

Examen **Analyse Multi-résolution et Ondelettes**

26 janvier 2012

Durée de l'examen : 2h. Portables éteints et rangés.
Documents autorisés : photocopiés de cours (éventuellement annotés).
Les deux parties sont indépendantes et à rendre sur des copies séparées.

Le barème, sur 40, n'est donné qu'à titre indicatif, et est susceptible d'être modifié.

1 1^{ère} Partie : Représentation des signaux et des images (20 points)**Exercice 1 — Résolution spatiale et fréquentielle (10 points)**

On cherche à mettre en place un système d'analyse sonore dont l'objectif est de détecter la succession de deux sons successifs. Les deux sons, représentés par les signaux $x_1(t)$ et $x_2(t)$, ont des hauteurs différentes. On considérera le cas simpliste où $x_1(t)$ et $x_2(t)$ sont des fréquences pures, de fréquences f_1 et $f_2 = 2 \cdot f_1$. Chaque signal est fenêtré sur 4 périodes. Le processus de fenêtrage sera modélisé par une fonction porte.

1. Donner la formule de $x_1(t)$ et $x_2(t)$.
2. Donner la formule des Transformées de Fourier de $x_1(t)$ et $x_2(t)$, notées respectivement $X_1(f)$ et $X_2(f)$.
3. On considère maintenant le signal dans lequel $x_c(t)$ correspondant à mettre "bout à bout" les deux signaux $x_1(t)$ et $x_2(t)$. Donner la définition formelle de $x_c(t)$. Exprimer la transformée de Fourier $X_c(f)$ de $x_c(t)$ et la représenter graphiquement.
4. Les composantes fréquentielles de $x_1(t)$ et $x_2(t)$ sont-elles bien séparées dans $X_c(f)$? **N.B.** : On considérera que la fonction sinus cardinal est nulle en dehors de son premier lobe.
5. L'application visée consiste à détecter automatiquement les cas où $x_1(t)$ apparaît avant $x_2(t)$. Peut-on utiliser directement $X_c(f)$ dans ce cadre applicatif? Justifier.
6. On considère maintenant que le calcul de la transformée de Fourier à court terme avec une fenêtre rectangulaire de taille T . On calcule des transformées sans superposition, *i.e.* en translatant la fenêtre temporelle de sa largeur T . Peut-on réaliser l'application demandée avec une fenêtre de largeur :
 - $T = \frac{5}{f_1}$? Justifier.
 - $T = \frac{2}{f_1}$? Justifier.
 - $T = \frac{1}{f_1}$? Justifier.

Donner la forme des spectrogrammes dans chacun des cas. De quel principe le compromis illustré ici est-il la conséquence?

Exercice 2 — Base d'ondelettes orthonormées (4 points)

On considère l'espace vectoriel des fonctions de carré intégrable sur $[0, 1[: E = L^2([0, 1[)$
Soit la fonction ϕ définie sur $[0, 1[$ par :

$$\phi(t) = \begin{cases} 1 & \text{pour } 0 \leq t < 1 \\ 0 & \text{autrement} \end{cases} \quad (1)$$

et $\phi_k^j(t) = \phi(2^j t - k)$:

$$\phi_k^j(t) = \begin{cases} \sqrt{2^j} & \text{pour } \frac{k}{2^j} \leq t < \frac{k+1}{2^j} \\ 0 & \text{autrement} \end{cases} \quad (2)$$

Soit V^j l'espace vectoriel des fonctions constantes sur $\{[\frac{k}{2^j}, \frac{k+1}{2^j}[]_{k \in \{0; 2^j-1\}}$ (dimension 2^j)

1. Proposer une famille de fonctions $(\phi_k^j)_{k \in \{0; 2^j-1\}}$ formant une base orthonormée d'un espace vectoriel W^j telles que $V^{j+1} = V^j \oplus W^j$.
2. Exprimer les (ϕ_k^{j+1}) en fonction des (ϕ_k^j) et des (ψ_k^j) , et inversement
3. La relation liant coefficients dans $V^j \oplus W^j$ à ceux dans V^{j+1} est-elle la même que celle liant les fonctions de base ?
4. Exprimer les coefficients (s_k^{j+1}) en fonction des (s_k^j) et des (d_k^j) , et inversement

Exercice 3 — Algorithme K-SVD pour la compression (6 points)

En TME, des applications de l'algorithme K-SVD ont été vues pour des problèmes de débruitage et de reconstruction de données manquantes. On propose ici d'étudier un article de recherche [1] qui utilise l'algorithme K-SVD pour la compression d'images faciales. Voici le résumé de l'article :

"In this work we propose a new method for compressing facial images, based on the K-SVD algorithm. We train K-SVD dictionaries for predefined image patches, and compress each new image according to these dictionaries. The encoding is based on sparse coding of each image patch using the relevant trained dictionary, and the decoding is a simple reconstruction of the patches by linear combination of atoms. An essential pre-process stage for this method is an image alignment procedure, where several facial features are detected and geometrically warped into a canonical spatial location. We present this new method, analyze its results and compare it to several competing compression techniques."

... Et un autre extrait détaillant la procédure d'apprentissage :

"The K-SVD training process is an off-line procedure, preceding any image compression. The training produces a set of K-SVD dictionaries that are then considered fixed for the image compression stage. A single dictionary is trained for each 15×15 patch over a set of examples that will be referred to as the learning set. Prior to the training process for each patch, the mean patch image of the examples in the learning set is calculated and subtracted from all the examples in this set."

1. Quelle différence principale y a-t-il dans l'apprentissage du dictionnaire proposé ici par rapport au dictionnaire parcimonieux appris sur des bases d'images naturelles en TME ?
– Expliquer la nécessité et le principe de l'étape d'alignement. Pourquoi n'était-elle pas nécessaire dans l'apprentissage fait en TME ?
2. Décrire la phase hors-ligne de l'algorithme d'apprentissage de dictionnaire parcimonieux. Quelle est la sortie de l'algorithme ?

- Donner les grandes étapes de la chaîne complète de compression une fois le dictionnaire appris, et faire un schéma.
 - Pourquoi va-t-on pouvoir compresser efficacement le codage effectué pour chaque patch ?
3. La figure 1, issue de l'article présente les performances comparatives de la méthode de compression par rapport à l'état de l'art.
- Par rapport à JPEG ou JPEG 2000, quelle est la principale différence de la méthode pouvant expliquer les meilleures performances ?
 - La méthode VQ (vector quantization) consiste à remplacer l'apprentissage de dictionnaire par une étape de K-Means, et le codage par un assignement (hard) au plus proche voisin. Expliquer les différences avec la méthode présentée dans l'article, en quoi le K-SVD peut-il être vu comme une généralisation ? Pourquoi le codage est-il, plus fin ?

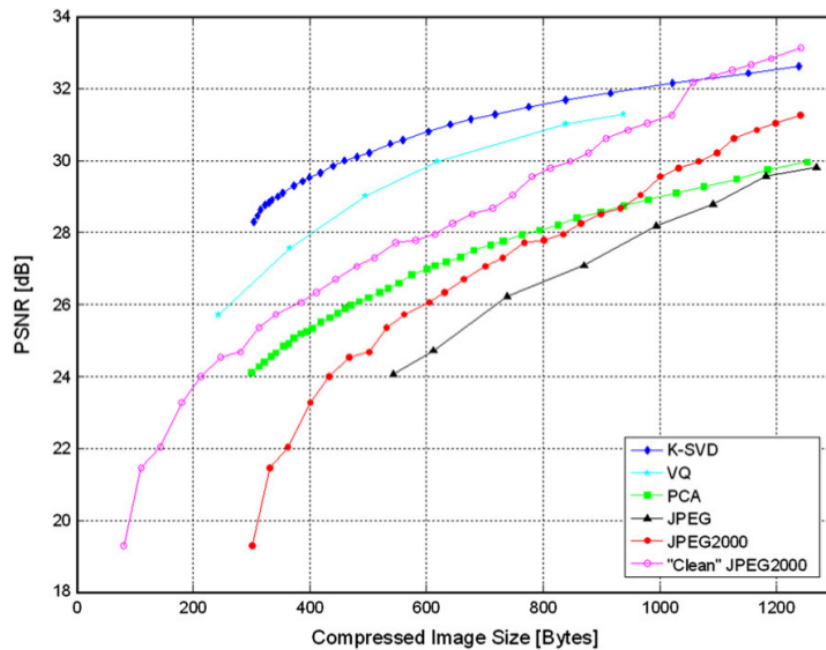


FIGURE 1 – Résultats comparatifs de la méthode de compression, issu de [1]

2 2^{nde} partie : Espaces d'échelles (20 points)

Exercice 4 — Compréhension du cours (6 points)

Répondre, **en justifiant**, aux questions suivantes :

1. On considère le filtre discret moyennneur défini sur une fenêtre carrée de taille t :
 - Rappeler sa définition mathématique appliquée à une image (*i.e.* exprimer $T_t(I)$). Ce filtre est-il linéaire ?
 - Respecte-t-il la causalité ? l'unimodalité ? la normalisation ?
 - Localise-t-il bien les contours ?
 - Est-il un noyau d'espace d'échelle admissible ?
2. On considère maintenant une représentation linéaire d'une image définie sur un domaine continu. Le noyau gaussien dérivé est-il un noyau d'espace d'échelle ?
3. Quand on discrétise le noyau gaussien, il y a plusieurs précautions à prendre. Quelles sont-elles ?
4. On rappelle la définition du détecteur de Harris-Laplace appliqué sur une image I :

$$\begin{aligned}R(L_t) &= \det(H(L_t) - \tau \operatorname{tr}(H(L_t))) \\ H(L_t) &= G_\sigma \star (\nabla L_t^T \nabla L_t)\end{aligned}$$

où L_t est la représentation à l'échelle t de l'image I , G_σ le noyau gaussien de variance σ^2 , \det est le déterminant d'une matrice et tr , sa trace. Rappeler les rôles des paramètres t , σ et τ .

Exercice 5 — Discrétisation (6 points)

- Le schéma FTCS de l'équation de la chaleur est-il explicite ou implicite ? Est-il stable ? Si oui, à quelle conditions ?
- Même question pour le schéma BTCS.
- Soit le schéma numérique suivant :

$$u_j^{n+1} = \frac{1}{2}(u_{j+1}^n + u_{j-1}^n) + \frac{\Delta t}{2\Delta x}(u_{j+1}^n - u_{j-1}^n)$$

où $u_j^n = u(x_0 + \Delta x j, t_0 + \Delta t n)$ et u est une fonction continue et dérivable sur \mathbb{R}^2 .

- De quelle équation aux dérivées partielles ce schéma est-il l'approximation ?
- Montrer que ce schéma est stable (au sens de l'analyse de Fourier) modulo une condition sur Δx et Δt que l'on donnera (indication : se souvenir que $\cos^2 \theta + \sin^2 \theta = 1$).

Exercice 6 — Applications (8 points)

Le but de cet exercice est de définir un espace d'échelles pour les images multispectrales. Dans une image multispectrale à n bandes, on trouve un n -uplets de valeurs à chaque pixel, et à chacune de ces valeurs correspond une acquisition dans l'une des n bandes disponibles. Une telle image est donc représentée par une fonction définie sur \mathbb{R}^3 , $I(x, y, \lambda)$, deux variables (x, y) pour le domaine spatial et une variable λ pour le domaine spectral.

Dans un premier temps, on considérera que I est une fonction continue sur ses 3 variables. On souhaite maintenant construire une représentation multi-échelles en espace et une représentation multi-échelles en λ .

1. On fixe la variable λ à λ_0 , définir la famille de fonctions $(x, y) \mapsto L^1(x, y, \lambda_0, t)$ telle qu'elle soit une représentation multi-échelles linéaire et continue de $(x, y) \mapsto I(x, y, \lambda_0)$ et t est un paramètre d'échelle spatial. Donner l'équation aux dérivées partielles dont L^1 est la solution.
2. On fixe maintenant la position (x, y) à (x_0, y_0) , définir la famille de fonctions $\lambda \mapsto L^2(x_0, y_0, \lambda, s)$ telle qu'elle soit une représentation multi-échelles linéaire et continue de $\lambda \mapsto I(x_0, y_0, \lambda)$ et s est un paramètre d'échelle spectral. Donner l'équation aux dérivées partielles dont L^2 est la solution.
3. Finalement, définir la famille de fonctions $(x, y, \lambda) \mapsto L(x, y, \lambda, t, s)$ paramétrées par s et t telle que si on fixe s on a une représentation multi-échelle linéaire de paramètre t et si on fixe t on a également une représentation multi-échelle linéaire de paramètre s . Donner le système d'équations aux dérivées partielles que vérifie L .

On souhaite maintenant appliquer cette représentation aux images couleurs RGB. Une image RGB est une image à trois bandes, une bande rouge, une bande verte et une bande bleue. Donc $n = 3$.

Pour appliquer ce qui précède, on choisit d'**échantillonner** le noyau gaussien et on s'intéresse à la représentation multi-échelles sur le paramètre de bande λ . Puisque l'espace des bandes est de petite dimension (3), on utilisera la convolution **circulante**¹ pour calculer les représentations multi-échelles sur le paramètre s .

1. Exprimer la représentation L à l'échelle spectrale s par une convolution discrète sur $L^1(x, y, t)$. Quelle est la taille du noyau de convolution ?
2. Application numérique : donner les valeurs du noyau de convolution pour les trois valeurs d'échelles suivantes :
 - $s = 1$ ($\exp(-1) \approx 0.32204$),
 - $s = 10$ ($\exp(-1/10) \approx 0.90484$),
 - $s = 100$ ($\exp(-1/100) \approx 0.9904$).
3. Lorsque s est grand, quel filtre classique la convolution précédente approche-t-elle ?
4. Qu'en déduisez-vous sur l'aspect des couleurs de L lorsque s est grand ? A quoi peut servir une telle représentation ?

Références

- [1] Ori Bryt and Michael Elad. Compression of facial images using the k-svd algorithm. *J. Vis. Commun. Image Represent.*, 19 :270–282, May 2008.

1. c'est-à-dire que le domaine sur λ est supposé périodique : $I(x, y, i) = I(x, y, (i + 3) \text{ modulo } 3) \forall i \in \mathbb{Z}$