PROIECT PROCESARE DE IMAGINI COMPRESIE JPEG

Una dintre marile provocări ale procesării de imagini este lucrul cu cantități mari de date. Deoarece majoritatea imaginilor au rezoluții mari, împărțite pe mai multe canale. O imagine este compusă din pixeli, iar valoarea acestora le determină culoarea. Această abordare simplă fata de imagini este reflectată în formatul bitmap (.bmp) pentru imagini. Este rar folosit în zilele noastre, deoarece dimensiunea lui face dificil lucrul cu el.

Pentru a reduce dimensiunea mare a acestor imagini și pentru a le face mai ușor de folosit pentru aplicații de procesare (sau aplicații AI), aplicăm diferiți algoritmi de compresie pe aceste imagini bitmap de bază pentru a obține înapoi un fișier care este mai mic ca dimensiune, dar care conține totuși conținutul original al imaginii într-un mod pe care îl putem decoda.

Compresia imaginii poate fi fie cu pierderi, fie fără pierderi, în funcție de dacă pierdem sau nu informații despre imagine atunci când îi aplicăm compresie. Algoritmii de compresie au si o caracteristica numita compression ratio, exprimata in procente, care defineste cat de mare este imaginea comprimata in comparatie cu originala.

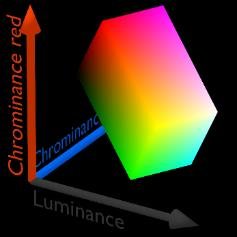
Algoritmii cu pierderi au, în general, rapoarte de compresie mai bune decât algoritmii fără pierderi, deoarece putem omite sau aproxima anumite caracteristici ale imaginii, reducand astfel dependenta de date explicite pentru imagine.

Unul dintre cele mai folositi algoritmi de compresie este compresia jpeg(Joint Photographic Experts Group). Este un algoritm lossy, deci pierdem calitate vizuala, dar poate ajunge la rapoarte de compresie foarte bune, dar acesta depinde de nivelul de detaliu dorit.

Algoritmul urmareste urmatorii pasi, incepand cu imaginea .bmp.

1. **Conversia RGB -> YCbCr.**

Fiindca ochiul uman este mai sensibil la intensitate decat culoare, incepem prin a converti spatiul de culoare al imaginii in Y(luminance) Cb (Chrominance-Blue) si Cr(Chrominance-Red).



Spatiul de culoare Y Cb Cr

Sursa: https://www.researchgate.net/figure/YCbCr-Color-Space-In-the-YCbCr-component-of-Y-is-a-luminance-brightness-whereas\_fig4\_323025324

Pentru a converti imaginea RGB in Y Cb Cr, aplicam urmatoarele operatii asupra canalelor RGB, care contin valori intre 0 si 255.

Aceste operatii vor rezulta in valori tot intre 0 si 255 pe care le putem folosi mai departe.

1. **Downsampling(Reducerea detaliului)**

Pentru a reduce mai departe mărimea fișierului, se reduce rezoluția canalelor Cb si Cr, se reduce mărimea imaginii, păstrând totuși datele importante de luminance.

Astfel, vom imparti rezolutia canalelor Cr si Cb la 2, iar pentru a afisa imaginea ne va trebui upsampling facut pe canalele Cr si Cb pentru a aproxima pixelii lipsa din acele canale.

Upsamplingul se poate face in mai multe moduri:

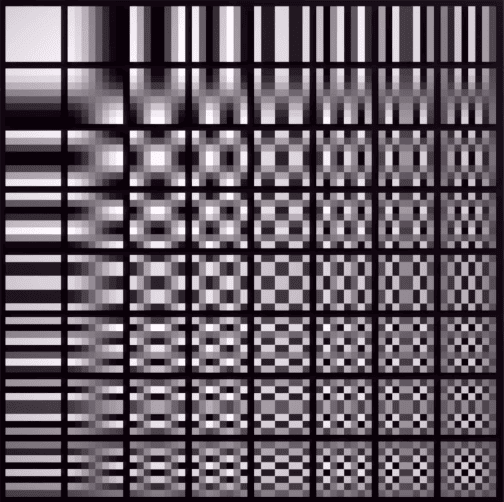
* Nearest neighbor
* Interpolare Biliniara
* Interpolare Bicubica

1. **Împărțirea in blocuri**

Imaginea este împărțita in blocuri de 8x8 pixeli, pentru a aplica urmatorul pas(Discrete Cosine Transform), pe o subsectiune a imaginii si nu pe toata imaginea. Fiindca am impartit deja rezolutia canalelor Cb si Cr la 2, impartind-o mai departe la 8 vom reduce marimea acelor canale de 16 ori, ceea ce ajuta mult la comprimare.

1. **Transformata Cosinus Discreta**

Aceste blocuri 8x8 pot fi reprezentate ca o multime de frecvente ale functiei cosinus, unde un maxim reprezinta o intensitate mare, iar un minim reprezinta o intensitate mica. Frecventele sunt distribuiute pe x si pe y, si prin combinatia acestora putem obitne orice imagine, daca atasam o pondere fiecarei componente si le adunam impreuna.



Tabelul de frecvente al transformatei, cu toate combinatiile de frecvente pe x si pe y

Sursa: https://www.baeldung.com/cs/jpeg-compression

Fiecare bloc 8x8 trece printr-o transformare discreta, ce transforma valorile pixelilor in componente de frecventa.

Pentru fiecare bloc 8x8, aplicam urmatoarea transformata:

Unde:

f(x,y) este valoarea unui pixel la pozitia (x,y)

F(u,v) este componenta de frecventa a transformatei

C(u) si C(v) sunt componente de normalizare, egali cu daca u = 0 sau v = 0, altfel 1

Dupa transformare, obtinem o matrice de ponderi, care, aplicate pe tabelul de transformari de mai sus si insumate, rezulta in imaginea originala, ca suma ponderata de frecvente.

La acest pas, nu se pierd informatii, iar imaginile transformate pot fi reconstruite aproape identic daca aplicam transformata inversa.

1. **Cuantizare**

Coeficienții transformatei cosinus sunt împărțiți la un coeficient de cuantizare dintr-un tabel de cuantizare standard, apoi sunt rotunjite in sus la cel mai apropiat intreg.

Tabelele de cuantizare sunt definite pentru luminanta(Y), si pentru canalele Chroma(Cr, Cb).

In G avem componentele de frecventa din DCT, iar in Q avem valorile matricii de cuantizare.

Tabelele standard de cuantizare pentru compresia jpeg sunt urmatoarele:

Pentru canalul Y:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 6 | 4 | 4 | 6 | 10 | 16 | 20 | 24 |
| 5 | 5 | 6 | 8 | 10 | 23 | 24 | 22 |
| 6 | 5 | 6 | 10 | 16 | 23 | 28 | 22 |
| 6 | 7 | 9 | 12 | 20 | 35 | 32 | 25 |
| 7 | 9 | 15 | 22 | 27 | 44 | 41 | 31 |
| 10 | 14 | 22 | 26 | 32 | 42 | 45 | 37 |
| 20 | 26 | 31 | 35 | 41 | 48 | 48 | 40 |
| 29 | 37 | 39 | 39 | 45 | 40 | 41 | 40 |

Pentru canalele de chroma:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 10 | 8 | 9 | 9 | 9 | 8 | 10 | 9 |
| 9 | 9 | 10 | 10 | 10 | 11 | 12 | 17 |
| 13 | 12 | 12 | 12 | 12 | 20 | 16 | 16 |
| 14 | 17 | 18 | 20 | 23 | 23 | 22 | 20 |
| 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 |
| 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 |
| 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 |
| 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 |

Se observa ca avem parametrii de cuantizare mult mai mari in coltul dreapta jos, si mult mai mici in coltul stanga sus. Acest lucru este datorat faptului ca detaliile de la frecventele mai mici sunt foarte greu de observat, si astfel ne permitem sa reducem detaliile acolo. De aici provine principala pierdere de date in compresia jpeg.

A number with numbers on it

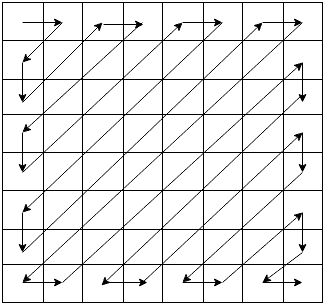
AI-generated content may be incorrect.

Bloc 8x8 dupa cuantizare, se observa cantitatea mare de zerouri in coltul dreapta jos, unde se afla coeficientii mai mari de cuantizare.

1. **Scanare ZigZag**

Dupa cuantizarea blocurilor, matricea contine cele mai importante valori(frecventele mici) in coltul stanga sus, iar cele mai putin importante in coltul dreapta jos.

Scanarea Zig-Zag reordoneaza matricea intr-o lista unidimensionala, parcursa de la coltul stanga sus la coltul dreapta jos. Deoarece frecventele cele mai importante sunt parcurse primele(la care se pierd mai putine detalii), cele neimportante ajung la final, unde avem sanse mari sa avem secvente lungi de 0 pe care le putem comprima folosind run length encoding.



Ordinea de parcurgere a bloculului cuantizat.

Sursa: https://www.baeldung.com/cs/jpeg-compression

1. **Crearea fisierului**

Datele comprimate, impreuna cu metadate, sunt grupate impreuna intr-un fisier cu extensia .jpg sau .jpeg. Care poate apoi fi deschis si afisat prin aplicarea algoritmului invers(decompresie).

Pentru ca fisierul sa se poata citeasca corect, trebuie incluse toate datele externe folosite, cum ar fi tabele de cuantizare, tabele de decodare huffman, etc.

In standardul JPEG, urmatoarele campuri sunt necesare:

1. Headerul JPEG(valoare magic, depinde de versiune, cel mai comun 0xFFD8)
2. eadTabele de cuantizare
   1. Header sectiune: 0xFFDB0043
   2. Tabelele de cuantizare pentru Y si C
3. Informatii despre imagine, printre care avem
   1. Header sectiune: 0xFFC00011
   2. Precizie(numar de biti): 0x08 pentru 8 biti
   3. Inaltimea imaginii
   4. Latimea Imaginii
   5. Numarul de canale(3 pentru YCbCr): 0x03
   6. Date despre componenta 1(Y):
      1. ID:0x01
      2. Raport subsampling
      3. Id tabel cuantizare
   7. Date despre componenta 2(Cb), aceleasi ca despre componenta 1
   8. Date despre componenta 3(Cr), aceleasi ca despre componenta 1
4. Tabele Huffman:
   1. Header sectiune: 0xFFC4
   2. Tabelele in sine
5. Detalii scanare:
   1. Header sectiune: 0xFFDA000C
   2. Numar componente: 0x03
   3. Asociere componenta, tabel huffman
      1. 0x01, 0x00 -> Componenta 1 foloseste tabelul 0
      2. 0x02, 0x01 -> Componenta 2 foloseste tabelul 1
      3. 0x03, 0x01 -> Componenta 3 foloseste tabelul 1
6. Datele comprimate ale imaginii.
7. Marker sfarsit fisier: 0xFFD9 (pentru toate fisierele JPEG)