

Partiel de compression de données

Les documents sont interdits, seules les calculettes sont autorisées, à l'exclusion de toute autre machine.

I – Entropie et ordre des données

Soit le fichier $F = \{ \text{BCADADBCBCADBCAD} \}$

1 – Calculez l'entropie d'ordre 1 de ces données.

A cet ordre le fichier est-il compressible ?

2 – Existe-t-il une entropie d'ordre supérieur qui révélerait une autre organisation des données ?

Le fichier serait-il alors compressible ?

II - Compression LZW

Le fichier : Lune.lzw , contient la séquence suivante où le caractère noté * est un caractère réservé de longueur 8 bits, prévenant le décompresseur qu'il doit ensuite lire les données sur 9 bits (les virgules ne font pas partie des données et ne sont là que pour faciliter la lecture):

d, o, *, 256, 256, r, e, m, i, 260, 256, 262, 260, 264, o

Le dictionnaire par défaut du décompresseur est constitué des 256 caractères ASCII. Le fichier se termine par un caractère de fin dont on ne tiendra pas compte.

1 – Décompressez ces données.

2 – Quel est le taux de compression de cette séquence.

3 - Dans le cadre d'une application où aucun caractère réservé n'est possible (un flux binaire codé sur 8 bits par exemple), quelle solution suggérez-vous pour s'en passer ?

III - Compression JPEG

1 – Quel est le rôle de la DCT (transformée en cosinus) dans l'algorithme de compression JPEG ?

2 – Après un rappel du rôle de la quantification dans la compression JPEG, expliquez la logique de la séquence zigzag lors de la linéarisation du spectre DCT quantifié ?

3 - Après compression d'une image I avec une matrice de quantification Q puis décompression, on obtient une image I' .

I' diffère-t-elle de I ?

On compresse ensuite I' avec le même algorithme et la même matrice de quantification Q . Après décompression on obtient l'image I'' .

I'' diffère-t-elle de I' ?

4 – Pour quelle raison principale JPEG utilise-t-il le codage des couleurs YUV et non RVB ?