

COMPRESIA IMAGINILOR ȘI TRANSFORMATĂ DISCRETĂ COSINUS

Gheorghe PURCI, Ion CAPCANARI, Gh. SOROCHIN

Universitatea Tehnică a Moldovei

Abstract: This paper describe the mathematical equation of the Discrete Cosine Transform and its uses with image compression, applied to JPEG quantification with 50% quality loss of initial image and decompression using inverse Discrete Cosine Transform.

Cuvinte cheie: Imagine, compresie, cuantizare, Transformata Discreta Cosinus, bloc

Introducere

Deoarece necesitățile și dependența noastră de calculatoare continuă să crească, avem nevoie de modalități eficiente de stocare a unor cantități mari de date. Spre exemplu, cineva are o pagină web sau un registru online unde sunt amplasate zeci sau sute de imagini. Evident că o cantitate mare de imagini va avea nevoie și de un spațiu de stocare respectiv. Deci apare problema insuficienței spațiului de stocare. Pentru soluționarea acestei probleme este nevoie de a comprima imaginile respective. La moment există o serie de metode disponibile pentru soluționarea problemei de comprimare. Aceste metode se împart în două categorii generale: lossless și lossy. Metoda de comprimare Transformata Discreta Cosinus (TDC) prezentată în această lucrare face parte din categoria „lossy”. TDC funcționează prin separarea imaginilor în părți diferite ale frecvențelor. În timpul etapei de cuantizare, frecvențele mai puțin importante sunt excluse. Iar frecvențele mai mult importante sunt folosite la decompresie. Ca rezultat imaginile reconstruite conțin anumite distorsiuni, aceste distorsiuni pot fi ajustate în timpul etapei de comprimare. Această metodă se folosește atât pentru imagini color cât și pentru cele alb-negru (binare).

Transformata Discreta Cosinus

Algoritmul procesului de compresie a imaginii cu Transformata Discreta Cosinus este urmatorul:

1. Imaginea este împărțită în blocuri a câte 8x8 pixeli.
2. De la stînga la dreapta, de sus în jos TDC se aplică pentru fiecare bloc.
3. Fiecare bloc va fi comprimat prin cuantizare.
4. Matricea de blocuri comprimate constituie imaginea ce va fi stocată într-o cantitate redusă de memorie.
5. La decompresia imaginii se aplică Inversa Transformata Discreta Cosinus (ITDC).

Ecuția TDC calculează imaginea conform intrărilor (i,j).

$$D(i, j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p(x, y) \cos \left[\frac{(2x+1)i\pi}{2N} \right] \cos \left[\frac{(2y+1)j\pi}{2N} \right] \quad (1.1)$$

(1.2)

Unde: $p(x, y)$ este un element al imaginii (un pixel) reprezentat de matricea p ; N este dimensiunea matricii și este egal cu 8. Prin urmare $D(i,j)$ va arata ca în ecuația (1.3):

$$D(i, j) = \frac{1}{4} C(i)C(j) \sum_{x=0}^7 \sum_{y=0}^7 p(x, y) \cos \left[\frac{(2x+1)i\pi}{16} \right] \cos \left[\frac{(2y+1)j\pi}{16} \right] \quad (1.3)$$

Deoarece TDC utilizează funcții cosinus, matricea rezultată va depinde de frecvențele orizontale, diagonale și verticale.

Pentru a obține forma de matrice a ecuației (1.1), vom folosi ecuația (1.4).

$$DCT(i, j) = \begin{cases} \frac{1}{\sqrt{N}} & \text{if } i = 0 \\ \sqrt{\frac{2}{N}} \cos \left[\frac{(2j+1)i\pi}{2N} \right] & \text{if } i > 0 \end{cases} \quad (1.4)$$

Pentru un bloc 8x8 pixeli în baza ecuației (1.4) a fost obținută următoarea matrice:

$$DCT = \begin{pmatrix} 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 & 0.3536 \\ 0.4904 & 0.4157 & 0.2778 & 0.0975 & -0.0975 & -0.2778 & -0.4157 & -0.4904 \\ 0.4619 & 0.1913 & -0.1913 & -0.4619 & -0.4619 & -0.1913 & 0.1913 & 0.4619 \\ 0.4157 & -0.0975 & -0.4904 & -0.2778 & 0.2778 & 0.4904 & 0.0975 & -0.4157 \\ 0.3536 & -0.3536 & -0.3536 & 0.3536 & 0.3536 & -0.3536 & -0.3536 & 0.3536 \\ 0.2778 & -0.4904 & 0.0975 & 0.4157 & -0.4157 & -0.0975 & 0.4904 & -0.2778 \\ 0.1913 & -0.4619 & 0.4619 & -0.1913 & -0.1913 & 0.4619 & -0.4619 & 0.1913 \\ 0.0975 & -0.2778 & 0.4157 & -0.4904 & 0.4904 & -0.4157 & 0.2778 & -0.0975 \end{pmatrix}$$

Sa luat în considerare faptul că valorile pixelilor unei game de alb-negru sunt de la 0 la 255. Unui pixel i se atribuie o nuanță gri care poate avea valoarea de la 0 la 255. Culoarea neagră este reprezentată de valoarea 0, iar culoarea alb pur, de valoarea 255. În dependență de rezoluția imaginii, pot fi de la sute pînă la mii de blocuri 8x8 pixeli.

TDC este proiectat pentru valori a unui pixel cu diapazonul $-128 \div 127$. Rezultă că blocul imaginii originale va fi adunat cu -128 conform ecuației (1.5).

$$M(i, j) = \sum_{i=0}^7 \sum_{j=0}^7 \text{ImgOriginala}(i, j) - 128 \quad (1.5)$$

$$\text{Original} = \begin{pmatrix} 168 & 177 & 197 & 214 & 206 & 171 & 135 & 107 \\ 189 & 205 & 231 & 250 & 241 & 209 & 167 & 114 \\ 212 & 234 & 255 & 255 & 255 & 247 & 200 & 138 \\ 255 & 248 & 255 & 255 & 255 & 255 & 219 & 154 \\ 255 & 249 & 255 & 255 & 255 & 255 & 223 & 158 \\ 214 & 233 & 255 & 255 & 255 & 244 & 205 & 150 \\ 191 & 204 & 224 & 234 & 227 & 211 & 181 & 135 \\ 172 & 181 & 197 & 211 & 204 & 194 & 169 & 127 \end{pmatrix} \quad M = \quad M = \begin{pmatrix} 40 & 49 & 69 & 86 & 78 & 43 & 7 & -21 \\ 61 & 77 & 103 & 122 & 113 & 81 & 39 & -14 \\ 84 & 106 & 127 & 127 & 127 & 119 & 72 & 10 \\ 127 & 120 & 127 & 127 & 127 & 127 & 91 & 26 \\ 127 & 121 & 127 & 127 & 127 & 127 & 95 & 30 \\ 86 & 105 & 127 & 127 & 127 & 116 & 77 & 22 \\ 63 & 76 & 96 & 106 & 99 & 83 & 53 & 7 \\ 44 & 53 & 69 & 83 & 76 & 66 & 41 & -1 \end{pmatrix}$$

Realizarea TDC ce se obține prin multiplicare conform ecuației (1.6).

$$D = DCT * M * DCT' \quad (1.6)$$

$$D = \begin{pmatrix} 657,3750 & 142,6678 & -206,4436 & 58,7859 & -39,8750 & 23,4702 & -4,1915 & 3,8011 \\ -15,8009 & 22,8682 & -22,0292 & -1,6961 & 8,2445 & -4,0462 & 0,2224 & 0,7316 \\ -185,2380 & -16,2231 & -15,9576 & -29,0072 & 23,1018 & -6,0983 & -3,7489 & -2,1520 \\ -8,2397 & 5,2529 & 0,5690 & -6,2527 & 7,8118 & -2,8027 & 0,0509 & 1,9857 \\ -23,6250 & 1,5207 & 32,4323 & 10,0600 & 7,6250 & 4,8556 & 0,2313 & -0,9192 \\ -9,5237 & 0,4714 & 7,2492 & -1,8331 & 1,2228 & -2,1269 & 1,9751 & 0,0458 \\ -1,3395 & -7,8871 & -3,4989 & -3,8412 & -7,2690 & -6,1836 & 2,9576 & -0,0257 \\ -3,3051 & -1,2988 & 4,3008 & 0,2533 & -2,1089 & -0,6532 & 0,1152 & -0,4886 \end{pmatrix}$$

Blocul $D(i, j)$ este o matrice bidimensională. În colțul de stînga-sus este elementul $D(0, 0)$ ce reprezintă cea mai mică frecvență a imaginii originale. Mișcîndune pe linii sau coloane coeficienții TDC corelează frecvențe mai mari și mai mari ale blocului de imagine, elementul $D(7, 7)$ corespunde celei mai mari frecvențe. Cu cît valoarea unui element al matricii obținute este mai aproape de zero cu atît și frecvența de apariție a imaginii este mai mare. Este important de remarcat că ochiul uman este mai sensibil la frecvențe joase, iar rezultatele din pasul de cuantizare va reflecta acest lucru.[1]

Compresia prin cuantizarea

În procesul de comprimare se pot selecta matrici de cuantizare specifice necesităților, aceste matrici sunt folosite de JPEG. Deci în dependență de cerințele față de calitatea imaginii ce dorim să obținem, se va alege nivelul calității ce variază de la 1 la 100. În cazul în care se alege primul nivel, imaginea va avea cea mai mică

calitate și cea mai mare compresie. În cazul alegerii nivelului de calitate 100, imaginea va avea cea mai înaltă calitate și cea mai mică compresie. Drept urmare raportul calitate / compresie îl alegem potrivit necesitatilor corespunzătoare.

Au fost efectuate calculele cu o matrice de cuantizare standart JPEG, cu un nivel de calitate de 50 %, această matrice face atât compresie de înaltă calitate cât și o calitate excelentă a imaginii decomprimată.[1]

$$Q50 = \begin{vmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 122 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 57 & 81 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{vmatrix}$$

$$Q10 = \begin{vmatrix} 80 & 60 & 50 & 80 & 120 & 200 & 255 & 255 \\ 55 & 60 & 70 & 95 & 130 & 255 & 255 & 255 \\ 70 & 65 & 80 & 120 & 200 & 255 & 255 & 255 \\ 70 & 85 & 110 & 145 & 255 & 255 & 255 & 255 \\ 90 & 110 & 185 & 255 & 255 & 255 & 255 & 255 \\ 120 & 175 & 255 & 255 & 255 & 255 & 255 & 255 \\ 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 \\ 255 & 255 & 255 & 255 & 255 & 255 & 255 & 255 \end{vmatrix} \quad Q90 = \begin{vmatrix} 3 & 2 & 2 & 3 & 5 & 8 & 10 & 12 \\ 2 & 2 & 3 & 4 & 5 & 12 & 12 & 11 \\ 3 & 3 & 3 & 5 & 8 & 11 & 14 & 11 \\ 3 & 3 & 4 & 6 & 10 & 17 & 16 & 12 \\ 4 & 4 & 7 & 11 & 14 & 22 & 21 & 15 \\ 5 & 7 & 11 & 13 & 16 & 12 & 23 & 18 \\ 10 & 13 & 16 & 17 & 21 & 24 & 24 & 21 \\ 14 & 18 & 19 & 20 & 20 & 20 & 20 & 20 \end{vmatrix}$$

Cuantizarea este realizată prin împărțirea fiecărui element din matricea $D(i,j)$ la $Q(i,j)$ și apoi rotunjirea la cea mai apropiată valoare întregă conform ecuației (1.7)[1].

$$C(i,j) = \text{round} \left(\frac{D(i,j)}{Q(i,j)} \right) \quad (1.7)$$

$$C = \begin{vmatrix} 41 & 13 & -21 & 4 & -2 & 1 & 0 & 0 \\ 0 & 2 & -2 & 0 & 0 & 0 & 0 & 0 \\ -13 & -1 & -1 & -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{vmatrix}$$

Colțul din stînga sus reprezintă frecvențele joase, iar zerourile reprezintă frecvențele mari mai puțin importante.

Pentru a comprima imaginea sunt citați fiecare pixel în zig-zag precum este prezentat în fig.1. Avantajul constă în numărul mare de zerouri ce duce la o compresie foarte înaltă.[1]

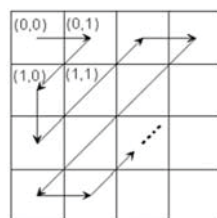


Figura 1 Citirea pixelilor în zig-zag

Decompresie

Decompresia imaginii începe prin decodificarea fluxului de biți care reprezintă matricea $C(i,j)$ cuantificată. Fiecare element al $C(i,j)$ este apoi înmulțit cu elementul inițial de cuantizare $Q(i,j)$ conform ecuației (1.8).[1]

$$R(i, j) = Q(i, j) * C(i, j) \quad (1.8)$$

$$R = \begin{pmatrix} 656 & 143 & -210 & 64 & -48 & 40 & 0 & 0 \\ 0 & 24 & -28 & 0 & 0 & 0 & 0 & 0 \\ -182 & -13 & -16 & -24 & 40 & 0 & 0 & 0 \\ -14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -18 & 0 & 37 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

ITDC se aplică matricii $R(i, j)$ care este rotunjită la cel mai apropiat număr întreg, apoi se adună valoarea 128 conform ecuației (1.9).[1]

$$N = \text{round}(DCT' * R * DCT) + 128 \quad (1.9)$$

$$N = \begin{pmatrix} 175 & 175 & 203 & 223 & 206 & 180 & 144 & 102 \\ 190 & 200 & 232 & 251 & 238 & 215 & 169 & 110 \\ 220 & 233 & 258 & 265 & 257 & 249 & 203 & 131 \\ 247 & 255 & 263 & 255 & 250 & 261 & 226 & 153 \\ 245 & 251 & 258 & 249 & 245 & 258 & 227 & 155 \\ 216 & 226 & 246 & 252 & 246 & 245 & 207 & 140 \\ 190 & 195 & 221 & 237 & 228 & 215 & 181 & 130 \\ 180 & 175 & 195 & 212 & 200 & 185 & 164 & 130 \end{pmatrix} \quad \text{Original} = \begin{pmatrix} 168 & 177 & 197 & 214 & 206 & 171 & 135 & 107 \\ 189 & 205 & 231 & 250 & 241 & 209 & 167 & 114 \\ 212 & 234 & 255 & 255 & 255 & 247 & 200 & 138 \\ 255 & 248 & 255 & 255 & 255 & 255 & 219 & 154 \\ 255 & 249 & 255 & 255 & 255 & 255 & 223 & 158 \\ 214 & 233 & 255 & 255 & 255 & 244 & 205 & 150 \\ 191 & 204 & 224 & 234 & 227 & 211 & 181 & 135 \\ 172 & 181 & 197 & 211 & 204 & 194 & 169 & 127 \end{pmatrix}$$

Acesta este un rezultat relativ bun, avînd în vedere că 70% din coeficienții TDC au fost egalați cu 0. Se observă că imaginea inițială nu diferă mult de imaginea ce a fost comprimată. Luînd în calcul că sunt 256 de nuanțe de gri într-o imagine alb-negru și ochiul uman abia recunoaște o deosebire la intervalul de 10 nuanțe.

Concluzie

Ca rezultat, a folosirii TDC și matricii de cuantizare JPEG, putem obține un raport calitate-compresie necesar necesitatilor noastre, fără a avea mari pierderi a calității imaginii inițiale. Folosind una din cele 3 matrici de cuantizare, sa observat, cînd nivelul de calitate a matricii de cuantizare este mai mare, obținem o calitate mai înaltă a imaginii cu o compresie mai mică. Dacă nivelul de calitate a matricii de cuantizare este mai mic, se va obține o compresie mai mare a imaginii, dar se va pierde din calitatea imaginii inițiale. Ținînd cont că matricea Q este aplicată pentru fiecare bloc 8x8 cu nivelul de calitate diferit apare necesitatea de a cerceta în continuare comportamentul fiecărei matrici $Q(i, j)$ cu nivelul de calitate de la 1 la 100 pe diferite porțiuni a imaginii.

Bibliografie

1. Ken Cabeen and Peter Gent „*Image Compression and the Discrete Cosine Transform*”
Web page. : <https://www.math.cuhk.edu.hk/~lmlui/dct.pdf>
2. Constantin Vertan „*Prelucrarea si analiza Imaginilor*” 1999