

Prise en compte de la Chaîne Complète d'Acquisition/Compression pour le Décodage Optimal d'Images

S. Tramini⁽¹⁾, M. Antonini⁽¹⁾, M. Barlaud⁽¹⁾, G. Aubert⁽²⁾
E-Mail: tramini@i3s.unice.fr

⁽¹⁾ I3S laboratoire, CNRS UPRES-A 6070
Université de Nice-Sophia Antipolis,
2000 route des lucioles bâtiment Algorithmes/Euclide
Sophia Antipolis, 06410 Biot - France

⁽²⁾ J.A. Dieudonné laboratoire, UMR 6621,
Université de Nice-Sophia Antipolis,
Parc Valrose,
06108 Nice Cedex 2- France

Résumé : Les systèmes d'imagerie entraînent l'apparition de flou et le codeur réduit le débit binaire pour la transmission ou le codage. Ces opérations se caractérisent par une perte d'information, ainsi que par l'apparition d'artefacts. Supprimer ces artefacts permettrait un meilleur rendu visuel pour la donnée reconstruite. Contrairement aux méthodes classiques, qui effectuent séparément le décodage et le débruitage, nous proposons dans le présent article une méthode de **décodage/déconvolution conjoint** pour la reconstruction d'images. Le but de ce travail est la suppression de ces artefacts en prenant en compte l'ensemble de la chaîne d'acquisition/compression.

Abstract : Imaging systems involves blur and the coder reduces the binary rate for transmission or storage. These operations remove pertinent information contained by the image, and introduce annoying artifacts. Removing these artifacts allows higher visual quality for the reconstructed data. Unlike usual techniques, which made separately decoding and post-processing, we propose a **joint decoding/deblurring** method for image reconstruction. The goal of this work is to overcome the introduction of these negative effects by taking into account all the acquisition chain model.

1. Introduction

Les systèmes d'imagerie (caméras et appareils photo numériques, satellites) entraînent une perte de l'information due à l'optique (flou, artefacts dus aux capteurs électroniques). De plus, la capacité mémoire liée au stockage ou à la transmission étant limitée, l'utilisation d'un codeur est nécessaire afin de réduire le débit binaire. L'usage d'un codeur a pour conséquence d'engendrer la création d'artefacts nuisibles à la qualité visuelle. La suppression de ces artefacts permettrait un meilleur confort visuel dans la perception des données restaurées.

engendre. Dans cet article, nous proposons de traiter simultanément le décodage et la déconvolution sur des données bruitées ayant subies une opération de compression. Le but de ce travail est la suppression des artefacts liés à l'acquisition et la compression par la prise en compte de la globalité du système..

2. Modélisation du système traité

2.1. Notation

Notons Ω , le support de l'image f , défini comme un ensemble ouvert borné de \mathbb{R}^2 . En variables continues l'image f que l'on souhaite reconstruire peut être représentée par des fonctions définies sur $\Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$: à chaque pixel $(x, y) \in \mathbb{R}^2$ est associé son niveau de gris $f(x, y)$. En traitement d'image, l'image d'origine (notée f) est reliée à l'image observée \tilde{p} (voir **figure 1**), par une transformation qui fait intervenir deux processus décrits dans les paragraphes (2.2) et (2.3).

2.2. Processus d'acquisition

Le premier processus est relatif à **l'acquisition de l'image**. Les imperfections des capteurs de nombreux systèmes d'imagerie introduisent une dégradation que l'on peut modéliser par une convolution avec un filtre passe-bas H .

L'opérateur $H : L^2(\Omega) \rightarrow L^2(\Omega) : f(x, y) \rightarrow Hf(x, y)$, est supposé linéaire et continu. De plus, un bruit électronique b vient s'ajouter à l'image rendue floue. On peut donc modéliser la transformation permettant de passer de l'image de départ f à l'image floue et bruitée (notée f_b), par un modèle linéaire de la forme suivante :

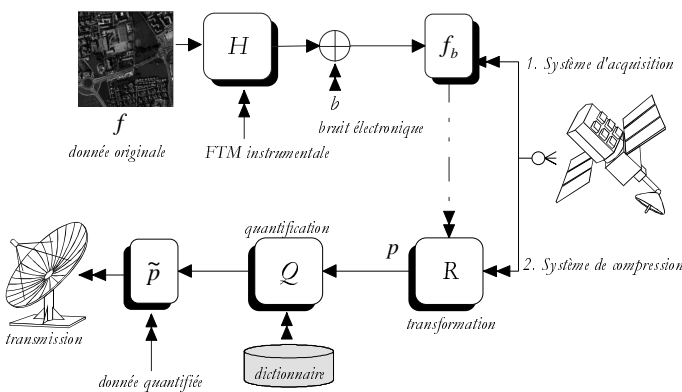


Figure 1 : représentation d'un système d'imagerie

Les méthodes classiques procèdent en deux étapes : dans un premier temps, l'image est décodée ; puis, afin de supprimer le flou et d'éliminer les artefacts, l'image subit un post-traitement. Malheureusement ces méthodes n'offrent pas de bons résultats. L'inconvénient majeur de telles approches est le fait qu'elles ne prennent pas en compte les caractéristiques du système d'acquisition/transmission et du bruit qu'il

$$f_b = Hf + b, \quad (1.)$$

Selon les experts de système d'imagerie [1], le bruit d'acquisition global b peut être assimilé à un bruit borné centré et non stationnaire. Ainsi si nous notons $-\overline{w}$ et \overline{w} les bornes du signal $b(x, y)$, nous observons au site (x, y) :

$$-\overline{w} \leq b < \overline{w} \quad (2.)$$

La variance du bruit au pixel est définie par [1] :

$$\sigma_b^2 = A + B Hf + C Hf^2 \quad (3.)$$

La connaissance de f_b , version floue et bruitée de f , nous permet de dire que l'image de départ f appartient à l'intervalle $I_b = [f_b - \overline{w}, f_b + \overline{w}]$. En utilisant les équations (2-3) et en supposant que la distribution du bruit b est uniforme, nous pouvons déterminer les bornes au site (x, y) par :

$$\overline{w} = \sqrt{3(A + B Hf + C Hf^2)} \quad (4.)$$

2.3. Processus de compression

Le second processus est relatif à l'opération de **compression de données**. Soit \tilde{p} l'image quantifiée dans le domaine transformé (voir **figure 1**). Alors, si l'on suppose que le bruit de quantification est additif [2], on peut écrire l'équation suivante :

$$\tilde{p} = \mathcal{Q}(Rf_b) = Rf_b + \varepsilon(Rf_b), \quad (5.)$$

L'opérateur $R : L^2(\Omega) \rightarrow L^2(\Omega^R) : f(x, y) \rightarrow Rf(\xi, \zeta)$, est supposé linéaire et continu (transformée en ondelettes dyadique bi-orthogonale).

La transformation R et le quantificateur \mathcal{Q} sont donnés et l'énergie du bruit de quantification ε est connue dans chaque sous-bande. Dans le présent article, nous supposons :

- l'opérateur R utilisé correspond à une transformation qui possède des bonnes propriétés de décorrélation [3],
- \tilde{p} résulte d'une quantification optimale des données transformées [4].

Si l'on considère le cas d'une quantification scalaire uniforme, la fonction de quantification est définie par

$$\mathcal{Q} : \mathbb{R} \longrightarrow q\mathbb{Z}; \quad x \longrightarrow mq \quad \text{si } x \in [mq - \frac{q}{2}, mq + \frac{q}{2}[\quad (6.)$$

où q désigne le pas de quantification et $m \in \mathbb{Z}$.

En ce qui concerne l'erreur de quantification, nous savons qu'elle appartient pour chaque sous-bande j à l'intervalle ci-dessous :

$$\left(-\frac{q}{2}\right)_j \leq \varepsilon_j < \left(\frac{q}{2}\right)_j \quad (7.)$$

Définissons pour toute sous-bande j , l'intervalle de quantification suivant (q^j désignant le pas de quantification pour une sous-bande j) :

$$I_{q,j} = [\tilde{p}_j - q_j/2, \tilde{p}_j + q_j/2[\quad (8.)$$

Pour tout échantillon appartenant à (8), l'opérateur \mathcal{Q} définit un unique représentant \tilde{p}_j . La connaissance de \tilde{p}_j , valeur quantifiée de $p_j = (Rf_b)_j$ nous permet de dire que p_j est un réel appartenant à (8).

2.4. Modèle de dégradation

En combinant ainsi les deux processus, on peut modéliser la transformation complète permettant de passer de l'image de départ f à l'image observée \tilde{p} , par un modèle additif de la forme suivante :

$$\tilde{p} = \mathcal{Q}(R(Hf + b)) = R Hf + \eta, \quad (9.)$$

où $\eta = Rb + \varepsilon$ représente le bruit de dégradation dû à la fois à la quantification et aux défauts du système d'acquisition. L'idée est ici de considérer conjointement les deux transformations de compression et d'acquisition. Dans cette optique, nous étendons nos deux intervalles précédemment définis I_q et I_b à un unique intervalle I . Ce dernier modélise à la fois l'erreur de quantification et le bruit uniforme non stationnaire. La difficulté réside en le fait de borner la transformée d'un signal b non stationnaire. Nous considérons dans cet article la variance moyenne au pixel du signal b donnée par :

$$\sigma_{b, moyenne}^2 = \frac{1}{|\Omega|} \int_{\Omega} A + B Hf + C Hf^2(x, y) dx dy \quad (10.)$$

Nous définissons tout comme dans le paragraphe (2.2) des bornes ; cependant cette fois-ci les bornes ne dépendent pas du site (x, y) .

$$\overline{w}_{moyenne} = \sqrt{3 \sigma_{b, moyenne}^2} \quad (11.)$$

Dans ce qui suit, nous utilisons deux propriétés liées à la nature de l'opérateur R : 1) la transformée R conserve l'énergie :

$$\sigma_{b, moyenne}^2 = \sigma_{Rb, moyenne}^2 \quad (12.)$$

2) la propriété d'orthogonalité nous donne l'égalité suivante :

$$\sigma_{Rb, moyenne}^2 = \sum_j \pi_j \sigma_{Rb, j}^2 \quad (13.)$$

L'hypothèse faite ici consiste à utiliser les propriétés de la décomposition de façon à exprimer la distorsion totale comme une combinaison linéaire des distorsions dans chaque sous-bande j . La condition d'orthogonalité n'étant ici pas satisfaite, une correction par un facteur de pondération π_j est nécessaire si l'on souhaite minimiser l'erreur quadratique moyenne [5] (*Si la décomposition est orthogonale, les coefficients de pondération π_j valent l'unité*).

Si l'on considère que dans une sous bande la distribution du signal transformé est de type Gaussienne généralisée, nous pouvons déterminer les bornes $\overline{w}_{R,j}$ du signal transformé dans chaque sous bande j .

En utilisant (1), (4), (5), (7) et (11) à (13), nous pouvons définir le nouvel intervalle au site (ξ, ζ) :

$$I_j = [\tilde{p}_j - \theta_j, \tilde{p}_j + \theta_j] \quad (14.)$$

où les bornes θ_j sont définies pour une sous-bande j par :

$$\theta_j = \frac{q_j}{2} + \bar{\omega}_{R,j} \quad (15.)$$

3. Déconvolution-Décodage optimal

3.1. Optimisation sous contrainte

Le terme *problème inverse* consiste à inverser la transformation représentée en figure 1, autrement dit : retrouver l'image de départ f à partir de l'image observée \tilde{p} . Ce problème inverse est dit mal posé (au sens d'Hadarnard) du fait que la seule connaissance des données observées est insuffisante pour assurer l'existence, l'unicité et la stabilité de la solution. Il est donc nécessaire d'introduire une contrainte supplémentaire, un a priori sur la solution. C'est la régularisation.

Ainsi, l'image estimée est donnée par :

$$\hat{f} = \arg \max_{\mu \in \mathbb{R}^{2^+}} \min_{f \in \Omega} J(f, \mu) \quad (16.)$$

où la fonctionnelle J est la somme d'un terme d'attache aux données, et de termes de régularisation. Ainsi, l'idée principale est de minimiser le critère suivant en f :

$$\begin{aligned} J(f, \mu) = & \sum_j \int_{\Omega_j^2} 2\sigma_{\epsilon,j}^{-2} \pi_j (RHf - \tilde{p})^2(\xi, \zeta) d\xi d\zeta \\ & + \int_{\Omega} (G * \lambda)(x, y) \psi(|\nabla f|)(x, y) dx dy \\ & + \sum_j \sum_{i=1}^2 \int_{\Omega_j^k} \mu_{i,j} g_{i,j}(f) \end{aligned} \quad (17.)$$

$$\text{avec} \quad \begin{cases} g_{1,j}(f) = (RHf - \tilde{p})_j - (q_j + 2\theta_j)/2 < 0, \\ g_{2,j}(f) = (\tilde{p} - RHf)_j - (q_j + 2\theta_j)/2 \leq 0 \end{cases} \quad (18.)$$

Le premier terme exprime la connaissance que nous avons du modèle de formation des données. Nous introduisons dans le premier terme de la fonctionnelle une pondération qui permet de prendre en compte la non stationnarité du bruit, ainsi que ses caractéristiques dans les différentes sous-bandes. $\sigma_{\epsilon,j}^2$ correspond à la puissance du bruit de quantification dans la sous-bande j introduite lors de la quantification au codeur. De même que dans le paragraphe 2.4, la condition d'orthogonalité n'étant ici pas satisfaite, une correction par un facteur de pondération π_j est nécessaire.

Le second terme exprime les hypothèses faites sur la solution afin de régulariser le problème. Cette contrainte permet de préserver les contours tout en lissant les zones homogènes, [6;7;8;9]. La fonction de potentiel ψ est supposée convexe. Le paramètre λ a été introduit comme paramètre de régularisation afin de contrebalancer les effets du

terme d'attache aux données et des a priori sur la solution. Nous considérons ce paramètre comme une fonction dépendant du pixel. Ce paramètre permettra de réaliser une régularisation différente suivant les zones d'intérêt (dans ce cas précis, nous en distinguons deux : les zones de textures et les zones homogènes). Cela nous permettra de définir un lissage fort pour les zones homogènes et un lissage faible pour les zones de textures. Cette carte de λ peut être déterminée de diverses façons; la méthode choisie est un traitement morphologique sur le gradient (combinaisons de dilatation et d'érosion).

La fonction λ étant non différentiable, nous la régularisons par une convolution avec une fonction Gaussienne G .

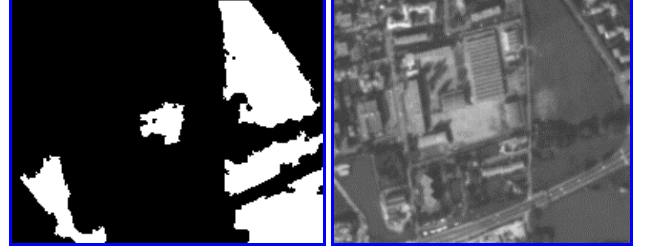


Figure 2: Image λ obtenue après traitement sur le gradient de l'image Nîmes floue et bruitée (zone noire = texture, zone blanche = homogène).

Le troisième terme nous assure de trouver une solution f de telle sorte que RHf appartienne au « bon » intervalle (15) et de conserver ainsi le même train binaire qu'après codage. De plus l'introduction de cette contrainte revient à étudier le critère dual (16), et de par la même minimiser (17) en f et maximiser (17) en μ . L'étude de ce critère est effectuée dans [10].

3.2. Equations normales

La solution optimale est obtenue lorsque

$$\frac{1}{2} \frac{\partial}{\partial f} J(f, \hat{\mu}) \Big|_{f=\hat{f}} = 0 \quad (19.)$$

Nous en déduisons les équations d'Euler-Lagrange :

$$\begin{aligned} H^* R^* \frac{1}{2\sigma_{\epsilon}^2} RH\hat{f} - (G * \lambda) \operatorname{div} \left(\frac{\psi'(|\nabla \hat{f}|)}{(|\nabla \hat{f}|)} \right) \\ - (G * \lambda) \left(\frac{\psi'(|\nabla \hat{f}|)}{(|\nabla \hat{f}|)} \right) \nabla \hat{f} + H^* R^* (\mu_1 - \mu_2) = H^* R^* \frac{1}{2\sigma_{\epsilon}^2} \tilde{p} \end{aligned} \quad (20.)$$

avec les conditions de Neumann aux bords.

Ainsi, une solution à ce système d'équations peut être obtenue en utilisant un algorithme basé sur une régularisation semi-quadratique [7] couplé à une méthode de descente du gradient de type UZAWA [10].

4. Résultats expérimentaux

Nous présentons ici des résultats en restauration d'image satellite entachée d'un bruit électronique issu de l'acquisition et d'un bruit de quantification lié à la compression. Les résultats expérimentaux ont été

effectués sur l'image Nîmes de taille 728x728 pixels codée sur 10 bits/pixel (bpp). La fonction de transfert utilisée est une donnée confidentielle fournie par le constructeur (Centre Nationale d'Étude Spatiale) [1]. L'image Nîmes floue et bruitée est une donnée simulée par le CNES. La transformation R utilisée est la transformée en ondelettes dyadique avec les filtres 9-7 [3]. Le taux de compression atteint est de 12.6 : 1. Les images sont codées par quantification scalaire en utilisant l'algorithme optimal d'allocation des bits défini en [4]. La figure 3 correspond à un décodage "classique", la figure 4 correspond à la figure 3 déconvoluée par [7], et la figure 5 correspond à notre algorithme de décodage/déconvolution conjoint optimal.

5. Conclusion

Ces résultats illustrent le comportement de la méthode proposée :

- Elle prend en compte la globalité de la chaîne d'acquisition / compression.
- Elle abandonne la notion de bruit Gaussien par l'introduction de la notion de bruit borné.
- Elle est adaptée à la statistique d'un bruit local.
- Elle est spatialement adaptative (approche locale).
- Elle apporte un gain visuel (voir figure 6).

Remerciements : *Les auteurs souhaitent remercier le CNES pour avoir fourni l'image de Nîmes ainsi que des renseignements sur la chaîne d'acquisition satellitaire.*

6. références

- [1] B. Rougé, technical report, 1996. CNES. Toulouse, France.
- [2] A. Gersho, R. Gray, "Vector Quantization and Signal Compression", Boston: Kluwer Academic Publishers, 1990.
- [3] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies "Image Coding Using Wavelet Transform", IEEE Transaction on Image Processing, Vol.1, No.2, 1992, pp. 205-220.
- [4] P. Raffy, M. Antonini, M. Barlaud, "Optimal subband bit allocation procedure for very low bit rate image coding", Electronics Letters, Vol. 34 No. 7, pp 647, 2nd April 1998.
- [5] B. Usevitch, "Optimal Bit allocation for Biorthogonal Wavelet Coding", in *Proc. Data Compression Conference* (Snowbird, Utah), pp. 387-395, March 1996.
- [6] G. Aubert, and L. Lazaroia, "A variational Method in Image Recovery", Research Note 423, Laboratoire Jean-Alexandre Dieudonné, 1995, To appear in SIAM Journal of Num. Anal.
- [7] P. Charbonnier, L. Blanc-Féraud, G. Aubert, M. Barlaud, "Deterministic Edge-Preserving Regularization in Computed Imaging", IEEE Transaction on Image Processing, Vol.5, No.12, 1997.
- [8] L.Alvarez, P.L. Lions and J.M. Morel, "Image selective smoothing and edge detection by nonlinear diffusion", SIAM J. Num. Anal. 29 (1992) 845.
- [9] S. Tramini, M. Antonini, M. Barlaud, "Intraframe Image Decoding based on a Nonlinear Variational Approach", International Journal of Imaging Systems and Technology, Vol. 9, pp. 369-380, 1998.

- [10] S Tramini, M. Antonini, M. Barlaud, G. Aubert, "Quantization Noise Removal for Optimal Transform Decoding" International Conference on Image Processing, Chicago, Illinois, USA, October 1998, Vol.3.



Figure 3 : Décodage de Nîmes floue et bruitée en utilisant le décodeur "classique"



Figure 4 : déconvolution de l'image de la figure 3 en utilisant une méthode semi-quadratique [8]



Figure 5 : Restauration optimale : décodage/déconvolution simultané