

## Algoritmul de compresie DCT

Prof. dr. Ion IVAN, prep. Daniel VERNIȘ  
Catedra de Informatică Economică, A.S.E. București

*La începutul anilor '80, cercetarea asupra unor noi tehnici de compresie specifice imaginilor promitea să surclaseze cu mult tehnicile de compresie cunoscute pînă atunci. În anul 1991 grupul JPEG dă publicității un standard pentru compresia de imagini care oferea, pentru imagini destul de complexe, rate de compresie cu mult mai bune decît algoritmi clasici, avînd un anumit grad de pierdere a informației.*

*Cuvinte cheie : compresie de date, imagini, JPEG, algoritmi*

### 1. Introducere

Scopul JPEG a fost să dezvolte o metodă de compresie a imaginilor cu ton continuu. În 1987 grupul JPEG a desfășurat un proces de selecție bazat pe o apreciere directă a calității unei imagini compresate și a restrîns cele 12 metode propuse initial la 3. În acest timp au fost formate 3 grupuri de lucru pentru îmbunătățirea acestor metode iar în ianuarie 1988, pe baza unui proces de selecție mult mai riguros, a fost dezvăluită o nouă propunere bazată pe 8X8 DCT, care avea cea mai bună calitate a imaginii.

În momentul selectării ei, metoda bazată pe DCT a fost doar potențial definită pe cîteva dintre modurile de operare. Între anii 1988 și 1990 grupul JPEG și-a asumat sarcina de a defini, documenta, simula, testa și valida această metodă.

Standardul propus conține cele patru moduri de operare identificate anterior. Pentru fiecare mod, au fost specificate unul sau mai multe procedee de compresie și decompresie. Aceste procedee diferă de la un mod la altul în funcție de precizia eșantionului din imaginea sursă cu care pot opera sau de entropia metodei de codificare folosită. Nu există restricții cu privire la faptul că orice implementare trebuie să includă atît codorul cît și decodorul. Multe aplicații au doar unul din cele două. Cele patru moduri de operare și variantele lor de compresie au rezultat din scopul JPEG de a fi generic și din diversitatea de formate de imagine care se întîlnesc în aplicații.

### 2. Pașii de prelucrare ai compresiei bazate pe DCT

Figurile 1 și 2 prezintă pașii procesării care sunt nucleul modurilor de operare bazate pe DCT. Aceste figuri ilustrează cazul special al unei compresii de imagine în tonuri de gri. Compresia DCT poate fi asemănată cu compresia unui șir de 8X8 blocuri de eșantioane dintr-o imagine în nuanțe de gri. Compresia imaginilor color poate fi aproximată cu compresia unui grup de imagini în nuanțe de gri. Compresia se poate face independent pentru fiecare imagine în parte. O altă modalitate este prelucrarea imaginilor pe blocuri de eșantioane avînd dimensiunea 8\*8, simultan.

Pentru tipurile de compresie care operează în mod secvențial, diagramele simplificate indică cum funcționează compresia unei singure componente. Fiecare bloc 8X8 de intrare, își urmează drumul prin toți pașii de procesare și produce o ieșire în formă comprimată în interiorul unui șir de date.

### 3. 8X8 FDCT și IDCT

La intrarea în procesul de compresie, eșantioanele din imaginea sursă sunt grupate în blocuri de 8X8, transformate din întregi fără semn aparținînd intervalului  $[0, 2^p - 1]$  în întregi cu semn în intervalul  $[-2^{p-1}, 2^{p-1} - 1]$  și introduși în **Forward\_DCT (FDCT)**.

La ieșire, **Inverse\_DCT (IDCT)** se obțin blocuri de dimensiuni 8X8 utilizate pentru a reconstitui imaginea inițială.

Următoarele ecuații reprezintă definițiile matematice ale FDCT și IDCT :

$$F(u, v) = \frac{1}{4} C(u)C(v) \left[ \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) * \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \right]$$

$$f(x, y) = \frac{1}{4} \left[ \sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v) F(u, v) * \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \right]$$

unde  $C(u), C(v) = \frac{1}{\sqrt{2}}$  pentru  $u, v = 0$  și  $C(u), C(v) = 0$  în rest.

DCT este asemănător cu Transformata Fourier Discretă DFT. Se poate deduce compresia bazată pe DCT prin analogie cu FDCT ca analizor armonic sau IDCT ca sintetizator armonic. Orice bloc de 8X8 din eşantioanele imaginii sursă este de fapt un semnal discret de 64 de puncte care este o funcție de două argumente  $x$  și  $y$ . FDCT consideră acest semnal ca o intrare și îl descompune în 64 de semnale ortogonale de bază. Fiecare conține una din cele 64 de frecvențe spațiale bidimensionale unice, care cuprind spectrul semnalului de intrare. Ieșirea FDCT-ului este setul de 64 de amplitudini bazate pe semnal sau coeficienții DCT ale căror valori sunt unic determinate de

către semnalul de intrare de 64 de pixeli. Valorile coeficienților DCT pot fi privite drept cantități relative de frecvențe spațiale bidimensionale conținute în semnalul de intrare de 64 de pixeli.

Coeficientul cu frecvența 0 în ambele dimensiuni este numit coeficient DC, iar ceilalți 63 de coeficienți sunt numiți coeficienți AC. Deoarece valorile eşantioanelor variază foarte încet de la un punct la altul de-a lungul unei imagini, pașii algoritmului FDCT reprezintă baza pentru obținerea rezultatelor comprimate prin concentrarea celor mai multe semnale în frecvențe spațiale reduse. Pentru un bloc 8X8 dintr-o imagine sursă tipică, cele mai multe dintre frecvențele spațiale au amplitudinea zero sau aproape zero și nu este necesară codificarea lor.

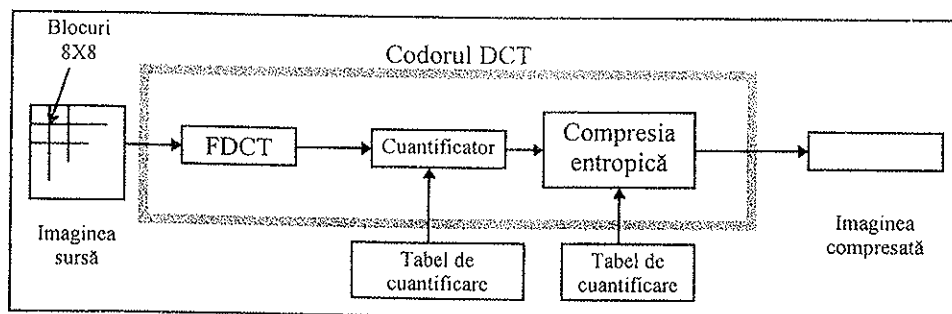


Fig. 1. Pașii de procesare în cadrul compresiei bazate pe DCT

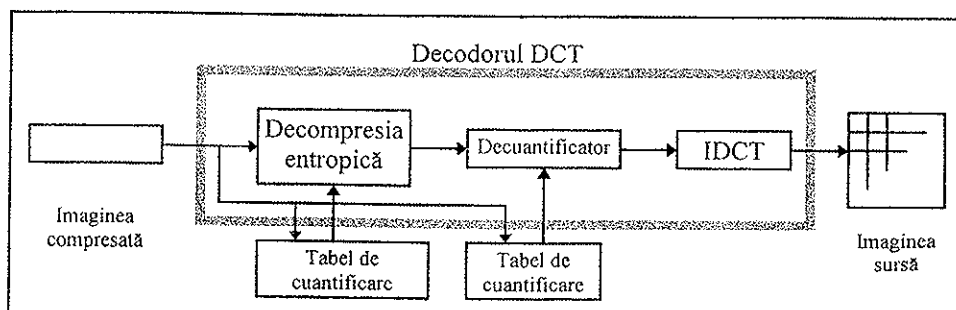


Fig. 2. Pașii de procesare în cadrul decompresiei bazate pe DCT

La decompresie, IDCT parcurge pașii de procesare în ordine inversă. Se iau cei 64 de

coeficienți DCT, care în acest moment au fost cuantificați, și recostituie o ieșire de 64

de puncte sub forma unui semnal prin însumarea semnalelor de bază. Matematic, DCT este o punere în corespondență 1 la 1 a unor vectori de 64 de pixeli între imagine și domeniul de frecvențe. Dacă FDCT și IDCT ar putea fi calculate cu o precizie perfectă și dacă coeficienții DCT nu ar fi cuantificați ca în descrierea următoare, semnalul original de 64 de puncte ar putea fi reconstituit exact. În principiu, DCT nu introduce nici o pierdere în eșantioanele din imaginea sursă, el reușeșind doar să producă o transformare a valorilor coeficienților DCT de la o reprezentare în dubla precizie la una în simplă precizie.

Cîteva proprietăți ale implementărilor practice pentru FDCT și IDCT ridică problema: ce anume ar trebui să se ceară de la standardul JPEG. O proprietate fundamentală este că ecuațiile FDCT și IDCT conțin funcții transcendente. Nici o implementare fizică nu le poate calcula cu o precizie perfectă. Datorită importanței aplicațiilor bazate pe DCT și a relațiilor lui ce DFT, au fost prezentați mulți algoritmi prin care FDCT și IDCT pot fi aproximativ calculate. Cercetările din domeniul algoritmilor DCT rapizi continuă și nu există un singur algoritm care să fie optim pentru toate implementările. Ce este optim în software pentru un procesor pentru scopuri generale este puțin probabil să fie optimal în firmware pentru un DSP programabil și sigur este suboptimal pentru un VLSI dedicat.

Chiar în lumina preciziei finite a intrărilor și ieșirilor DCT, implementări ale aceluiași algoritm, FDCT sau IDCT, care sunt proiectate independent, cu toate că acestea diferă minimal, doar în modul de reprezentare a termenilor cosinus, a rezultatelor intermediare sau în felul în care se rotunjesc valorile fracționare, vor produce ieșiri ușor diferite pentru aceleași intrări.

Pentru a conserva libertatea de inovare sau modificare a implementărilor, JPEG nu specifică un algoritm unic de calcul a FDCT sau IDCT în standardul propus. Acest lucru împiedică oarecum, deoarece două codoare sau decodoare diferite vor produce în general ieșiri diferite pentru intrări identice.

Standardul JPEG rezolvă această problemă prin specificarea unui test de precizie, ca parte componentă a codoarelor sau decodoarelor bazate pe DCT; acest lucru pentru a se asigura împotriva funcțiilor de bază cosinus calculate imprecis, ceea ce ar duce la degradarea calității imaginii.

Pentru orice model bazat pe DCT, propunerea JPEG specifică modele de compresie/decompresie separate pentru imagini cu eșantioane de 8 respectiv 12 biți pe componentă. Modelele pe 12 biți sunt folosite pentru imagini care necesită resurse de calcul mai mari pentru a se putea obține precizia dorită a FDCT sau IDCT. Imaginile cu alte precizii de eșantion pot fi potrivite prin modele pe 8 sau pe 12 biți, dar acest lucru trebuie făcut în afara standardului JPEG. Este responsabilitatea aplicațiilor să decidă cum să includă sau să încadreze un eșantion de 6 biți într-o interfață a cadrului de lucru de 8 biți, cum să despacheteze ieșirile decodorului și cum să codifice informațiile necesare.

#### 4. Cuantificarea

După ce ies din FDCT, cei 64 coeficienți DCT sunt cuantificați uniform, folosind tabela de cuantificare care conține 64 de elemente, specificate de utilizator sau aplicate ca o intrare pentru algoritmul de compresie. Fiecare element poate avea o valoare întreagă de la 1 la 255, ceea ce specifică mărimea pasului cuantificatorului pentru coeficientul DCT corespunzător. Scopul cuantificării este obținerea unei compresii mai bune prin reprezentarea coeficienților DCT cu o precizie mult mai mare decât este necesar pentru a obține calitatea dorită a imaginii. Scopul pasului de procesare este ignorarea informației care nu este semnificativă vizual.

Cuantificarea este o corespondență *mulți la unu* și prin urmare fundamental ea va înregistra pierderi. Aceasta este principala sursă de pierderi din algoritmi de compresie bazați pe DCT. Cuantificarea se realizează prin împărțirea fiecărui coeficient DCT la pasul de cuantificare corespunzător, urmată de rotunjirea la cel mai apropiat întreg:

$$F^Q(u, v) = \text{Parte întreaga} \left( \frac{F(u, v)}{Q(u, v)} \right).$$

Valoarea de ieșire este normalizată cu ajutorul mărimii pasului de cuantificare.

Decuantificarea este funcția inversă, ceea ce, în acest caz, înseamnă că normalizarea este înlăturată prin multiplicarea cu mărimea pasului de cuantificare, ceea ce va întoarce un rezultat a cărui reprezentare este folosită ca intrare pentru IDCT.

$$F^Q(u, v) = F(u, v) * Q(u, v)$$

Cînd scopul este obținerea unei cît mai mari compresii a imaginii, fără modificări vizibile, fiecare mărime a pasului trebuie, în mod ideal, aleasă astfel încît să minimizeze posibilitatea de eroare a funcției cosinus corespunzătoare. Pragurile sunt funcție de caracteristicile imaginii sursă, caracteristicile de afișare și distanța de vizualizare.

Pentru aplicațiile în care aceste variabile pot fi în mod rezonabil bine definite, experimentele psihovizuale pot fi realizate pentru a determina cele mai bune praguri. Experimentul descris de H. Lohscheller în 1984, a condus la un set de tabele de cuantificare pentru imaginile și display-urile CCIR-607. Acestea au fost folosite în mod experimental de către grupul JPEG și au apărut în standardele ISO ca o problemă de informare, nu ca o cerință.

### 5. Compresia DC și secvențele în zig-zag

După cuantificare, coeficientul DC este tratat diferit față de ceilalți 63 coeficienți

AC. Coeficientul DC este o măsură a valorii medii a celor 64 eșantioane ale imaginii. Deoarece de obicei există o puternică corelație între coeficienții DC ai blocurilor adiacente de 8X8, coeficientul DC cuantificat este codificat prin diferență față de termenul DC al blocului precedent, așa cum apare în figura 3. Acest tratament special merită făcut, din moment ce coeficienții DC conțin în mod frecvent o fracțiune semnificativă din energia totală a imaginii.

În cele din urmă toți coeficienții cuantificați sunt ordonați într-o secvență de zig-zag, de asemenea prezentată în figura 3. Ordonarea ajută la ușurarea calculului entropiei prin plasarea coeficienților cu frecvențe reduse înaintea celor cu frecvență mare.

### 6. Codificarea entropică

Ultimul pas din cadrul procesului de codificare bazat pe DCT este codificarea entropică. Acest pas produce pierderi adiționale prin codificarea coeficienților DCT mult mai compact, bazîndu-se pe caracteristicile lor statistice. Standardul JPEG specifică două metode de compresie entropică: compresia Huffman și compresia aritmetică.

Modelul secvențial de bază folosește compresia Huffman dar tipologii de compresie cu ambele metode sunt specificate pentru toate modurile de operare. Este foarte util să se considere codificarea entropică ca un proces în 2 pași.

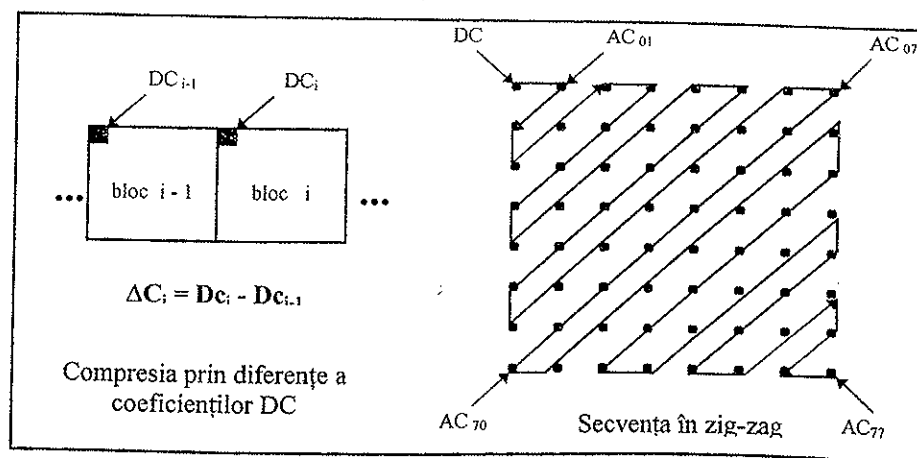


Fig. 3. Pregătirea coeficienților DCT pentru codificarea entropică

a valorii  
imaginii.  
puternică  
locurilor  
cuanti-  
față de  
așa cum  
special  
nții DC  
semni-

ntificați  
zag, de  
lonarea  
ei prin  
reduse

ui de  
ficarea  
adițio-  
DCT  
tracte-  
JPEG  
entropia

sește  
com-  
ficate  
oarte  
că ca

Primul pas convertește secvența în zig-zag a coeficienților cuantificați într-o secvență intermediară de simboluri. Al doilea pas convertește simbolurile în șiruri de date în care simbolurile nu mai au granițe identificabile extern. Forma și definiția simbolurilor intermediare sunt dependente atât de modul de operare bazat pe DCT cât și de metoda de codificare entropică.

Compresia Huffman cere ca unul sau mai multe seturi de tabele de coduri Huffman să fie specificate de aplicație. Aceleași tabele folosite la compresia unei imagini sunt folosite și la decompresia ei. Tabelele pot fi definite și folosite în interiorul aplicației ca tabele implicite sau calculate în mod specific pentru o imagine dată într-un pas anterior compresiei pe baza principiilor cunoscute de la algoritmul Huffman. Asemenea alegere este lăsată la dispoziția aplicațiilor care implementează standardul. JPEG specifică faptul că nu se cere un set predefinit de tabele Huffman.

Dimpotrivă, metoda compresiei aritmetice nu necesită introducerea exterioară a unor tabele deoarece este capabilă să se adapteze statisticilor imaginii pe parcursul compresiei. Dacă se dorește, tabelele de condiționare statistică pot fi introduse pentru o ușoară îmbunătățire a eficienței, dar acest lucru nu este impus. Compresia aritmetică produce o compresie cu 5 - 10% mai bună decât cea realizată de Huffman pentru multe imagini pe care a fost testat JPEG-ul.

O parte a programatorilor cred că este mai complexă decât Huffman pentru anumite implementări. Complexitatea s-a dovedit a fi o metrică ne semnificativă pentru metoda de comparare a compresiilor.

Dacă unica diferență între cele două modele JPEG este codificarea entropică, atunci transcodarea între cele două este posibilă prin simpla decodare entropică cu una dintre metode urmată de recodarea entropică folosind cealaltă metodă.

## 7. Concluzii

Pentru imaginile color cu scene relativ complexe, toate modelele de operare bazate pe DCT produc în mod tipic nivelele de

calitate prezentate mai jos. Aceste modele sunt doar niște linii de ghidare, calitatea și compresia putând varia semnificativ în concordanță cu scenele și caracteristicile imaginilor sursă. Unitățile biți/pixel reprezintă numărul total de biți din imaginea compresată, incluzând componentele cromatice. Clasificarea după nivelele de calitate este următoarea :

- 0,25 - 0,5 biți/pixel: calitate moderată până la bună, suficientă pentru unele aplicații ;
- 0,5 - 0,75 biți/pixel: calitate de la bună la foarte bună, suficientă pentru multe aplicații
- 0,75 - 1,5 biți/pixel: calitate excelentă pentru cele mai multe aplicații ;
- 1,5 - 2,0 biți/pixel: de obicei nu se disting diferențele față de original, suficientă pentru aplicațiile cele mai pretențioase.

Compresia JPEG pentru tonuri continue nu este o soluție care va rezolva toate problemele ridicate de imaginile digitale și de integrarea lor totală în aplicațiile care în final le vor folosi.

De exemplu, dacă două aplicații nu pot schimba o imagine necompresată pentru că folosesc spații de culoare incompatibile, dimensiuni incompatibile, atunci o metodă obișnuită de compresie nu va fi soluția acestei probleme.

Pentru aplicațiile care au nevoie de resurse mari de stocare sau transmisie evaluarea tehnică completă, testarea, selecția, validarea și munca de documentare pe care membrii grupului JPEG au depus-o a produs un standard aprobat internațional care va rezista tuturor testelor de calitate și timp. Odată cu apariția tot mai multor aplicații grafice, implementările pe sisteme deschise sau chiar rețele a standardului JPEG nu vor mai fi privite ca un alt tip de date și vor fi ceva normal.

## Bibliografie

- [Be90] Bell T.C., Cleary J.G., Witten I.N. Image Compression, Prentice Hall Englewood Cliffs, N.J. 1990
- [Dc95] \*\*\*  
Data Compression Conference, Snowbird, UT, 28-30 martie 1995, ISBN 0-8186-7012-6, Editura IEEE Compu Press

- [Iv95] I.Ivan, D. Verniș  
Analiza comparată a algoritmilor de compresie date, PC World, nr.12, Decembrie 1995
- [Iv95] I.Ivan, D. Verniș  
Comprimarea datelor, PC Report, nr. 9, Septembrie 1996
- [Iv96] I.Ivan, D. Verniș  
Evaluarea seturilor de date destinate compresării, Revista de statistică, nr.1, Ianuarie 1996
- [Iv98] I.Ivan, D. Verniș  
Compresia de date, Editura CISON, 1998
- [Ve96] Daniel Verniș  
Comunicarea "Algoritmi de compresie pentru șiruri de date numerice", Cercul de Structuri de Date, A.S.E. Bucuresti, 1996.
- [Ve97] Daniel Verniș  
Implementarea algoritmilor de compresie folosind rețele neuronale, Studiu ASE, 1997.