Codage par transformée et par sous-bandes

Marco Cagnazzo

Département Traitement du Signal et des Images
TELECOM ParisTech

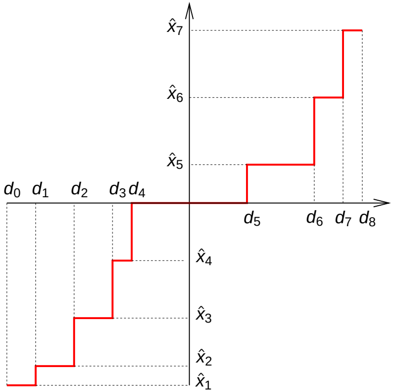
15 Décembre 2010

Plan

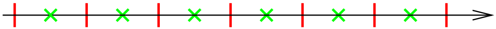
Plan

Rappel : Quantification scalaire

$$Q : x \in \mathbb{R} \to y \in \mathcal{C} = \{\hat{x}_1, \hat{x}_2, \ldots \hat{x}_L\} \subset \mathbb{R}$$



Quantification uniforme



- Simple
- Minimise l'erreur maximale
- Optimale pour v.a. uniforme

Courbe D(R) pour une v.a. uniforme :

$$D = \sigma_X^2 2^{-2R}$$

Courbe D(R) pour une v.a. non uniforme en *haute résolution* :

$$D = K_X \sigma_X^2 2^{-2R}$$

Quantification uniforme
Exemple de quantification

Image Originale, 24 bpp

Quantification uniforme

Exemple de codage

Débit 21 bpp PSNR 47.19 dB TC 1.143

Quantification uniforme
Exemple de codage

Débit 18 bpp PSNR 42.38 dB TC 1.333

Quantification uniforme

Exemple de codage

Débit 15 bpp PSNR 36.97 dB TC 1.600

Quantification uniforme
Exemple de codage

Débit 12 bpp PSNR 31.40 dB TC 2.000

Quantification uniforme
Exemple de codage

Débit 9 bpp PSNR 29.26 dB TC 2.667

Quantification uniforme

Exemple de codage

Débit 6 bpp PSNR 27.83 dB TC 4.000

Quantification uniforme

Exemple de codage

Débit 3 bpp PSNR 25.75 dB TC 8.000

Quantification

Conclusion

- Quantification : au centre de la compression avec perte
- Opération non réversible
- Centrale dans le compromis débit-distorsion
- Approximations à haute résolution:
  - $D \sim 2^{-2R}$
  - $D \sim \sigma^2$
- La seule quantification est insuffisante à assurer des bonnes performances de compression : **tout échantillon a la même importance !**

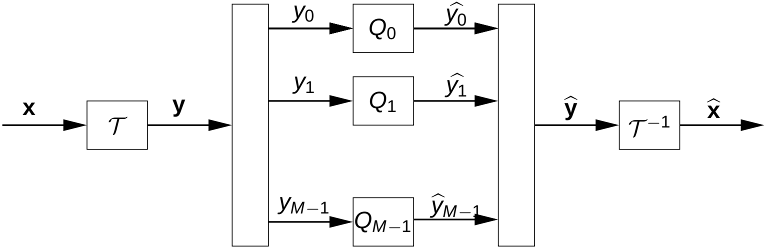
Plan

Principes de la transformée linéaire

- Transformation linéaire : changement de base
- Représentation alternative du signal
  - Séparation des données entre importants et pas importants (concentration de l'energie)
  - Déterminer les informations importantes pour notre perception
  - Allocation des ressources entre données importantes et peu importantes
- Mise en évidence des caractéristiques
- Réduire la corrélation

Transformée linéaire

Paradigme du codage par transformée



On passe du vecteur $\mathbf{x}$ à $\mathbf{y} = \mathcal{T}\mathbf{f}$: on veut un vecteur plus "facile" à quantifier : peu de coefficients "importants", beaucoup de coefficients "insignifiants"

Transformations unitaires

$\mathcal{T}$ unitaire $\Rightarrow \mathcal{T}^{-1} = \mathcal{T}^H$

avantages :

1. inversion immédiate
2. conservation de la norme : $\|\mathbf{Y}\| = \|\mathbf{X}\|$

$\Rightarrow$ Distorsion sur **Y** = distorsion sur **X** :

$$\mathrm{E}\left[\|\mathbf{Y}-\widehat{\mathbf{Y}}\|^2\right] = \mathrm{E}\left[(\mathbf{Y}-\widehat{\mathbf{Y}})^H(\mathbf{Y}-\widehat{\mathbf{Y}})\right] = \mathrm{E}\left[(\mathbf{X}-\widehat{\mathbf{X}})^H\mathcal{T}^H\mathcal{T}(\mathbf{X}-\widehat{\mathbf{X}})\right]$$

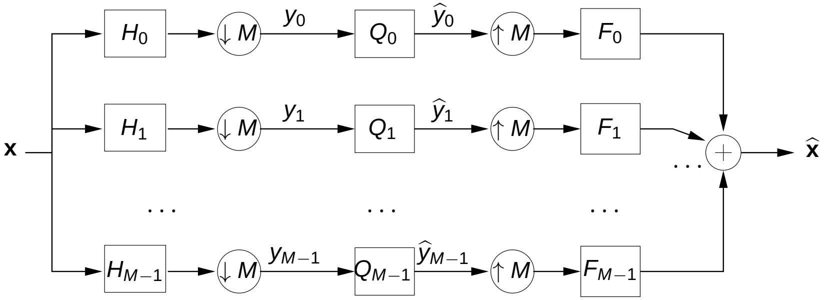
$$= \mathrm{E}\left[\|\mathbf{X}-\widehat{\mathbf{X}}\|^2\right]$$

Propriété fondamentale pour décider l'allocation des ressource dans le domaine transformée

Principes du codage par sous-bandes

- Analyse temps-fréquence du signal
- Implantation par banc de filtres
- Échantillonnage critique
- Équivalence avec la transformée linéaire

Codage en sous-bandes



$$y_k(m) = \sum_{\ell=0}^{M-1} h_k(\ell)x(mM-\ell) \qquad \mathbf{y} = \mathcal{T}\mathbf{x}$$

$$\mathbf{x} = \mathcal{P}\mathbf{y} \qquad \text{RP} : \mathcal{P}\mathcal{T} = \mathcal{I}$$

Codage en sous-bandes

Filtres d'analyses et de synthèse

- Déterminer les filtres en sorte que la condition de RP soit respectée est difficile
- Solution simple : bancs de filtres modulée

$$h_k(n) = d_k(n)h(n) \qquad f_k(n) = d_k(n)f(n)$$

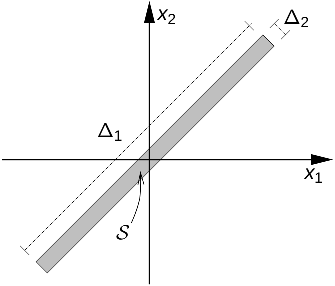
$$d_k(n) = \sqrt{\frac{2}{M}} \cos\left[(2k+1)(2n+1+N-M)\frac{\pi}{4M}\right]$$

$$h(n) = f(n) = \sin\left[(2n+1)\frac{\pi}{4M}\right]$$

- Choix de N et M : compromis entre résolution en fréquence et en temps

Codage par transformée
Exemple

Couple de v.a. fortement corrélées



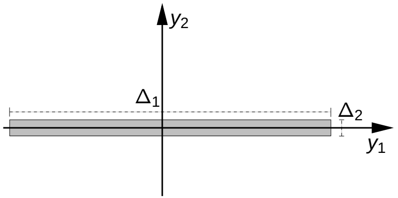
$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} \frac{1}{\Delta_1\Delta_2} & \text{si } (x_1,x_2) \in \mathcal{S} \\ 0 & \text{si } (x_1,x_2) \notin \mathcal{S} \end{cases}$$

$$\Delta_1 \gg \Delta_2$$

$$X_1 \sim X_2 \sim \mathcal{U}[-\frac{\Delta_1}{2\sqrt{2}}, \frac{\Delta_1}{2\sqrt{2}}]$$

Codage par transformée
Transformée : rotation de 45 degrés

Après transformation : v.a. indépendantes



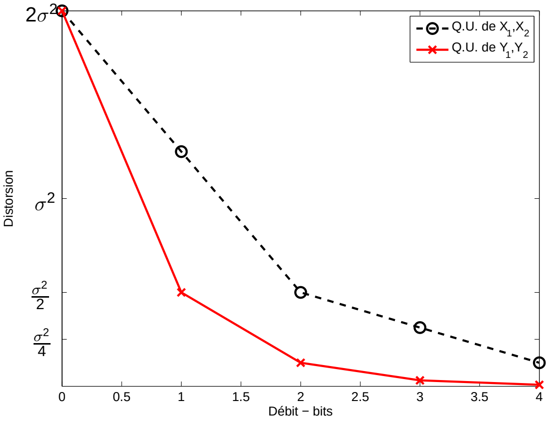
$$f_{Y_1,Y_2}(y_1,y_2) = \begin{cases} \frac{1}{\Delta_1\Delta_2} & \text{si } (y_1,y_2) \in \mathcal{S} \\ 0 & \text{si } (y_1,y_2) \notin \mathcal{S} \end{cases}$$

$$Y_1 \sim \mathcal{U}[-\frac{\Delta_1}{2}, \frac{\Delta_2}{2}]$$

$$Y_2 \sim \mathcal{U}[-\frac{\Delta_2}{2}, \frac{\Delta_2}{2}]$$

Codage par transformée

Performances RD de la quantification après transformée



Plan

Allocation des ressource

- Composants de $\mathbf{y} = \mathcal{T}\mathbf{x}$ interprétées comme réalisations de M processus aléatoires stationnaires de variance σ_k^2
- Cas de transformée linéaire unitaire : $\mathcal{T}$ est une matrice carrée et la distorsion est la même dans le 2 domaines

$$\begin{aligned}
\mathcal{D} &= \mathrm{E}\left[\|\mathbf{X} - \widehat{\mathbf{X}}\|^2\right] = \mathrm{E}\left[\|\mathbf{Y} - \widehat{\mathbf{Y}}\|^2\right] \\
&= \mathrm{E}\left[\sum_{k=0}^{M-1}(Y_k - \widehat{Y}_k)^2\right] = \sum_{k=0}^{M-1}\mathrm{E}\left[(Y_k - \widehat{Y}_k)^2\right] \\
&= \sum_{k=0}^{M-1}\mathcal{D}_k
\end{aligned}$$

- On connaît la relation entre $\mathcal{D}_k$ et le débit de quantification b_k (hypothèse HR) :

$$\mathcal{D}_k = c_k\sigma_k^2 2^{-2b_k}$$

Allocation des ressource
Définition du problème

- Allocation optimale : minimiser $\mathcal{D}$ sous contrainte de débit

$$\min \mathcal{D}(\mathbf{b}) = \sum_{k=0}^{M-1} c_k \sigma_k^2 2^{-2b_k} \qquad \text{soumis à } \sum_{k=0}^{M-1} b_k \leq B$$

- Technique de Lagrange. Minimiser le critère :

$$J(\mathbf{b}, \lambda) = \sum_{k=0}^{M-1} c_k \sigma_k^2 2^{-2b_k} + \lambda \left(\sum_{k=0}^{M-1} b_k - B \right)$$

- Solution (formule de Huang-Schulteiss) :

$$b_k^* = \frac{B}{M} + \frac{1}{2} \log \left[\frac{c_k \sigma_k^2}{c_{\mathrm{GM}} \sigma_{\mathrm{GM}}^2} \right]$$

Allocation des ressources

Interprétation du résultat

- Les ressources sont reparties uniformément ($\bar{b} = B/M$) avec une correction qui dépend des variances
- Valeur de la distorsion des composantes :

$$\mathcal{D}_k^* = c_k \sigma_k^2 2^{-2b_k^*} = c\sigma_k^2 2^{-2\bar{b}} \frac{c_{\mathrm{GM}}\sigma_{GM}^2}{c_k\sigma_k^2} = c_{\mathrm{GM}}\sigma_{GM}^2 2^{-2\bar{b}}$$

$$\mathcal{D}_{\mathcal{T}} = \sum_{k=0}^{M-1} \mathcal{D}_k^* = Mc_{\mathrm{GM}}\sigma_{\mathrm{GM}}^2 2^{-2\bar{b}}$$

- Données Gaussiennes : $c_{\mathrm{GM}} = c_k = c_{\mathcal{N}}$ et donc

$$b_k^* = \bar{b} + \frac{1}{2}\log\left[\frac{\sigma_k^2}{\sigma_{GM}^2}\right] \qquad \mathcal{D}_k^* = c_{\mathcal{N}}\sigma_{\mathrm{GM}}^2 2^{-2\bar{b}}$$

$$\mathcal{D}_{\mathcal{T}} = Mc_{\mathcal{N}}\sigma_{\mathrm{GM}}^2 2^{-2\bar{b}}$$

Gain de codage

- Soit X un vecteur aléatoire de M données d'entrée (son, image...)
- Hypothèse : les composantes de X sont i.d., p.e. Gaussiennes $\mathcal{N}(0, \sigma_X^2)$
- Sans transformée, le mieux qu'on puisse faire est quantification et allocation optimale des ressources (PCM). La distorsion est :

$$\mathcal{D}_{\mathrm{PCM}} = M c_{\mathcal{N}} \sigma_X^2 2^{-2\bar{b}}$$

Gain de codage

- Le *gain de codage* d'une transformée $\mathcal{T}$ est défini comme le rapport entre la distorsion qu'on aurait sans transformée, et la distorsion qu'on peut atteindre avec la transformée :

$$G_{\mathcal{T}} = \frac{\mathcal{D}_{\mathrm{PCM}}}{\mathcal{D}_{\mathcal{T}}} = \frac{\sigma_X^2}{\sigma_{\mathrm{GM}}^2} = \frac{\sigma_{\mathrm{AM}}^2}{\sigma_{\mathrm{GM}}^2}$$

- La transformée doit rendre les variances des composantes du vecteur Y les plus inégales possible
- La moyenne géométrique d'un ensemble de nombres positifs est toujours inférieur ou égal à la moyenne arithmétique

Allocation des ressources
Algorithmes pratiques

La formule de Huang-Schulteiss

- Peut donner des valeurs négatifs
- Peut donner des valeurs fractionnaires

Donc on utilise des algorithmes sous-optimales

- Algorithme de Huang-Schulteiss modifié
- Algorithme *greedy*

Allocation des ressources
Algorithme de Huang-Schulteiss modifié

1. On calcule b_k^* avec H.-S.
2. Si certaines b_k^* sont négatifs, on répète l'algorithme en enlevant les σ_k^2 concernée ; ces variables seront codées avec zéro bitses variables seront codées avec zéro bits
3. Le pas précédent est répété jusqu'à quand il n'y a plus de valeurs négatifs
4. Les valeurs trouvé sont arrondis à l'entier inférieur
5. Le débit résiduel est alloué aux coefficients avec l'erreur maximum

Algorithme *greedy*

1. Initialisation
   - $b_k = 0 \quad \forall k \in \{0, 1, \dots M-1\}$.
   - $D_k = \quad \sigma_k^2 \forall k \in \{0, 1, \dots M-1\}$.
2. Tant que $\sum_k b_k \leq B$
   - $\ell = \arg\max_k D_k$
   - $b_\ell \leftarrow b_\ell + 1$
   - $D_\ell \leftarrow D_\ell / 4$

Plan

Transformation optimale : Transformations de Karhunen-Loève (TKL)

Définition : transformation permettant de décorréler les données
$\mathbf{R_Y} = \mathrm{E}\{\mathbf{YY}^H\}$: matrice diagonale
On utilise la diagonalisation de $\mathbf{R}_X$: si les composantes de $\mathbf{X}$ sont linéairement indépendantes, $\mathbf{R}_X$ est définie positive et admet la représentation :

$$\mathbf{R}_f = \mathbf{U}\Lambda\mathbf{U}^T$$

où : $\mathbf{U} = \left[\mathbf{u}_1 \,\vdots\, \mathbf{u}_2 \,\vdots\, \ldots \,\vdots\, \mathbf{u}_N\right]$ matrice des vecteurs propres orthonormaux et $\Lambda = \mathrm{Diag}(\lambda_1, \ldots, \lambda_N)$ est la matrice diagonale des valeurs propres.
La TKL est définie par :

$$\mathbf{T} = \mathbf{U}^T$$

Propriété de la TKL

- Transformation unitaire

$$\mathbf{T}^H\mathbf{T} = \mathbf{I}$$

- Transformation décorrélante

$$\mathrm{E}[Y_i Y_j] = \lambda_i \delta_{ij}$$

- Meilleure concentration des énergies :

$$\forall \mathbf{T} \text{ transformée unitaire, si}$$

$$\mathbf{Y} = \mathbf{U}^T\mathbf{X} \text{ et } \mathbf{Z} = \mathbf{T}\mathbf{X} \text{ alors}$$

$$\sum_{i=1}^{N} \overline{Y_i^2} \geq \sum_{i=1}^{N} \overline{Z_i^2}$$

- Transformée optimale pour variables Gaussiennes

Optimalité de la KLT

- $\mathbf{X} = \mathcal{N}(\mathbf{0}, \mathbf{R}_X)$
- $\mathbf{Y} = \mathbf{U}^T\mathbf{X}$ et $\mathbf{Z} = \mathbf{T}\mathbf{X}$
- $\mathbf{R}_Y = \mathbf{U}^T\mathbf{R}_X\mathbf{U}$ et $\mathbf{R}_Z = \mathbf{A}\mathbf{R}_X\mathbf{A}^T \Rightarrow$ $\det(\mathbf{R}_X) = \det(\mathbf{R}_Y) = \det(\mathbf{R}_Z) = \Pi_{k=0}^{M-1}\sigma_{k,\mathbf{Z}}^2$
- $\mathcal{S} = \mathrm{Diag}(\sigma_{1,\mathbf{Y}}, \sigma_{2,\mathbf{Y}}, \ldots, \sigma_{M,\mathbf{Y}})$
- $W_i = Y_i/\sigma_{i,\mathbf{Y}} \Rightarrow \mathbf{Y} = \mathcal{S}\mathbf{W}$
- $\Pi_{k=0}^{M-1}\sigma_{k,\mathbf{Z}}^2 = \det(\mathbf{R}_Y) = \Pi_{k=0}^{M-1}\sigma_{k,\mathbf{Y}}^2 \det(\mathbf{R}_W) \leq \Pi_{k=0}^{M-1}\sigma_{k,\mathbf{Y}}^2$ car le déterminant de la matrice d'autocorrélation d'un vecteur Gaussien centré à variances unitaires est toujours inférieur ou égal à un.

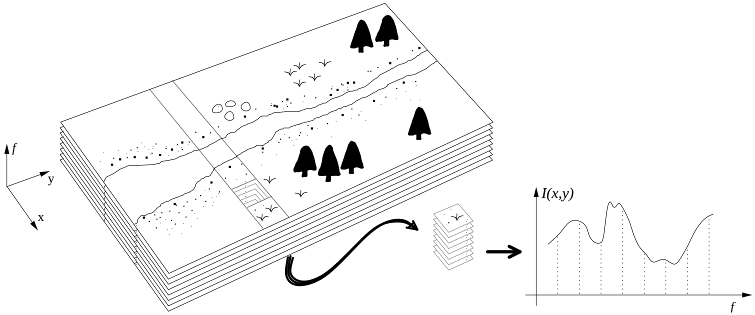
Bilan de la TKL

Avantages	Inconvénients
Décorrélation	Dépendante des données
Concentration des énergies	Complexité élevée
Optimale dans le cas Gaussien	Difficulté de l'estimation

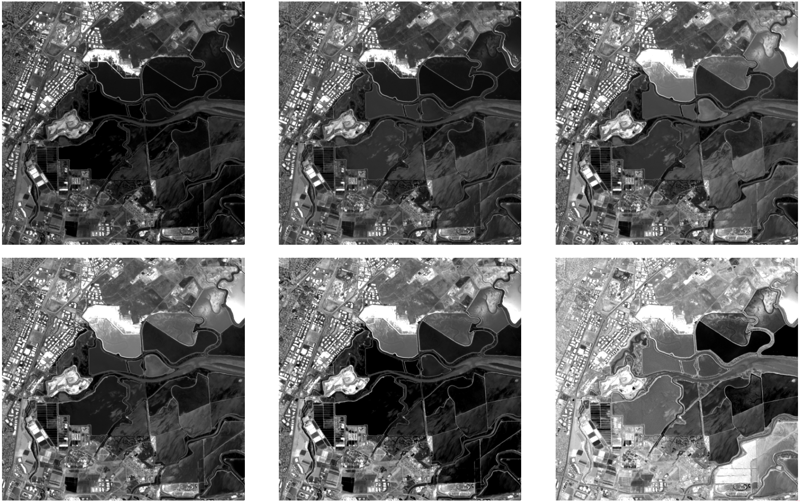
Images naturelles : non stationnaires
La TKL est rarement utilisée dans la compression d'image, excepté des cas particuliers

Exemple d'application : TKL

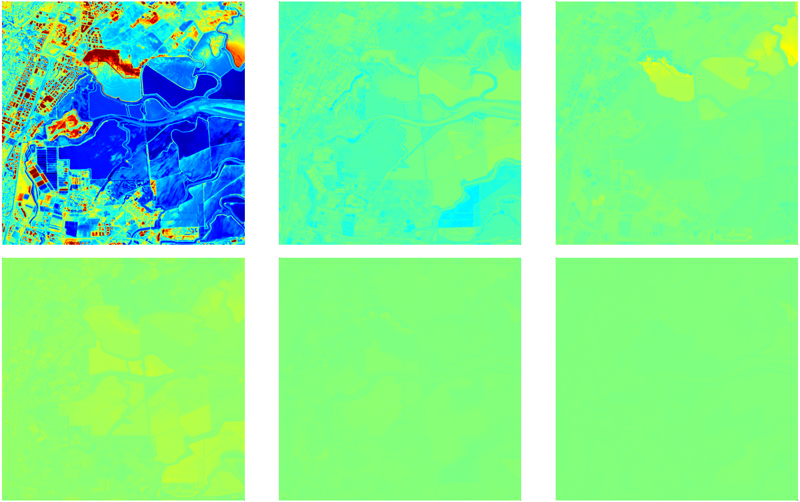
Images multispectrales



Exemple d'application : TKL



Exemple d'application : TKL



Exemple d’application : TKL

Concentration de l’énergie dans les “images-propres”

Image propre	1	2	3	4	5	6
% Energy	86.37	7.65	4.16	1.67	0.12	0.03

Codage par transformée : Bilan

- Changer la représentation du signal
- Concentrer l'énergie (information) en peu de coefficients (décorrélation)
  - Transformée optimale : KLT
  - Complexe, peu robuste (stationnarité)
  - Approximations de la KLT
- Allouer différemment les ressources de codage en fonction de l'importance des coefficients (allocation de débit)
  - Technique optimale : formule de Huang et Schulteiss
  - Version modifiée pour la pratique
  - Algorithme greedy
  - Algorithmes euristiques