

# ANALYSE EN COMPOSANTES INDÉPENDANTES ET COMPRESSION DE DONNÉES

Michel Narozy\* et Michel Barret  
Équipe Systèmes de Traitement des Signaux  
Supélec  
2, rue E. Belin  
57070 Metz, France

emails : Michel.Narozy@metz.supelec.fr, Michel.Barret@supélec.fr

\*Ce travail a été financé en partie par la région Lorraine

## Résumé

Dans [1], Narozy *et al* ont montré que la recherche de la transformation linéaire optimale en codage par transformée peut se ramener, sous l'hypothèse haute résolution, à un problème d'analyse en composantes indépendantes (ACI) modifié. À partir de ce nouveau point de vue, ils ont proposé deux nouveaux algorithmes, GCGsup et ICAorth, pour calculer respectivement la transformation linéaire optimale et la transformation orthogonale optimale. Ces algorithmes ont été obtenus en modifiant l'algorithme d'ACI ICAinf de Pham. Dans cet article, nous comparons les performances des transformations retournées par GCGsup et ICAorth avec celles de la transformée de Karhunen-Loève (TKL) en compression à haut et moyen débits. Les tests ont été effectués sur une image naturelle et une image médicale.

## 1 Introduction

Les données multimédia (photo, audio et vidéo) non compressées réclament une capacité de stockage et une bande passante considérables. En dépit des progrès réalisés en capacité de stockage, en puissance de traitement des processeurs et en performance des systèmes de communications numériques, la demande en capacité de stockage et en bande passante de transmission reste toujours supérieure au potentiel des technologies disponibles.

Le traitement des données brutes, c'est-à-dire dans le domaine visuel pour les images ou le domaine audible pour les sons, ne permet pas d'obtenir des résultats de compression satisfaisants car on ne peut pas exploiter d'une manière efficace les redondances présentes dans les signaux. Il faut donc réorganiser l'information contenue dans les données de manière à pouvoir éliminer les redondances ; on parle alors de transformation des données par des fonctions mathématiques. Les transformations habituellement utilisées en codage par transformée [2, 3] (par exemple, la transformée de Karhunen-Loève ou la transformée en cosinus discrètes) ont toutes pour conséquence de décorréler les données d'origines et d'augmenter ainsi les performances des étapes de quantification et de codage

entropique situées en aval.

L'analyse en composantes indépendantes (ACI) [4] permet quant à elle d'aller plus loin que la décorrélation en proposant un nouvel espace de représentation dans lequel les données seront aussi indépendantes que possible. Dans [1], Narozy *et al* ont montré que la recherche de la transformation linéaire optimale en codage par transformée peut se ramener, sous l'hypothèse haute résolution, à un problème d'analyse en composantes indépendantes modifié. À partir de ce nouveau point de vue, ils ont proposé deux nouveaux algorithmes, GCGsup et ICAorth, pour calculer respectivement la transformation linéaire optimale et la transformation orthogonale optimale. Ces algorithmes ont été obtenus en modifiant l'algorithme d'ACI ICAinf de Pham. Remarquons que d'autres travaux mêlant ACI et compression de données existent [5, 6].

La deuxième section de cet article présente des résultats importants en codage par transformée sous l'hypothèse haute résolution. Dans la troisième section, nous faisons un rappel des principaux résultats présentés dans [1]. Enfin, dans la partie expérimentale, nous comparons les performances des transformations retournées par GCGsup et ICAorth avec celles de la transformée de Karhunen-Loève en compression à haut et moyen débits. Les tests ont été effectués sur une image naturelle et une image médicale.

## 2 Codage par transformée

Le schéma d'un codeur-décodeur par transformée est présenté à la figure 1. Dans un codeur par transformée, on construit d'abord un vecteur  $\mathbf{X}(m) = [X_0(m), X_1(m), \dots, X_{N-1}(m)]^T$  de dimension  $N \in \mathbb{N}^*$  à partir des échantillons du signal que l'on cherche à compresser. Dans le cas d'une image, par exemple, cette dernière sera découpée en blocs de  $\sqrt{N} \times \sqrt{N}$  pixels et chacun de ces blocs sera traité individuellement. En parcourant les lignes d'un bloc de gauche à droite et de haut en bas, on peut former avec ces pixels un vecteur colonne de  $N$  échantillons. Les composantes de  $\mathbf{X}(m)$  sont généralement fortement corrélées et un

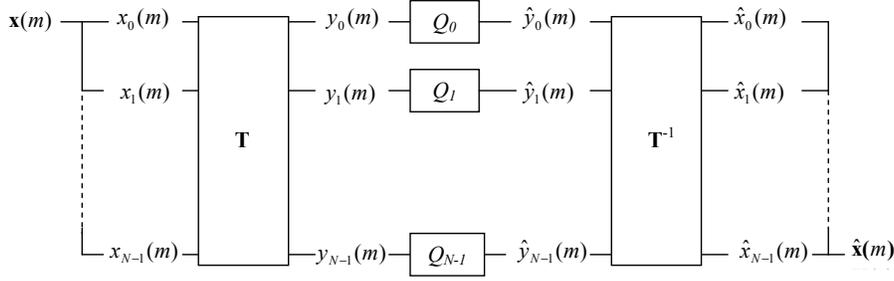


FIG. 1. Schéma d'un codeur-décodeur par transformée.

simple codage pour chacune d'elles est insuffisant. L'idée consiste à appliquer une transformation linéaire  $\mathbf{T}$  – orthogonale le plus souvent – sur le vecteur  $\mathbf{X}(m)$  pour obtenir un vecteur  $\mathbf{Y}(m) = \mathbf{TX}(m)$  ayant des composantes décorréelées qui seront ensuite codées (c'est-à-dire quantifiées et codées sans pertes à l'aide d'un codeur entropique) séparément les unes des autres. Pour obtenir une estimation du signal d'origine, la transformation inverse  $\mathbf{T}^{-1}$  est appliquée au vecteur  $\hat{\mathbf{y}}(m) = [\hat{y}_0(m), \hat{y}_1(m), \dots, \hat{y}_{N-1}(m)]^t$  pour donner le vecteur reconstruit  $\hat{\mathbf{x}}(m) = [\hat{x}_0(m), \hat{x}_1(m), \dots, \hat{x}_{N-1}(m)]^t$ . Bien qu'il n'existe aucun théorème garantissant dans le cas général une compression plus efficace avec des composantes décorréelées, il s'avère en pratique (et sous certaines conditions en théorie) qu'un codage séparé des composantes transformées et décorréelées est bien plus efficace qu'un codage direct des échantillons du signal d'entrée.

Dans cet article, nous nous plaçons dans le cas particulier — très probablement sous-optimal — où la transformation inverse correspond à l'inverse mathématique de la transformation directe. C'est une contrainte que nous nous imposons, mais qui pourrait être relâchée dans le cadre d'une étude future.

**Distorsion et débit asymptotiques** Sous l'hypothèse haute résolution [2], on montre que l'erreur quadratique moyenne entre le signal d'origine et le signal reconstruit est donnée par [1] :

$$D_{\mathbf{T}} = \frac{1}{N} E \left[ \left\| \mathbf{X} - \hat{\mathbf{X}} \right\|^2 \right] \quad (1)$$

$$\approx \frac{1}{N} \sum_{i=1}^N w_i c_i \sigma_{Y_i}^2 2^{-2R_i}, \quad (2)$$

où  $R_i$  désigne le nombre de bits alloué au quantificateur de la  $i^{\text{ème}}$  composante transformée  $Y_i$ ,  $\sigma_{Y_i}^2$  est la variance de  $Y_i$ ,  $w_i$  est égal au carré de la norme de la  $i^{\text{ème}}$  colonne de  $\mathbf{T}^{-1}$  et  $c_i = \frac{1}{12} 2^{h(\tilde{Y}_i)}$  où  $h(\tilde{Y}_i)$  désigne l'entropie différentielle de la variable aléatoire  $\tilde{Y}_i = (Y_i - E[Y_i]) / \sigma_{Y_i}$ .

D'autre part, on montre que le débit à allouer dans la  $i^{\text{ème}}$  composante transformée de façon à minimiser l'erreur

quadratique moyenne de reconstruction sous la contrainte d'un débit moyen de sortie cible  $R_{cible}$  est donné par [1] :

$$R_i \approx R_{cible} + \frac{1}{2} \log_2 \frac{w_i c_i}{\left( \prod_{i=1}^N w_i c_i \right)^{1/N}} + \frac{1}{2} \log_2 \frac{\sigma_{Y_i}^2}{\left( \prod_{i=1}^N \sigma_{Y_i}^2 \right)^{1/N}}. \quad (3)$$

En remplaçant  $R_i$  dans (2) par l'expression précédente, on obtient l'expression de la distorsion asymptotique en fonction du débit cible :

$$D_{\mathbf{T}} \approx \left( \prod_{i=1}^N w_i \right)^{\frac{1}{N}} \left( \prod_{i=1}^N c_i \right)^{\frac{1}{N}} \left( \prod_{i=1}^N \sigma_{Y_i}^2 \right)^{\frac{1}{N}} 2^{-2R_{cible}}. \quad (4)$$

de laquelle on déduit immédiatement — par inversion — l'expression du débit asymptotique en fonction de la distorsion cible :

$$R_{\mathbf{T}} \approx \frac{1}{N} \sum_{i=1}^N h(Y_i) + \frac{1}{2N} \sum_{i=1}^N \log_2 w_i - \frac{1}{2} \log_2 12 D_{cible}. \quad (5)$$

**Gain de codage généralisé** Le gain de codage généralisé est défini comme le rapport entre les distorsions  $D_{\mathbf{I}}$  (où  $\mathbf{I}$  désigne la transformation identité) et  $D_{\mathbf{T}}$  :

$$G^* = \frac{D_{\mathbf{I}}}{D_{\mathbf{T}}} = \frac{\left( \prod_{i=1}^N c_i^* \right)^{\frac{1}{N}} \left( \prod_{i=1}^N \sigma_{X_i}^2 \right)^{\frac{1}{N}}}{\left( \prod_{i=1}^N w_i \right)^{\frac{1}{N}} \left( \prod_{i=1}^N c_i \right)^{\frac{1}{N}} \left( \prod_{i=1}^N \sigma_{Y_i}^2 \right)^{\frac{1}{N}}}. \quad (6)$$

Ce critère permet de quantifier la performance d'une transformation linéaire intervenant dans un codeur par transformée utilisant des quantificateurs scalaires haute résolution et effectuant une allocation optimale de débit entre les différentes composantes transformées. Aucune hypothèse sur le caractère gaussien ou non du signal n'est faite.

Remarquons que lorsque le signal d'origine est gaussien de variance  $\sigma_X^2$ , et pour une transformation orthogonale (dans ce cas tous les coefficients  $w_i$  sont égaux à 1), on retrouve le gain de codage « classique » [2] :

$$G^* = \frac{\sigma_X^2}{\left( \prod_{i=1}^N \sigma_{Y_i}^2 \right)^{\frac{1}{N}}}. \quad (7)$$

**Réduction maximale de débit généralisée** La réduction maximale de débit généralisée est définie comme la différence entre les débits  $R_I$  et  $R_T$  :

$$R_{max} = R_I - R_T. \quad (8)$$

Cette quantité mesure la réduction du débit de sortie que l'on peut espérer en utilisant la transformation  $\mathbf{T}$  dans un codeur par transformée utilisant des quantificateurs scalaires haute résolution et effectuant une allocation optimale de débit entre les différentes composantes transformées.

On montre [1] que la réduction maximale de débit généralisée peut se mettre sous la forme suivante :

$$R_{max} \approx \frac{1}{N} I(X_1; X_2; \dots; X_N) - \frac{1}{N} I(Y_1; Y_2; \dots; Y_N) - \frac{1}{2N} \log_2 \left( \frac{\det \text{Diag}(\mathbf{T}^{-t} \mathbf{T}^{-1})}{\det \mathbf{T}^{-t} \mathbf{T}^{-1}} \right) \quad (9)$$

$$= \frac{1}{2} \log_2 G^*, \quad (10)$$

où  $I(X_1; X_2; \dots; X_N)$  (resp.  $I(Y_1; Y_2; \dots; Y_N)$ ) représente l'information mutuelle entre les composantes du vecteur d'origine (resp. du vecteur transformé).

### 3 Transformation optimale et analyse en composantes indépendantes

Dans notre étude la transformation optimale est définie comme la transformation linéaire  $\mathbf{T}$  qui maximise le gain de codage généralisé, ou d'une manière équivalente, qui maximise la réduction de débit généralisée. D'après (10), il apparaît que chercher la transformation qui maximise  $R_{max}$  est équivalent à chercher la transformation qui minimise le critère suivant<sup>1</sup> :

$$C(\mathbf{T}) = I(Y_1; Y_2; \dots; Y_N) + \frac{1}{2} \log_2 \left( \frac{\det \text{Diag}(\mathbf{T}^{-t} \mathbf{T}^{-1})}{\det \mathbf{T}^{-t} \mathbf{T}^{-1}} \right). \quad (11)$$

**Lien avec l'ACI** Le critère d'optimisation (11) peut se décomposer en la somme de deux critères distincts :

$$C(\mathbf{T}) = C_{ACI}(\mathbf{T}) + C_O(\mathbf{T}) \quad (12)$$

où

$$C_{ACI}(\mathbf{T}) = I(Y_1; \dots; Y_N), \text{ et } C_O(\mathbf{T}) = \frac{1}{2} \log_2 \frac{\det \text{Diag}(\mathbf{T}^{-t} \mathbf{T}^{-1})}{\det \mathbf{T}^{-t} \mathbf{T}^{-1}}. \quad (13)$$

Le premier terme  $C_{ACI}(\mathbf{T})$  correspond au critère de l'information mutuelle habituellement utilisé en ACI. Comme le critère d'optimisation (global) est invariant par changement d'échelle, on peut toujours s'arranger pour que les colonnes de  $\mathbf{T}^{-1}$  soient de norme unité. Dans ce cas, le terme  $C_O(\mathbf{T})$  apparaît comme un terme de pénalisation qui pénalise les transformations non orthogonales. En général, la transformation linéaire optimale en codage par transformée est différente de la solution de l'ACI sauf lorsque cette dernière correspond déjà à une matrice orthogonale.

<sup>1</sup>L'information mutuelle  $I(X_1; X_2; \dots; X_N)$  étant constante quelle que soit la transformation utilisée, elle n'intervient pas dans la maximisation de la réduction de débit généralisée.

Remarquons que le critère  $C(\mathbf{T})$  est toujours supérieur ou égal à zéro, avec égalité si et seulement si la transformation  $\mathbf{T}$  est une transformation dont les colonnes sont deux à deux orthogonales et qui produit des composantes indépendantes<sup>2</sup>. Ainsi, lorsqu'une telle transformation existe, elle constitue la solution à la fois au problème d'ACI et au problème de compression. Malheureusement, la plupart des signaux rencontrés en pratique sont rarement obtenus par mélange orthogonal de signaux statistiquement indépendants.

**Influence de la mesure de distorsion** Le critère  $C(\mathbf{T})$  dépend du choix de la mesure de distorsion. Dans toute cette étude, nous avons choisi de prendre comme mesure de distorsion l'erreur quadratique moyenne entre le signal d'origine et le signal décodé. Il n'est pas étonnant que l'usage d'une telle mesure favorise les transformations orthogonales puisque celles-ci conservent l'énergie.

**Algorithmes** Dans [1], Narozny *et al* ont présenté deux nouveaux algorithmes, GCGsup et ICAorth, obtenus à partir d'une version modifiée de l'algorithme ICAinf de D. T. Pham pour calculer la transformation linéaire optimale et la transformation orthogonale optimale respectivement. La méthode d'optimisation choisie consiste à effectuer un développement de Taylor au second ordre de la fonction à minimiser autour du point courant et à minimiser la forme quadratique qui en résulte. Cela constitue un pas d'itération qui est à répéter jusqu'à la convergence. Afin d'exploiter la structure multiplicative du paramètre  $\mathbf{T}$  et ainsi obtenir des simplifications intéressantes, Pham propose de travailler avec des incréments multiplicatifs de ce paramètre. Plus précisément, il propose d'effectuer le développement de Taylor du critère jusqu'au second ordre autour de  $\mathcal{E} = 0$ , puis de minimiser ce développement pour obtenir la matrice  $\mathcal{E}$  et de déplacer le point courant  $\mathbf{T}$  en  $\mathbf{T} + \mathcal{E}\mathbf{T}$ .

## 4 Résultats expérimentaux

Dans cette section, nous proposons une comparaison des performances de la transformée de Karhunen-Loève avec celles des transformations retournées par GCGsup et ICAorth en compression à haut et moyen débits. Les signaux testés sont l'image Lena et l'image médicale AH17. Le débit est mesuré à l'aide d'une estimation de l'entropie d'ordre 1 des coefficients quantifiés. La distorsion est mesurée quant à elle à l'aide du pic du rapport signal à bruit (PRSB).

La figure 2 montre les courbes débits-distorsions obtenues avec l'image Lena. Nous considérons le cas  $N = 64$ . Dans la région où l'hypothèse haute résolution est

<sup>2</sup>De cette remarque, nous déduisons immédiatement que n'importe quelle TKL est une transformation linéaire optimale pour des signaux gaussiens puisqu'une telle transformation est orthogonale et produit des composantes transformées décorréelées — or décorrélation rime avec indépendance pour des signaux gaussiens.

bien vérifiée, c'est-à-dire pour des débits supérieurs à environ 1,5 bpp, les courbes débits-distorsions obtenues avec les différentes transformations testées sont des droites parallèles de pentes environ égales à 6 dB/bit. L'hypothèse haute résolution étant bien vérifiée, nous avons effectué l'allocation de bits dans chaque composante selon la formule (3). Les courbes débits-distorsions de la figure 2 ont été obtenues en faisant varier les débits cibles entre 1,5 bpp et 4 bpp avec un pas de 0,5 bpp. Dans cette région des débits, nous constatons que, quel que soit le débit considéré, le PRSB de GCGsup (respectivement ICAorth) est environ 0,43 dB (respectivement 0,37 dB) au dessus de celui de la TKL.

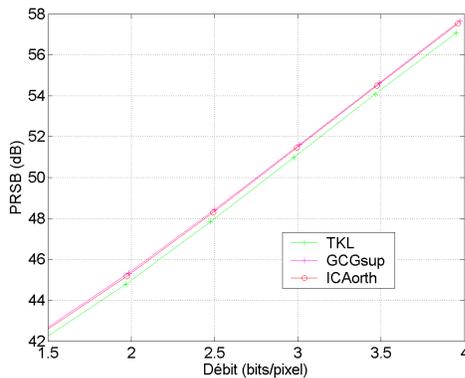


FIG. 2. Courbes débits-distorsions obtenues avec l'image Lena ( $N = 64$ ) en utilisant la TKL et les transformations retournées par GCGsup et ICAorth.

La figure 3 montre les courbes débits-distorsions obtenues avec l'image médicale AH17. Nous considérons également le cas  $N = 64$ . Dans la région où l'hypothèse haute résolution est bien vérifiée, c'est-à-dire pour des débits supérieurs à environ 1,5 bpp, nous constatons que, quel que soit le débit considéré, le PRSB de GCGsup (respectivement ICAorth) est environ 0,8 dB (respectivement 0,9 dB) au dessus de celui de la TKL.

## 5 Conclusion

Dans cet article, nous nous sommes intéressés aux performances en compression à haut et moyen débits des algorithmes d'ACI modifiés, GCGsup et ICAorth, présentés dans [1]. Dans un premier temps, nous avons présenté des résultats importants en codage par transformée sous l'hypothèse haute résolution. Ensuite, nous avons fait un rappel des principaux résultats présentés dans [1]. Enfin, dans la partie expérimentale, nous avons montré que les transformations retournées par les algorithmes GCGsup et ICAorth permettent d'obtenir de meilleures performances que la transformée de Karhunen-Loève. Dans nos tests, le débit est mesuré à l'aide d'une estimation de l'entropie

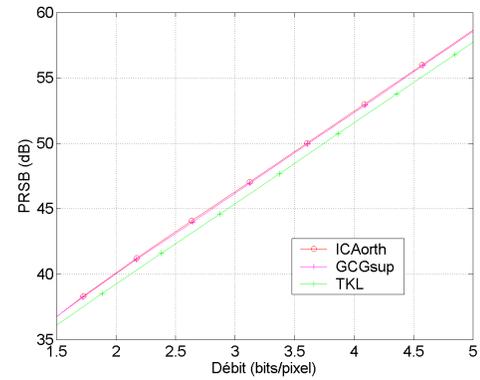


FIG. 3. Courbes débits-distorsions obtenues avec l'image médicale AH17 ( $N = 64$ ) en utilisant la TKL et les transformations retournées par GCGsup et ICAorth.

d'ordre 1 des coefficients quantifiés. La distorsion est mesurée quant à elle à l'aide du pic du rapport signal à bruit. Ces résultats ont été obtenus sur l'image Lena et l'image médicale AH17.

## Références

- [1] M. Narozny, M. Barret, D. T. Pham, I. P. Akam Bita, "Modified ICA algorithms for finding optimal transforms in transform coding", *4th Int. Symposium on Image and Signal Processing and Analysis*, Zagreb (Croatie), 15–17 sept. 2005.
- [2] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Boston, MA : Kluwer Academic Publishers, 1992.
- [3] V. K. Goyal, "Theoretical foundations of transform coding," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 9–21, Sept. 2001.
- [4] P. Comon, "Independent component analysis – a new concept?," *Signal Processing*, vol. 36, pp. 287–314, 1994.
- [5] A. T. Puga and A. P. Alves, "An experiment on comparing PCA and ICA in classical transform image coding," in *Proc. of the First Workshop on Blind Separation and Independent Component Analysis (ICA'98)*, pp. 105–108, 1998.
- [6] A. Ferreira and M. Figueiredo, "Class-adapted image compression using independent component analysis," in *IEEE International Conference on Image Processing - ICIP'2003*, Barcelona, Spain, Sept. 2003.
- [7] I. P. Akam Bita, M. Barret et D. T. Pham, "Compression of multicomponent satellite images using independent component analysis," *ICA 2006*, Charleston (USA), 5-8 March 2006.