată cu un banc de filtre separabile. Tabela 1 conține valorile entropiei înainte și după descompunere. Numeroțarea subbenzilor este conformă figurii 2b

Cuantizarea vectorială este, în momentul de față, una din cele mai răspândite metode de codare pentru semnalele uni- și multidimensionale datorită, atât performanțelor (calitate bună a semnalului reconstituit pentru factori de compresie ridicată), cât și modului extrem de simplu de decodare. Principiul metodei constă în a coda grupuri mici de eșantioane ale semnalului, constituite în vectori. Cuantizorul se poate defini ca o funcție [3],[4]:

$$VQ: A^k \rightarrow B$$

unde: k este dimensiunea vectorilor, iar A^k este spațiu euclidian. A este mulțimea valorilor (discrete ale) eșantioanelor intrării. B este un spațiu care conține vectori tot de dimensiune k , dar în număr mult mai mic (în comparație cu A^k) și poartă numele de tabelă de coduri. Aceștia sunt vectorii utilizați pentru reproducerea originalului. La codare, fiecare vector al intrării e comparat cu vectorii din B și înlocuit cu cel mai apropiat dintre aceștia, în metric -- de obicei -- euclidiană.

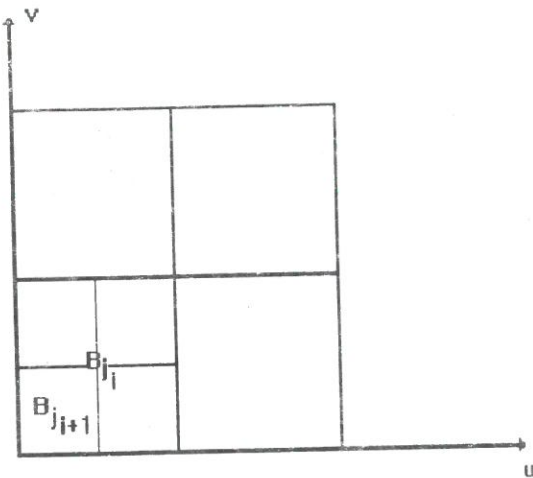


Figura 2a.

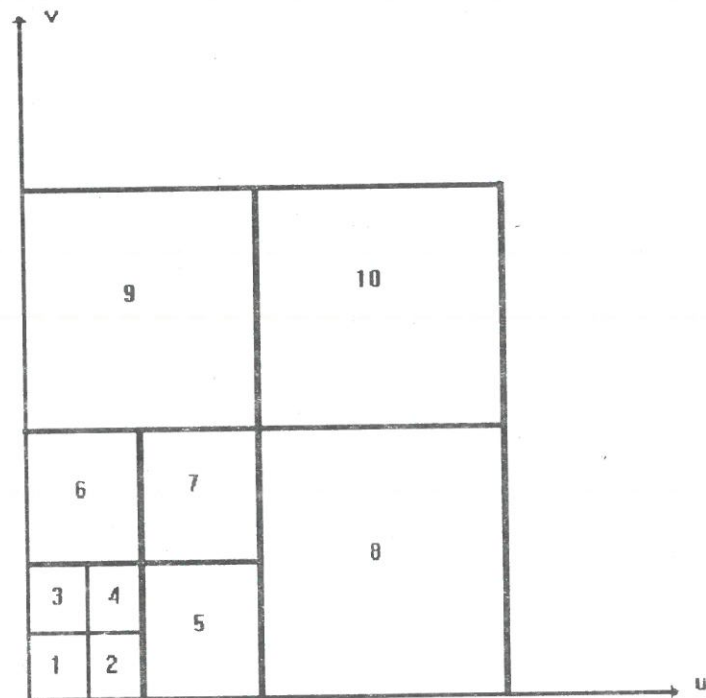


Figura 2b.

5, 6, 7 - detalii de ordin 2

5 - detalii pe orizontală Dx

6 - detalii pe verticală Dy

7 - detalii pe diagonală Dxy

8, 9, 10 - detalii de ordin 1

8 - detalii pe orizontală Dx

9 - detalii pe verticală Dy

10 - detalii pe diagonală Dxy

În figura 3 este ilustrată descompunerea ierarhică pe două niveluri (șapte subbenzi) ale unei imagini. Numărul de elemente al reprezentării este egal cu cel al imaginii original, deoarece reprezentarea este ortogonală.

Imagine	Entropia (bit / eşantion)							
	Original	Subbanda						
		1	2	3	4	5	6	7
Casa	6.9984	6.96	4.40	3.46	2.95	4.28	3.11	2.71
Sabine	7.1799	7.15	2.53	3.27	1.32	0.94	2.28	1.76
Vioara	6.7853	6.76	2.72	2.80	1.99	2.70	2.71	2.11
Sat	6.9702	6.92	4.40	3.69	2.88	4.00	3.11	2.02

Tabela 1. Codarea pe grupuri de subbenzi folosind cuantizarea vectorială

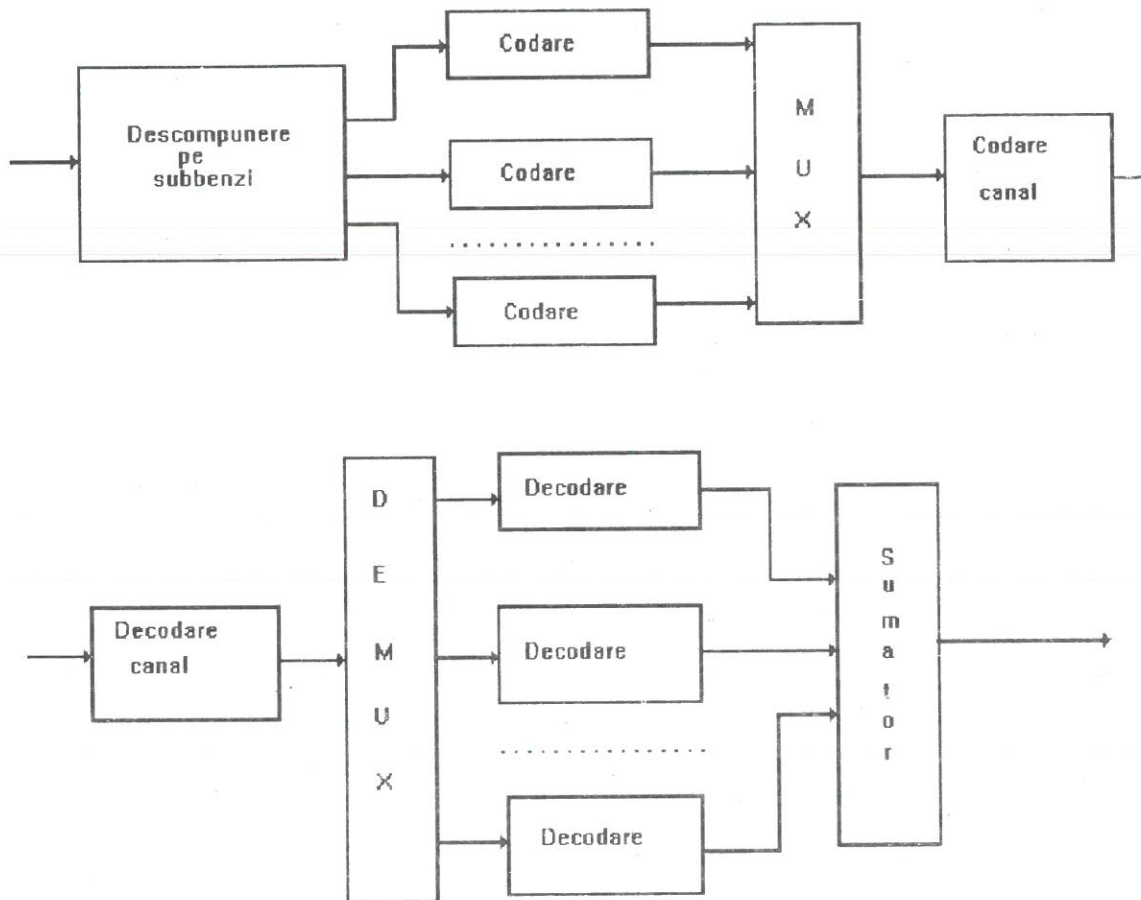


Figura 4.

Este un proces ireversibil prin care se introduc distorsiuni.

În loc de vector, se transmite adresa lui din tabela de coduri, pentru reconstituire fiind necesară, la recepție, o tabelă identică cu cea utilizată la codare. Generarea tabelii de coduri este o etapă, care, practic determină performanțele codării. Algoritmul cel mai des utilizat este cel conceput de Linde, Buzo și Gray (LBG) [5] (eventual cu unele perfecționări[6],[7]) și care reduce recursiv distorsiunea între o tabelă inițială de coduri și un set de vectori "de învățare". În acest scop se grupează vectorii de "învățare" în jurul celui mai apropiat vector din tabelă (în raport cu metrica folosită). Se atribuie apoi vectorilor tabelii valoarea centroidului grupului format în jurul fiecăruia. Procedul se reia cu tabela astfel modificată, până când distorsiunea globală scade sub un prag predefinit. Dezavantajele acestui algoritm sunt dependența soluției (tabela finală de

coduri) de tabela inițială, precum și posibilitatea stabilizării pe un minim local. Trebuie observat că, pentru a obține adaptarea tabelii de coduri la statistica semnalului, trebuie ca sursa să fie staționară, ceea ce presupune ca setul de vectori "de învățare" să fie foarte numeros.

Imaginea descompusă pe subbenzi de frecvență spațială se poate coda folosind metoda cuantizării vectoriale în fiecare subbandă în parte. Astfel este exploatată corelația spațială existentă în cadrul subbenzilor. O altă posibilitate este de a face cuantizarea vectorială interbenzi, în acest caz, într-un vector grupându-se elementele din fiecare subbandă, care au aceleași coordonate spațiale. Acest mod de codare ține seama de corelația spațială dintre subbenzi. Metoda a fost propusă în [8] pentru descompunerea uniformă pe subbenzi a imaginii.

4. Codarea reprezentărilor ierarhice pe subbenzi folosind metoda cuantizării vectoriale

Aplicarea ca atare a cuantizării vectoriale interbenzi reprezentărilor ierarhice nu este posibilă deoarece semnalul din subbenzi are dimensiuni diferite. Adaptarea tabeli de coduri la statistica semnalului din subbenzi este mai ușor de

detațiilor, se impune la codare o tratare diferită a semnalelor respective. Cuantizarea vectorială pe grupuri de subbenzi aflate pe același nivel de rezoluție pare mai adecvată, dacă se are în vedere și adaptarea codării (compresiei) la caracteristicile de percepție spațială ale sistemului vizual uman.

După cum se poate observa din figura 2.b, pe fiecare nivel al descompunerii ierarhice există trei subbenzi de detalii. Pe nivelurile de

Esantioane ale imaginii descompuse pe subbenzi

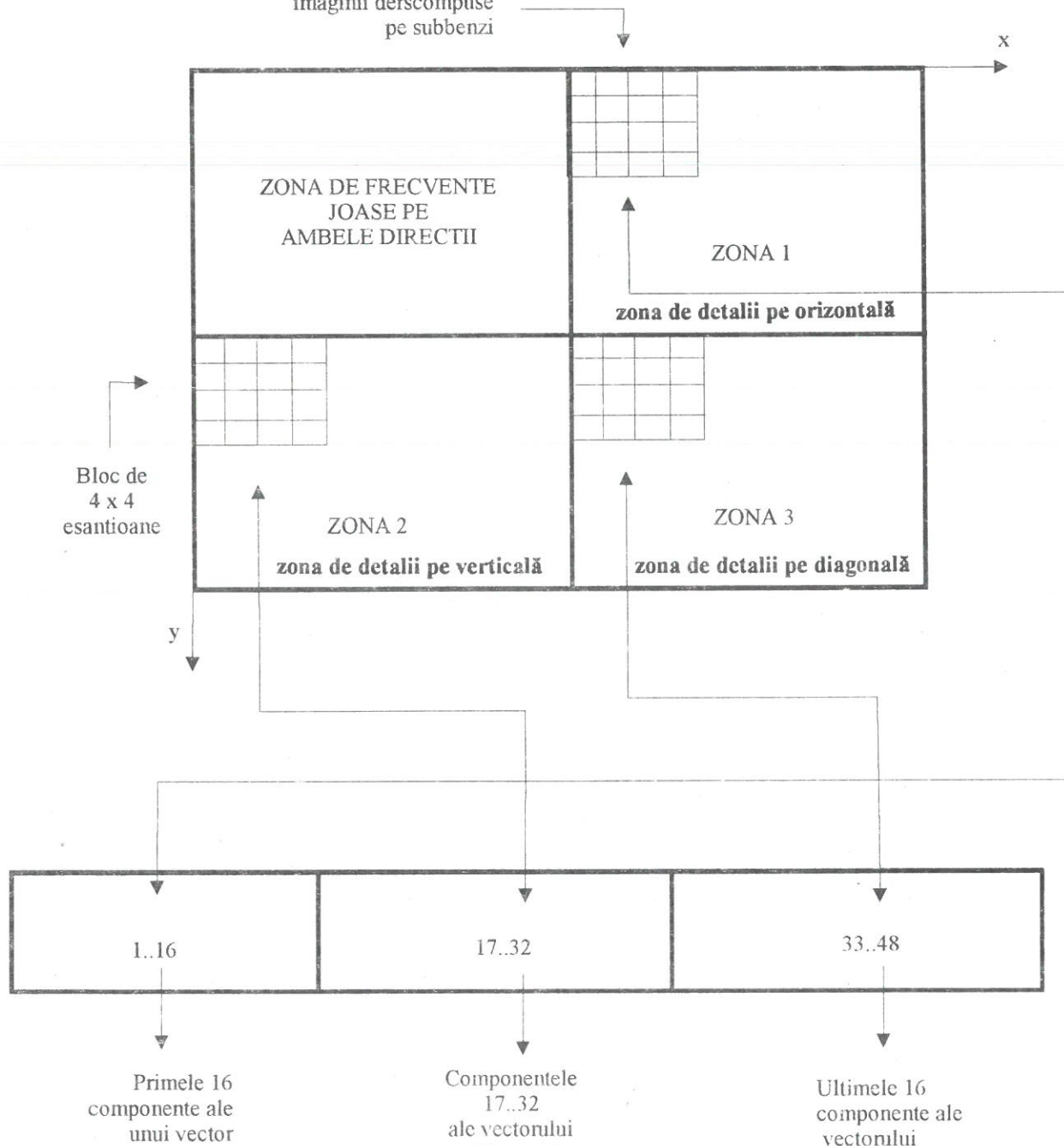


Figura 5.

realizat dacă vectorii conțin elemente din subbenzile aceleiași nivel de rezoluție al reprezentării. În plus, deoarece statistica aproximației - apropiată de aceea a imaginii original- este (de regulă) mult diferită de aceea a

rezoluție ridicată (subbenzile 5-7, 8-10 în figura 2.b), unde numărul de elemente este mare, codarea cu vectori 3-D este inefficientă. Valoarea mică a entropiei - relativă la original- (exemplificare pentru mai multe imagini cu statistici diferite în

tabela 1), arată că, după descompunerea în aceste subbenzi, rămâne suficientă corelația spațială. Pe aceste niveluri se poate spori eficiența codării prin creșterea numărului de componente al vectorului, constituându-l din blocuri care, în fiecare subbandă, au aceleași coordonate spațiale. În figura 5 este prezentat modul de formare al unui vector 48-D. Astfel cuantizarea vectorială devine inter-intrabandă. Chiar și așa, mai rămâne

2048 sau 1024 de coduri; în sfârșit, pentru subbenzile 8-10 se utilizează blocuri 48-D (blocuri de 4x4 pixeli) și o tabelă de 256 coduri. Generarea tabelelor s-a făcut folosind algoritmul LBG. La fiecare iterație, componentele vectorilor au fost prăguite în origine. Pragul utilizat diferă în funcție de orientare (mai mare pentru detaliile xy) și de nivelul de rezoluție (crește cu rezoluția spațială), ținându-se astfel seama și de caracteristicile

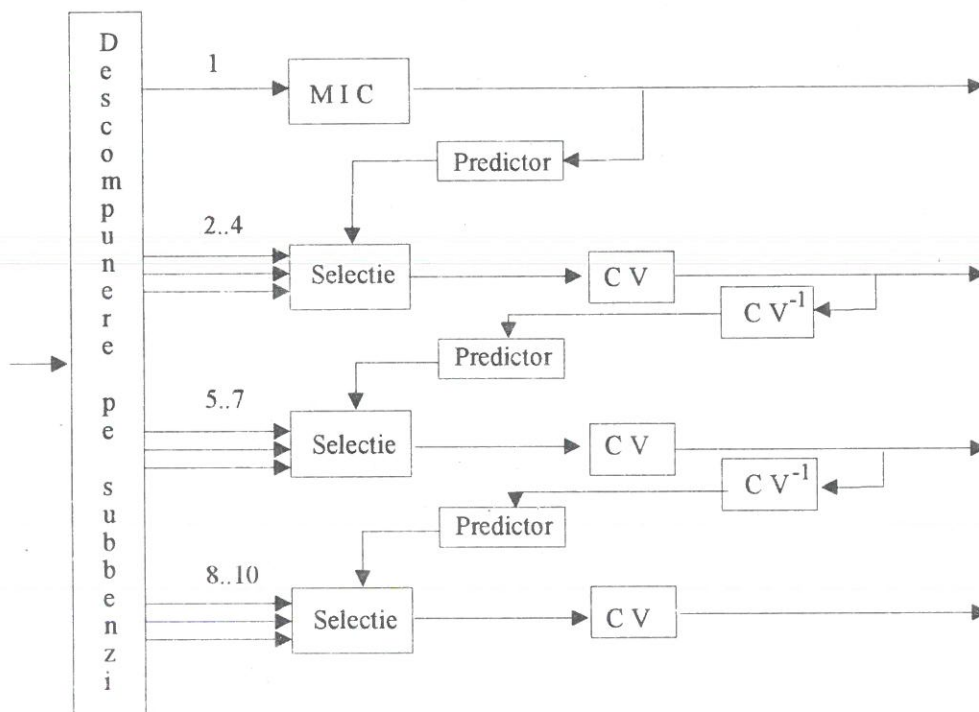


Figura 6. CV - cuantizor vectorial

corelație între informația de pe diferitele niveluri de rezoluție ale reprezentării. O parte din ea se poate elimina selectând informația ce urmează a fi codată pe baza celei de pe nivelul de rezoluție imediat inferior.

Pornind de la aceste considerații, se propune o metodă de codare pentru imagini reprezentate ierarhic pe subbenzi care este bazată pe cuantizarea vectorială. În figura 6 este dată schema bloc utilizată la codare, în cazul lucrului pe 10 subbenzi. Descompunerea se realizează cu bancuri de filtre unidimensionale, după algoritmul prezentat în figura 1.

Numerotarea subbenzilor s-a făcut conform figurii 2. Subbanda de joasă frecvență (numărul 1 -- aproximația) este codată cu MIC (modulația impulsurilor în cod) cu 8 bit/pixel. Deoarece această subbandă ocupă 1/64 din întreaga imagine, aportul ei la rata de bit este de 0.125 bit/pel. Restul de 9 subbenzi sunt codate prin cuantizare vectorială, grupate pe niveluri de rezoluție. Astfel, subbenzile 2-4 de pe nivelul de rezoluție scăzută sunt codate cu vectori 3-D utilizând o tabelă cu 256 sau 512 coduri; pentru subbenzile 5-7 se folosesc vectori 12-D (blocuri de 2x2 pixeli din fiecare subbandă) și o tabelă de

sistemului vizual uman. Lotul de antrenare utilizat cuprinde 8 imagini cu statistica variată. În scopul obținerii unei table simetrice, aceste imagini au fost rotite cu 90 grade și respectiv oglindite, rezultând 24 de imagini.

După cum se poate observa din schema bloc, pe fiecare nivel de rezoluție detaliile ce urmează a fi codate sunt selectate în funcție de valoarea (relativă la un prag) a detaliilor omoloage de pe nivelul de rezoluție imediat inferior. Pentru nivelul de rezoluție scăzută (subbenzile 2-4) selecția se realizează pe baza conținutului de detalii al aproximației (subbanda 1). Decizia se ia pe baza informației decodate (la codare) astfel încât decodorul poate localiza independent informația transmisă, fără o informație suplimentară. Acest mod de selecție a informației se bazează pe observația că, în general, muchiile conțin detalii de frecvențe scăzute (rezoluție spațială mică), dar și o serie de armonici în detaliile cu rezoluție spațială ridicată. În figura 7 este ilustrată selecția detaliilor pentru o imagine; cu negru sunt reprezentate pozițiile care se transmit (codate cu cuantizorul vectorial).

În tabela 2 sunt prezentate rezultatele obținute prin aplicarea metodei asupra unor

imagini din interiorul și din afara lotului de antrenare. În figura 8 sunt prezentate imaginile codate.

Tabela 2.

Imaginea codată	Rata de bit (biți / pixeli)	Raport semnal/zgomot (decibeli)
Sabine	0.507431	33.302476
Madonna	0.514816	33.365384
Lena	0.484055	30.467007
Harbour	0.467049	25.740779

Imaginile Sabine și Madonna sunt din lotul de antrenare, iar Lena și Harbour din afara acestuia. În tabela 3 este prezentată structura datelor transmise funcție de nivelul de rezoluție și ratele de bit obținute pentru imaginea LENA.

Tabela 3a. Structura datelor transmise funcție de nivelul de rezoluție și ratele de bit rezultate pentru imaginea LENA

LENA 1	Subbenzile			
	2-4	5-7	8-10	Global
Nr. total de vectori	4096	4096	4096	-
Nr. de vectori efectiv transmiși	3747	3293	1880	-
Rata de bit fără selecție detalii (bit/pel)	0.125	0.171 875	0.125	0.5468 75
Rata de bit cu selecție detalii (bit/pel)	0.114 349	0.138 1798	0.0573 7305	0.4361

Praguri utilizate pentru generarea tabelor de coduri:

- pentru subbenzile 2-4 :1.0 și 1.0 (256 coduri)
- pentru subbenzile 5-7 :1.5 și 1.5 (2048 coduri)
- pentru subbenzile 2-4 : 3.8 și 3.8 (256 coduri)

Praguri utilizate pentru predictor:

- pentru subbenzile 2-4 :1.0 și 1.0
- pentru subbenzile 5-7 :1.5 și 1.5
- pentru subbenzile 2-4 :3.8 și 3.8

Rata de bit totală : 0.4361 biți/pixel

Raport semnal/zgomot: 30.1556 decibeli

Tabela 3b. Structura datelor transmise funcție de nivelul de rezoluție și ratele de bit rezultate pentru imaginea LENA

LENA 2	Subbenzile			
	2-4	5-7	8-10	Global
Nr. total de vectori	4096	4096	4096	-
Nr. de vectori efectiv transmiși	3982	3599	2787	-
Rata de bit fără selecție detalii (bit/pel)	0.140 625	0.156 25	0.125	0.5468 75
Rata de bit cu selecție detalii (bit/pel)	0.136 7111 2	0.137 2909 5	0.0850 5254	0.4840 5457

Praguri utilizate pentru generarea tabelor de coduri:

- pentru subbenzile 2-4 :0.5 și 0.8 (512 coduri)
- pentru subbenzile 5-7 :1.0 și 1.2 (1024 coduri)
- pentru subbenzile 2-4 : 2.0 și 3.0 (256 coduri)

Praguri utilizate pentru predictor:

- pentru subbenzile 2-4 :0.5 și 0.8
- pentru subbenzile 5-7 :1.0 și 1.2
- pentru subbenzile 2-4 :2.0 și 3.0

Rata de bit totală : 0.484054565 biți/pixel

Raport semnal/zgomot: 30.46700652 decibeli

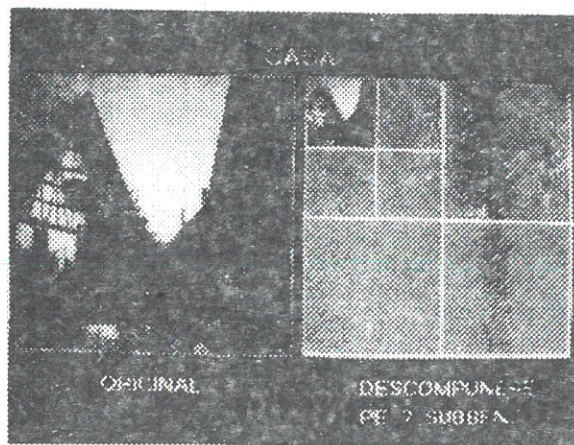


Figura 2

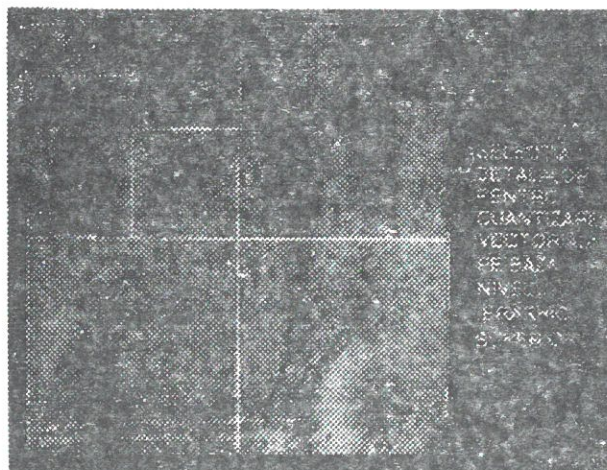


Figura 7

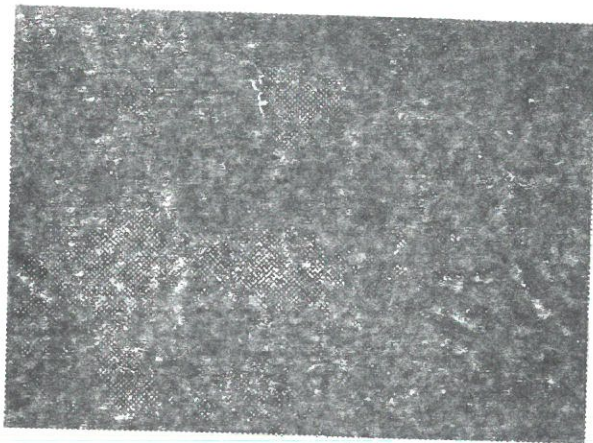


Figura 8.a

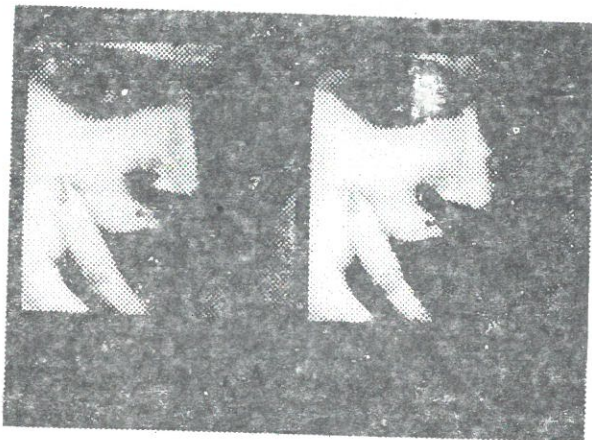


Figura 8.b



Figura 8.c



LENA
ORIGINAL

Figura 8.d



LENA 1
RATA DE BIT
0.4361

RAP S/ZG
30.1558

Figura 8.e



LENA 2
RATA DE BIT
0.4841

RAP S/ZG
30.4670

Figura 8.f

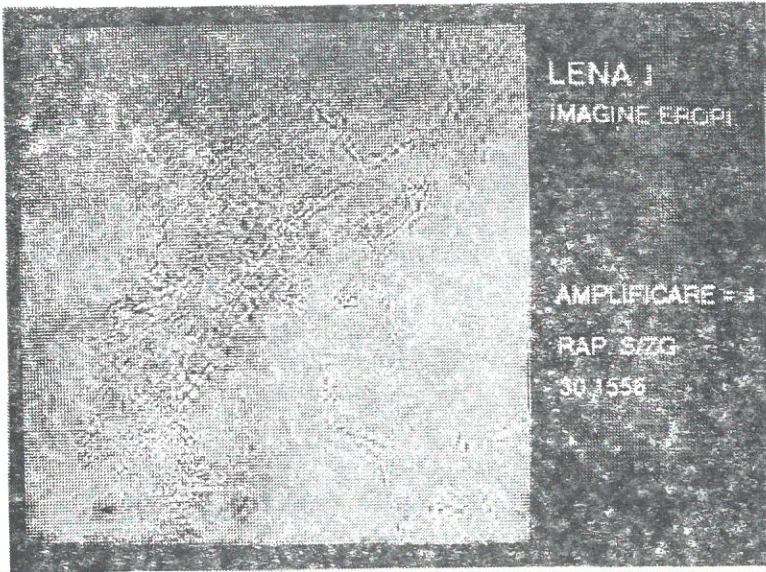


Figura 8.g

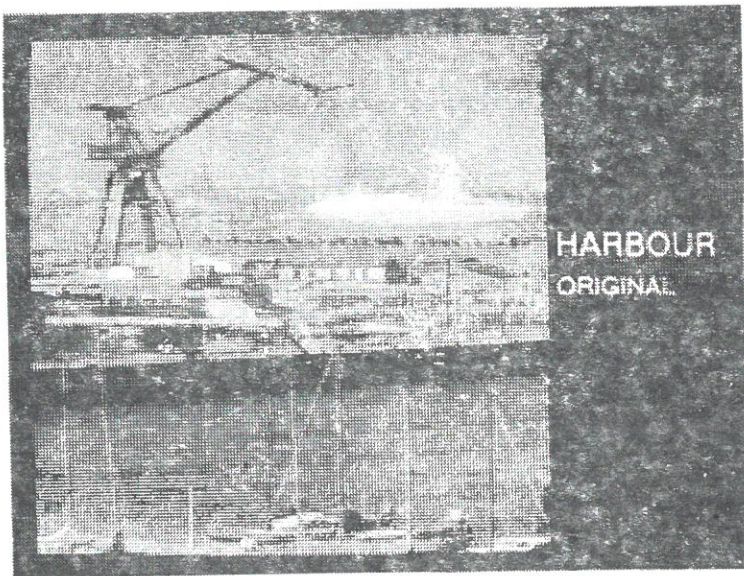


Figura 8.h

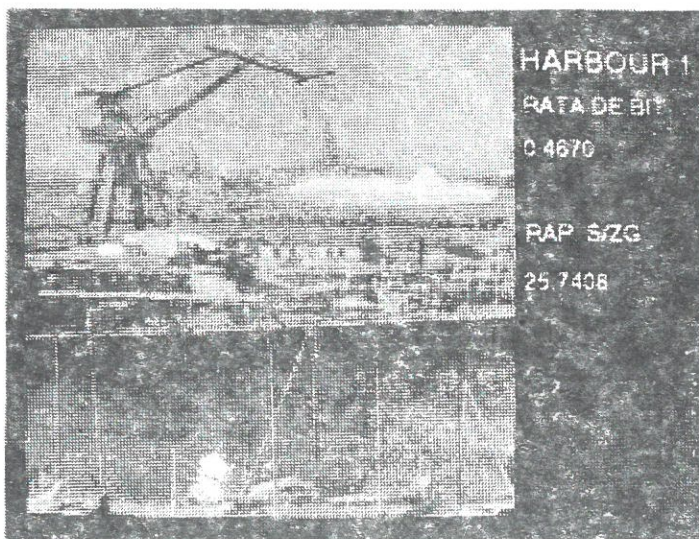


Figura 8.i

5. Concluzii

Se propune o metodă de codare interbandă, bazată pe cuantizarea vectorială a grupurilor de subbenzi, într-o descompunere ierarhică pe 3 niveluri de rezoluție. În figura 8 sunt prezentate imaginile original și codate din secvența de test și din afara ei. Pentru cele din afara lotului de antrenare sunt date și imaginile eroare. Această metodă, adaptată caracteristicilor de percepție vizuală, asigură factori de compresie mai buni de 16 valori experimentale cuprinse între 15.5396 și 18.3444, în condițiile obținerii unei calități bune a imaginii (raport s/z cu valori cuprinse între 30,155 dB și 33.3654 dB).

Bibliografie

1. WOODS, J.W., O'NEAL, S.D.: Subband Image Coding. In: *IEEE Transactions on ASSP*, Vol.34, No.5, 1986, pp.1278-1288.
2. FORCHHEIMER, R., KRONANDER, T.: Image Coding-From Waveforms to Animation. In: *IEEE Transactions on ASSP*, Vol.37, No.12, 1989, pp.2008-2023, 1989.
3. NASRABADI, N.M., KING, R.A.: Image Coding Using Vector Quantization: A Review". In: *IEEE Transactions on Communications*, Vol.36, No.8, 1988, pp.957-971.
4. SOLEIMANI, M.R., MORGERA, D.D.: A Fast MMSE Encoding Technique for VQ. In: *IEEE Transactions on Communications*, Vol.37, No.6, 1989, pp.656-658.
5. LINDE, Y.: An Algorithm for Vector Quantization Design. In: *IEEE Transactions on Communications*, Vol.28, No.1, 1980, pp.84-95.
6. CHANG, R.F.: A Fast Finite State Algorithm for VQ Design. In: *IEEE Transactions on Signal Processing*, Vol.40, No.1, 1980, pp.221-225.
7. PALIWAL, K.K., RAMASUBRAMANIAN, V.: Effect of Ordering the Codebook on the Efficiency of the Partial Distance Search Algorithm for VQ. In: *IEEE Transactions on Communications*, Vol.37, No.5, 1989, pp.538-540.
8. WESTERINK, P.H.: Subband Coding of Images Using Vector Quantization. In: *IEEE Transactions on Communications*, Vol.36, No.6, 1988, pp.713-719.