

# Compression de données

G. TOCHON

Juin 2017

*Durée totale : 2h*

Tous les documents sont interdits. L'utilisation d'une calculatrice non programmable est autorisée.  
Les 4 exercices sont indépendants.

## Exercice 1 (6 points)

On considère le fichier  $F = \{\beta, \beta, \gamma, \beta, \phi, \gamma, \phi, \phi, \gamma, \beta, \beta, \gamma, \tau, \gamma, \delta, \alpha, \phi, \alpha, \beta, \beta\}$  (les virgules ne font pas partie des données et ne sont là que pour faciliter la lecture).

On supposera que la distribution de probabilité des symboles est donnée par la fréquence relative d'apparition du symbole dans le fichier.

1. Calculez l'entropie  $H$  du fichier  $F$ .
2. Déterminez l'encodage d'Huffman des symboles.
3. Quelle serait la taille minimale d'un fichier  $F'$ , composé de 1000 symboles tirés du même dictionnaire que  $F$  et avec la même distribution de probabilité ? Quelle taille permet d'atteindre l'encodage d'Huffman ?

## Exercice 2 (8 points)

Le fichier *music.lzw* contient la séquence d'adresses suivante, où le caractère noté % est un caractère réservé (les virgules ne font pas partie des données et ne sont là que pour faciliter la lecture) :

@D [D], @D [O], @D [R], @D [E], @D [%], 256, @D [M], @D [I], 258, 260, 262, @D [E]

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

1. Rappelez l'utilité du caractère réservé pour la compression/décompression LZW.
2. Décompressez les données du fichier *music.lzw* (incluez le tableau de décompression dans votre copie).
3. Quelle compression permet d'atteindre LZW pour ce fichier ?

## Exercice 3 (6 points)

Détaillez les différentes étapes, en expliquant leur intérêt, de la compression JPEG pour une image couleur.

## Exercice 4 (10 points)

Chaque question admet une, et une seule, réponse correcte.

Une bonne réponse rapporte 1 point. Une mauvaise réponse rapporte -0.5 point.

1. Laquelle de ces affirmations sur les coefficients de la décomposition en série de Fourier est vraie ?

$$A \square c_n = \frac{2}{T} \int_{-T/2}^{T/2} f(t) e^{i2\pi \frac{n}{T} t} dt$$

$$C \square c_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-i2\pi \frac{n}{T} t} dt$$

$$B \square a_n = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin\left(2\pi \frac{n}{T} t\right) dt$$

$$D \square b_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) \sin\left(2\pi \frac{n}{T} t\right) dt$$

2. À quelle fréquence doit au moins être échantillonné un signal dont la fréquence maximale est 10 kHz pour éviter le repliement spectral ?

A ☐ 20 Hz

B ☐ 10 kHz

C ☐ 20 kHz

D ☐ 50 kHz

3. Quelle est la principale différence d'une interpolation bilinéaire pour une image par rapport à une interpolation par plus proche voisin ?

A ☐ Elle est plus floue.

B ☐ Elle est moins floue.

C ☐ Elle n'a plus la même taille.

D ☐ Elle est plus "pixelisée".

4. Laquelle de ces affirmations sur les transformées de Walsh-Hadamard (WHT), de Fourier discrète (DFT) et en cosinus discrète (DCT) est fausse ?

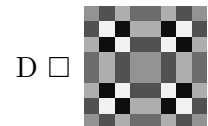
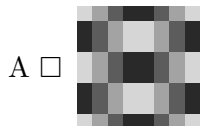
A ☐ La WHT compresse mieux l'énergie.

B ☐ La DCT est une transformée réelle.

C ☐ DCT et DFT sont fortement liées.

D ☐ La WHT n'est pas définie pour toutes les images.

5. Quelle est la fonction de base  $\mathcal{B}_{uv}$  de la DCT  $8 \times 8$  pour  $u = 0, v = 5$  ?



6. Quelle étape de l'algorithme JPEG introduit le plus de perte d'information ?

A ☐ Le calcul de la DCT

B ☐ La quantification de la DCT

C ☐ L'encodage d'Huffman

D ☐ Aucune, JPEG est un algo sans perte

7. Pour une même image, classez les formats du plus volumineux au moins volumineux :

A ☐ 1 - PNG    2 - TIF    3 - JPEG (qualité q=50)    4 - JPEG (qualité q=20)

B ☐ 1 - TIF    2 - PNG    3 - JPEG (qualité q=20)    4 - JPEG (qualité q=50)

C ☐ 1 - JPEG (qualité q=20)    2 - JPEG (qualité q=50)    3 - PNG    4 - TIF

D ☐ 1 - TIF    2 - PNG    3 - JPEG (qualité q=50)    4 - JPEG (qualité q=20)

8. Quelle est la couleur d'un pixel dont la représentation RGB est  $[40, 240, 10]$  ?



9. Quel espace colorimétrique est utilisé dans la compression JPEG couleur ?

A ☐ YUV

B ☐ RGB

C ☐ HSV

D ☐ SWAG

10. Combien de composantes principales possède une image RGB de dimensions spatiales 100 lignes  $\times$  200 colonnes ?

A ☐ 3

B ☐  $100 \times 200$

C ☐  $100 \times 200 \times 3$

D ☐ La réponse D