

Projet ACI²M

Application de l'analyse en composantes indépendantes au codage d'images multicomposantes

Coordinateur : M. Barret (Supélec)

Partenaires : D. T. Pham (LMC), A. Mohammad-Djafari & P. Duhamel (LSS), C. Lambert (CNES) et P. Comon (I3S)

Objectifs

Etudier l'intérêt de l'analyse en composantes indépendantes (ACI) en codage d'images multispectrales et hyperspectrales satellitaires.

Deux approches :

1. Cadre classique de la compression quasi sans perte avec un codage séparé des composantes après transformation.
2. Commencer par une analyse de l'image pour y séparer des objets (fusion d'images et segmentation jointe), puis coder par région d'intérêt.

Comparer les performances des méthodes utilisant l'ACI avec d'autres méthodes (sans ACI) et avec l'état de l'art.

Les difficultés à résoudre

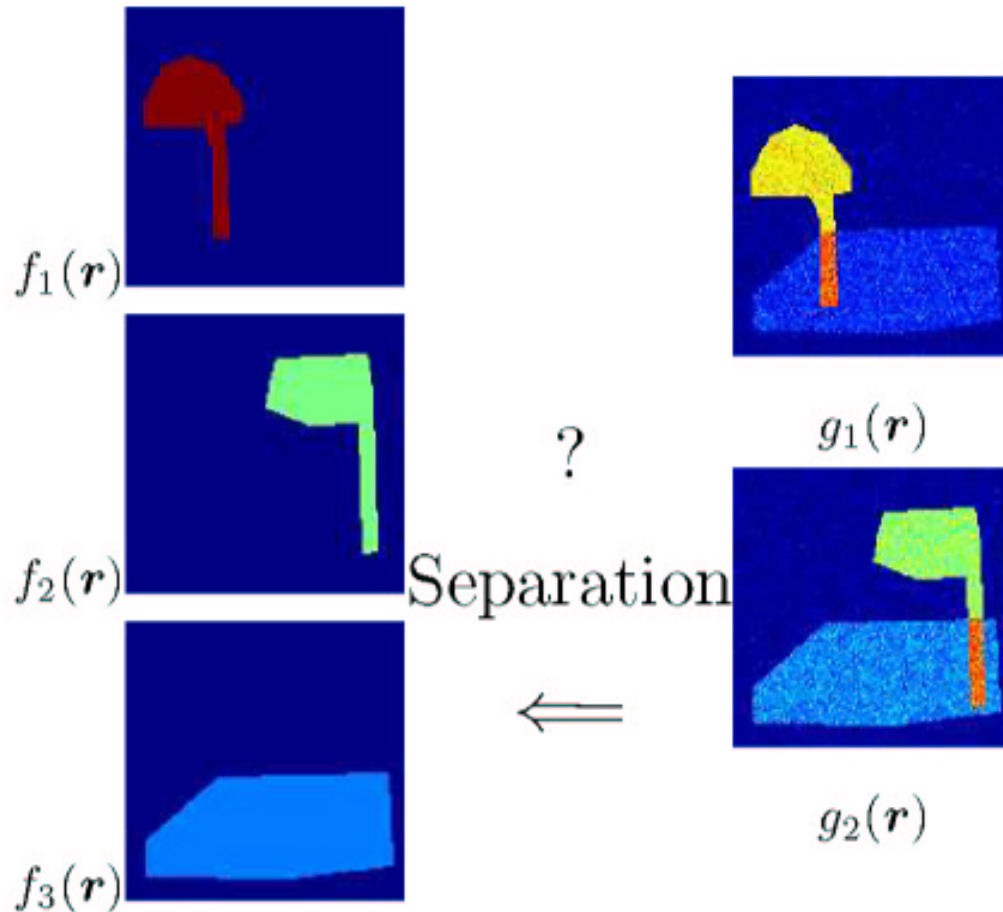
Pour la première approche :

1. Donner un cadre théorique pour comparer les performances (débit / distorsion) à faibles distorsions de différentes transformations, indépendamment (si possible) de la structure du codeur mis après la transformation. Notion de gain de codage d'une transformation.
2. Adapter les algorithmes d'analyse en composantes indépendantes pour maximiser le gain de codage de la transformation.
3. Vérifier expérimentalement la démarche. Analyser les résultats.

Pour la deuxième approche :

1. Adapter les méthodes bayésiennes de segmentation aux images multispectrales et hyperspectrales satellitaires.
2. Utiliser l'image segmentée pour faire de la compression.

Blind image separation and joint segmentation



$$g_i(\mathbf{r}) = \sum_{j=1}^N A_{ij} f_j(\mathbf{r}) + \epsilon_i(\mathbf{r})$$

$$\underline{g}(\mathbf{r}) = \{g_i(\mathbf{r}), i = 1, M\}$$

$$\underline{g}(\mathbf{r}) = \mathbf{A} \underline{f}(\mathbf{r}) + \underline{\epsilon}(\mathbf{r}),$$

$$\underline{g} = \{g_i(\mathbf{r}), i = 1, M\}$$

$$g_i = \{g_i(\mathbf{r}), \mathbf{r} \in \mathcal{R}\},$$

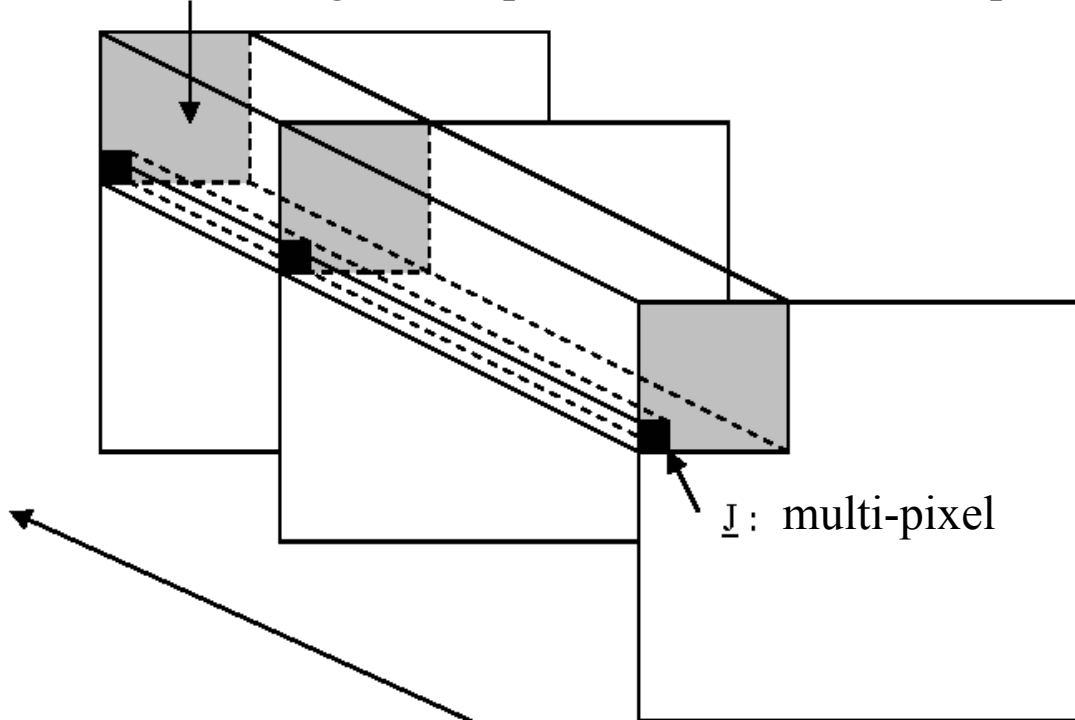
$$\underline{g} = \mathbf{A} \underline{f} + \underline{\epsilon}$$

A. Mohammad-Djafari & son équipe

Transformation appliquée à une image multicomposante

Convention :

\mathbf{J}^T : image multispectrale, bloc de T multi-pixels



N composantes de l'image multispectrale complète \mathbf{J}^L de dimension L

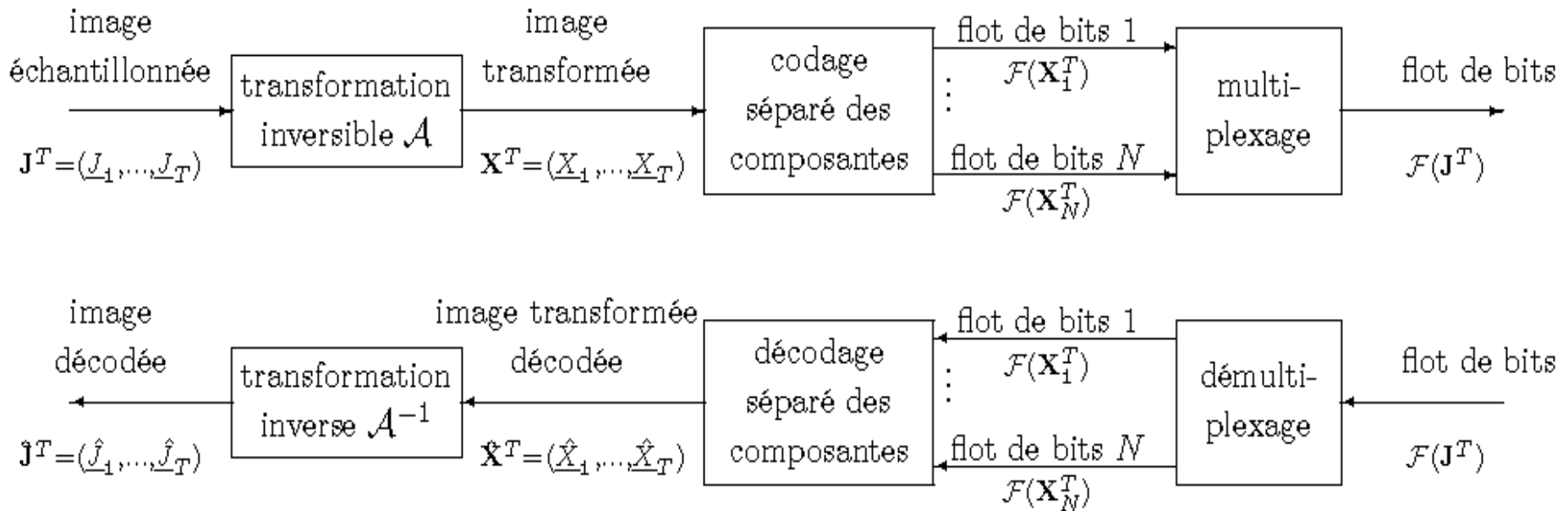
L'image entière \mathbf{J}^L est découpée en blocs $(\mathbf{J}^{T,i})_i$ que l'on supposera indépendants et identiquement distribués.

La même transformation inversible

$$\mathcal{A} : (\mathbb{R}^N)^T \rightarrow (\mathbb{R}^N)^T$$

est appliquée à chaque bloc $\mathbf{J}^{T,i}$, donnant le bloc transformé $\mathbf{X}^{T,i}$. Les blocs $(\mathbf{X}^{T,i})_i$ sont groupés pour former l'image \mathbf{X}^L .

Schémas classiques de compression d'images



Compression avec pertes. Après transformation, les fonctions de codage et de décodage agissent séparément sur les composantes.

Schémas classiques de compression d'images

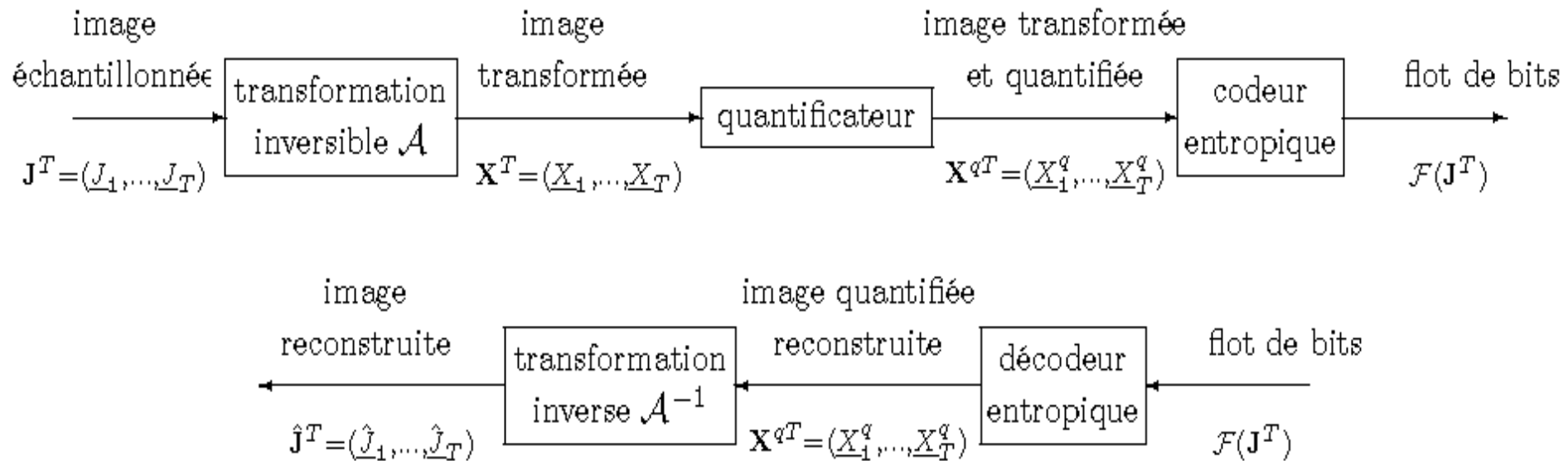


Schéma de compression avec pertes. Après transformation, chaque composante du bloc transformé \mathbf{X}^T subit une quantification suivie d'un codeur entropique. La quantification est scalaire ou vectorielle (de dimension K) et le codeur entropique d'ordre T/K .

Approximations classiques à faibles distorsions

Expression de la distorsion :

$$D(\mathbf{J}^T, \hat{\mathbf{J}}^T) = \frac{1}{N} \sum_{\hat{i}=1}^N w_{\hat{i}} \left(\frac{1}{T} \sum_{k=1}^T \mathbb{E} \left[(X_{\hat{i},k} - \hat{X}_{\hat{i},k})^2 \right] \right)$$

Expression du débit pour un codage optimal (la structure du codec étant fixée) :

$$R(D) \simeq \frac{1}{N} \sum_{\hat{i}=1}^N \frac{h(\mathbf{X}_{\hat{i}}^T)}{T} - \frac{1}{2} \log_2 D + \frac{1}{2N} \sum_{\hat{i}=1}^N \log_2 w_{\hat{i}} + R_0$$

avec

$$R_0 = \begin{cases} -\frac{1}{2} \log_2 [2\pi e] & \text{débit asymptotique de Shannon} \\ \frac{1}{2N} \sum_{\hat{i}=1}^N \log_2 G(\mathbf{U}_{\hat{i}}) & \text{débit d'une quantification vectorielle à haute résolution} \\ -\frac{1}{2} \log_2 12 & \text{débit d'une quantification scalaire à haute résolution.} \end{cases}$$

Gain de codage d'une transformation

Thèse de Michel Narozny

Les relations

$$h(\mathbf{X}^T) = h(\mathbf{J}^T) + \mathbf{E} [\log_2 |\text{Jac}_{\mathcal{A}}(\mathbf{J}^T)|] \quad \text{et}$$

$$\sum_{i=1}^N h(\mathbf{X}_i^T) = h(\mathbf{X}^T) + I(\mathbf{X}_1^T; \dots; \mathbf{X}_N^T)$$

impliquent

$$R_{\mathcal{A}}(D) \simeq \frac{I(\mathbf{X}_1^T; \dots; \mathbf{X}_N^T)}{NT} + \frac{\mathbf{E} [\log_2 |\text{Jac}_{\mathcal{A}}(\mathbf{J}^T)|]}{NT} + \frac{1}{2N} \sum_{i=1}^N \log_2 w_i + R_0(D)$$

où

$$R_0(D) = \frac{h(\mathbf{J}^T)}{NT} + R_0 - \frac{1}{2} \log_2 D$$

ne dépend pas de \mathcal{A} .

Gain de codage d'une transformation (2)

Thèse de Michel Narozny

- Gain de codage en débit d'une transformation \mathcal{A} .

$$\mathcal{G}(\mathcal{A}, D) = R_{\mathcal{A}_0}(D) - R_{\mathcal{A}}(D)$$

par rapport à une transformation étalon \mathcal{A}_0 .

- Le gain de codage ne dépend ni de la structure du codec, ni de la valeur de la distorsion (sous l'hypothèse de faibles distorsions).
- Pour une transformation linéaire \mathbf{A} et des blocs de dimension $T = 1$, maximiser le gain de codage revient à minimiser l'information mutuelle entre les composantes à laquelle s'ajoute une quantité qui mesure la distance à l'orthogonalité de la transformation.

Maximisation du gain de codage pour des transformations linéaires

Thèse de Michel Narozny

- Cas $T = 1$. Il s'agit de minimiser la conjonction de deux termes:

La « distance » à l'indépendance et

La « distance » à l'orthogonalité de \mathbf{A}

- D. T. Pham a modifié son algorithme ICAinf, pour maximiser le gain de codage : algorithmes Aopt et Qopt.

Vérification expérimentale (cas $T = 1$)

Thèse de Michel Narozny

Deux approches indépendantes pour $N = 3$:

1. Après transformation \mathcal{A} , sans préciser la structure du codeur situé en aval, on estime l'information mutuelle $I(\mathbf{X}_1^T; \dots; \mathbf{X}_N^T)$ avec deux estimateurs (Darbellay, Pham) et on calcule le terme supplémentaire (« distance » à l'orthogonalité).
2. Après transformation \mathcal{A} , on applique un quantificateur scalaire uniforme suivi d'un codeur entropique d'ordre 1 par composante et l'algorithme de Shoham et Gersho pour l'allocation optimale de bits entre composantes.

Pour de faibles distorsions, on obtient la même estimation du gain de codage.

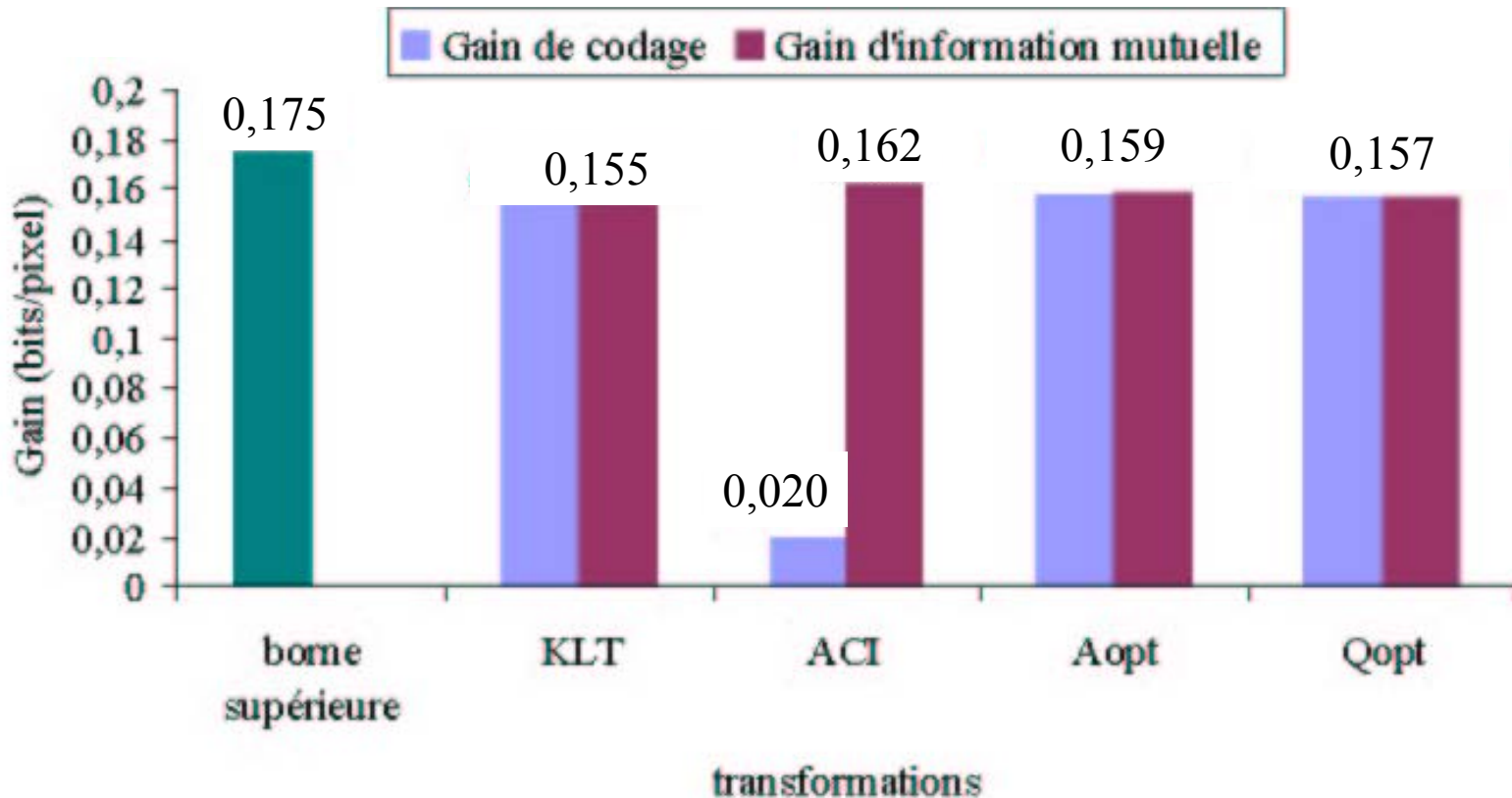
Résultats expérimentaux (cas $T = 1$)

Thèse de Michel Narozny

- Tests réalisés sur deux images multispectrales satellitaires Landsat (Paris et Rio) : 7 bandes, 8 bits/pixel, dimension 512 x 512.
- Chaque composante subit une décomposition en ondelettes 9/7.
- L'image \mathbf{J}^L à coder / décoder (avant transformation) est constituée des sous-bandes HL, HH ou LH (de dimension 256 x 256).
- La distorsion porte sur \mathbf{J}^L (l'ondelette 9/7 est quasi-orthogonale).
- La transformation étalon \mathcal{A}_0 est soit l'identité, soit la KLT.

Résultats expérimentaux (cas $T = 1$) (2)

Thèse de Michel Narozny



Paris, composantes 1, 2 et 3 ($N = 3$), sous-bande LH

Perspectives et « points durs »

Première approche :

- Les premiers tests montrent que la KLT donne de très bons résultats (cas $T=1$), le gain de codage maximal (pour une transformation linéaire) par rapport à la KLT valant quelques millièmes de bits/pixel.
- Poursuivre les tests sur des images fournies par le CNES en augmentant N (nombre de composantes) et T (nombre de pixels par composante), influence de la fonction coût ...
- Analyser les résultats, comprendre pourquoi la KLT fait si bien !

Deuxième approche :

- Segmentation d'images hyperspectrales par des méthodes bayésiennes (A. Mohammadpour / A. Mohammad-Djafari).