

Diffusion géométrique pour le masquage d'erreurs de quantification et de transmission sur des images JPEG couleur

Pascal BOURDON, Bertrand AUGEREAU, Christian CHATELLIER, Christian OLIVIER

Laboratoire SIC, FRE-CNRS 2731
Bât SP2MI, Bvd Marie et Pierre Curie, BP 30179,
F-86962 Futuroscope Chasseneuil Cedex, France
Tel : (33) 5 49 49 65 95 ; Fax : (33) 5 49 49 65 70

{bourdon, augereau, chatellier, olivier}@sic.univ-poitiers.fr

Résumé – Nous proposons ici une méthode de masquage d'erreurs de transmission *basée contenu*. Contrairement aux schémas classiques de correction d'erreurs de type FEQ/ARQ, notre méthode ne nécessite pas l'ajout de données de contrôle, et exploite directement la redondance spatiale de l'image source. Elle consiste en effet à interpoler les zones valides de l'image reçue dans les zones corrompues, à l'aide d'un processus de diffusion sous contraintes géométriques couplé à une approche multi-résolution. Outre les erreurs de transmission, ce modèle de diffusion nous permet de masquer également les erreurs de quantification (artefacts de compression).

Abstract – In this paper, we propose a *content-based* transmission error concealment method. Unlike classical FEC/ARQ error correction schemes, our method requires no extra control data, and take advantage of spatial redundancies contained in the source image. It consists in interpolating error-free decoded spatial information into corrupted areas, using a coupled multi-resolution/geometry-driven diffusion process. Beside transmission errors, this diffusion model can also be used to conceal quantization errors (compression artifacts).

1 Introduction

Malgré l'efficacité évidente de nouveaux standards de compression d'images, basés notamment sur des techniques de décomposition en ondelettes, comme JPEG2000 ou SPIHT, JPEG demeure l'un des standards les plus utilisés en télécommunications. Cependant, un problème récurrent dans tout système de communications est la perte ou dégradation possible des informations émises. Si aucune précaution n'est prise pour garantir une certaine qualité de service, l'un des inconvénients majeurs d'une transmission JPEG sur canal bruité est que la présence d'erreurs, aussi faible le Taux d'Erreurs Binaires (TEB) soit-il, peut avoir des conséquences désastreuses sur la qualité de l'image décodée (problèmes de désynchronisation, blocs manquants, etc) [1]. La raison d'une telle sensibilité aux erreurs est l'utilisation de codes à longueur variable. Bien que ces codes permettent d'atteindre des taux de compression très élevés, ils introduisent également une forte dépendance entre les mots de code. Du fait de cette dépendance, une simple erreur binaire aura tendance à se propager, et à rendre impossible la reconstruction correcte des données jusqu'à ce qu'un marqueur de synchronisation soit trouvé, en supposant que de tels marqueurs aient été insérés dans le flux de données. A titre d'exemple, nous pouvons observer, sur la figure 1, le résultat d'une simulation de transmission JPEG sur canal gaussien ($SNR = 11.0dB$, modulation MAQ-16) sans (Fig. 1.a), et avec (Fig. 1.b) marqueurs de synchronisation.

Comme nous pouvons le constater sur la figure 1.b, bien que l'utilisation de marqueurs de synchronisation permette de limiter la propagation des erreurs de transmission, elle ne les corrige pas pour autant. C'est pourquoi nous proposons ici un

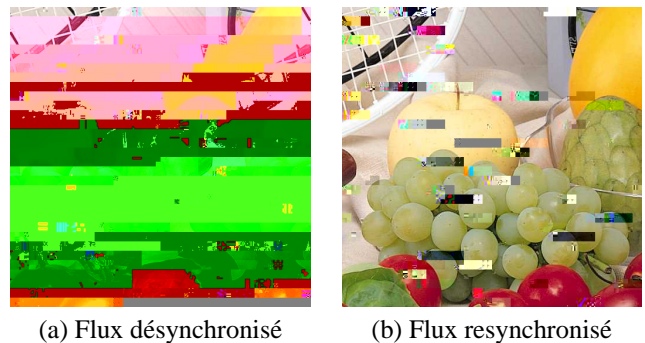


FIG. 1: *Resynchronisation de flux JPEG erronés* ($TEB \approx 6.0e^{-4}$)

algorithme de post-traitement, dont l'objectif n'est pas la restitution d'un flux binaire identique bit pour bit au flux émis, mais la restitution d'une image visuellement fidèle à l'originale (nous parlerons donc de *masquage* d'erreurs, et non de correction). Celui-ci consiste à interpoler les informations spatiales de l'image décodées sans erreurs dans les zones corrompues, en exploitant leurs propriétés géométriques [2]. Outre JPEG, cette méthode peut d'ailleurs s'appliquer à d'autres schémas de compression par blocs, comme MPEG.

Cet article est organisé de la façon suivante : dans la section 2, nous définissons le principe d'une diffusion sous contraintes géométriques et présentons notre modèle de reconstruction, qui se base sur le modèle d'*inpainting* par Variation Totale de Chan et Shen [3]. Puis, nous explicitons dans la section 3 le fonctionnement de notre approche multi-résolution. Enfin, des résultats de simulation seront proposés dans la section 4.

2 Diffusion géométrique

2.1 Principe

La reconstruction automatique de données manquantes, ou *inpainting*, se révèle comme une nouvelle tendance dans le domaine du traitement d'images par Equations aux Dérivées Partielles (EDP) [4, 3]. Il s'agit de reconstituer l'intégralité d'une ou de plusieurs zones perdues d'une image, en utilisant les données valides présentes au voisinage de ces zones. Bien que de nombreuses méthodes de reconstruction de données par EDP se basent aujourd'hui sur des modèles du 3^{eme} ordre, plus performants, les modèles du 2nd ordre, proches des modèles EDP de diffusion sous contraintes géométriques [5], peuvent eux aussi donner de bons résultats en *inpainting* [6]. De plus, il a été démontré qu'un modèle de diffusion par courbure moyenne (2nd ordre) était équivalent à une opération de filtrage médian [7]. Cette propriété nous semble particulièrement intéressante, puisqu'elle nous permet de proposer une approximation simple de notre modèle, à base de filtres médians, laquelle pourrait être implantée facilement sur carte DSP.

Soit $I(x, y) = I$ une image fixe en niveaux de gris, représentée par une fonction de $\Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ qui associe au pixel $(x, y) \in \Omega$ son niveau de gris $I(x, y)$; Ω est le domaine de définition de l'image. L'EDP de régularisation de Perona-Malik peut s'écrire de la façon suivante :

$$\begin{cases} \frac{\partial I(x, y, t)}{\partial t} = c_\xi I_{\xi\xi} + c_\eta I_{\eta\eta} \\ I(x, y, 0) = I_0(x, y) \end{cases} \quad (1)$$

où I_0 définit une version bruitée de l'image I , $I_{\xi\xi}$ et $I_{\eta\eta}$ les dérivées secondes de I dans les directions orthogonales ξ et η , et ξ et η respectivement les directions du gradient et de son vecteur orthogonal.

Une propriété intéressante du vecteur unitaire ξ est qu'en chaque point (x, y) , celui-ci est tangent à la courbure du contour de l'image. Ainsi, pour $c_\eta = 0$, l'équation (1) décrira la diffusion tangentielle pondérée (par c_ξ) de l'intensité lumineuse $I(x, y)$ le long d'un contour, durant un temps t . Intuitivement, la direction ξ nous permettrait donc d'interpoler des structures existantes à l'intérieur de zones manquantes, comme nous pouvons l'observer sur la figure 2.

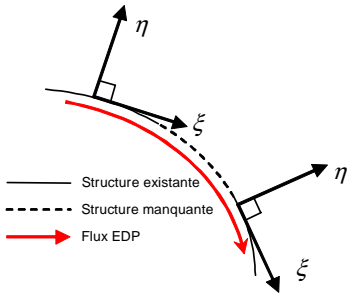


FIG. 2: Interpolation par diffusion le long d'une courbure

2.2 Modèle d'inpainting par Variation Totale

Dans [3], Chan et Shen proposent un modèle d'*inpainting* basé sur l'approche de restauration d'images par minimisation de Variation Totale (TV : *Total variation*) de Rudin *et al* [8]. Soit D la région de l'image à restaurer. $I_0|_{\Omega \setminus D}$ définit la portion incomplète, observée ou mesurée, de l'image originale sur le domaine Ω . Chan et Shen posent alors leur problème de reconstruction dans un cadre bayésien, et le présentent comme un problème de minimisation d'énergie. L'équation d'Euler-Lagrange qui lui est associée est la suivante :

$$\frac{\partial I}{\partial t} = \nabla \cdot \left[\frac{\nabla I}{\|\nabla I\|} \right] + \lambda_e (I - I_0) \quad (2)$$

valable sur l'intégralité du domaine Ω , où $\lambda_e = \lambda(1 - \chi_D)$ est le multiplicateur de Lagrange étendu, et χ_D la fonction caractéristique, ou masque, du domaine d'*inpainting* D .

L'équation (2) nous montre qu'un flux de chaleur purement anisotrope, orienté le long des courbures de l'image, est appliqué à l'intérieur du domaine d'*inpainting* D , dans le but de prolonger les lignes isophotes arrivant à ses frontières. Par ailleurs, nous démontrerons, à l'aide de résultats de simulation, que suivant les valeurs de λ , ce flux peut également être appliqué à l'intégralité de l'image, dans le but de masquer des erreurs de quantification (artefacts de compression).

2.3 Diffusion par tenseur

Tschumperlé *et al.* [9] ont récemment remarqué qu'une EDP basée sur des laplaciens orientés pouvait s'écrire de la façon suivante :

$$\frac{\partial I}{\partial t} = \text{trace}(\mathbf{T}\mathbf{H}) \quad (3)$$

où \mathbf{H} est la matrice hessienne de I et $\mathbf{T} \in \mathbb{R}^{2 \times 2}$ une matrice 2×2 définie positive qui définit les orientation et importance du lissage en chaque point de l'image au cours du processus de régularisation. \mathbf{T} d'ailleurs appelé *tenseur de diffusion*.

L'un des principaux avantages liés à l'utilisation de la formulation de Tschumperlé *et al.* pour implanter le modèle de *TV-inpainting* est son extension rapide au cas d'images couleur, et le fait qu'elle conduise à de meilleurs schémas de résolution numérique, comparé aux formulations basées sur l'opérateur $\text{div}(\cdot)$. Soit $\mathbf{I} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ une fonction vectorielle qui associe au pixel $(x, y) \in \Omega$ ses 3 composantes spectrales primaires $\mathbf{I}(x, y) = [I_1(x, y), I_2(x, y), I_3(x, y)]^T$ dans l'espace couleur Rouge-Vert-Bleu (RGB). Notre modèle de *TV-inpainting* couleur peut s'écrire de la façon suivante :

$$\begin{cases} \frac{\partial I_i}{\partial t} = \text{trace}(\mathbf{T}\mathbf{H}_i) + \lambda_e (I_i - I_{0_i}) \\ I_i(x, y, 0) = I_{0_i}(x, y) \end{cases} \quad \text{for } i = 1, 2, 3 \quad (4)$$

avec $\mathbf{T} = \frac{1}{\sqrt{\lambda_+^*}} \theta_-^* \theta_-^{*T}$, où λ_+^* , λ_-^* , θ_-^* , et θ_+^* sont respectivement les valeurs et vecteurs propres d'une version lissée du tenseur de structure multispectral de DiZenko $\mathbf{G} = \sum_{i=1}^3 \nabla I_i \nabla I_i^T$

(lisser le tenseur de DiZenzo nous permet d'éviter d'éventuels problèmes de stabilité théorique [10], et d'évaluer la géométrie locale à partir d'un voisinage plus étendu).

3 Un schéma multi-résolution

3.1 Motivations et principe

L'analyse multi-résolution d'une image consiste à explorer celle-ci à travers plusieurs niveaux de description (obtenus par sous-échantillonnage), de la plus faible résolution, qui représente l'image dans une version simplifiée et grossière, à des résolutions plus élevées, sur lesquelles vont apparaître des structures plus fines. Nous pouvons établir un parallèle avec la théorie de l'espace échelle, et l'utiliser pour modéliser les premières phases de la vision humaine. Nos motivations concernant l'élaboration d'un processus d'*inpainting* multi-résolution étaient les suivantes : un modèle d'*inpainting* doit certes prolonger les structures de l'image dans les zones manquantes, mais il doit le faire de la même façon qu'un observateur humain le ferait : des structures globales (basse résolution) aux plus fines (haute résolution). Cette méthode a deux avantages principaux : elle devrait d'abord nous permettre d'accélérer sensiblement le processus de reconstruction, les zones manquantes étant initialisées à chaque niveau de résolution par des valeurs *a priori* proches de la solution (dans beaucoup d'algorithmes de *inpainting*, ces zones manquantes sont initialisées par un bruit gaussien, afin d'assurer la convergence de l'algorithme vers une bonne solution). De même, le fait de considérer en premier lieu les structures globales de l'image devrait nous permettre de vérifier le principe de connectivité de la perception humaine mis en évidence par Chan et Shan [3] (Fig. 3), sans modifier pour autant notre modèle EDP.

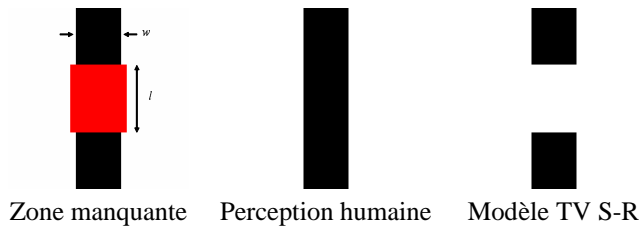


FIG. 3: Le TV-*inpainting* Simple-résolution (S-R) ne vérifie pas le principe de connectivité

3.2 Définition d'une grille multi-résolution

Notre objectif est donc de réduire les zones manquantes de l'image en réduisant sa taille, afin d'éviter les connections longue-distance. De petites zones d'*inpainting* (idéalement des zones de 1 pixel, puisque dans ce cas précis chaque pixel voisin constitue une donnée image valide) combinées à l'utilisation d'un tenseur de structure lissé devraient nous permettre de calculer des flots de diffusion cohérents. Nous pourrions alors restaurer les structures globales, et utiliser ce résultat afin d'initialiser les prochains niveaux de résolution, jusqu'à la restauration du dernier niveau (image pleine résolution).

Les transferts entre grilles de différentes résolutions seront

définis à l'aide des fonctions $(\cdot)_{\downarrow 2}$ et $(\cdot)_{\uparrow 2}$: la fonction $(\cdot)_{\downarrow 2}$ dénote l'opérateur de restriction, qui déplace la *matrice image* (nous sommes à présent dans le cas discret) d'une grille fine à une grille grossière, tandis que la fonction $(\cdot)_{\uparrow 2}$ dénote l'interpolation, qui consiste à effectuer l'opération inverse. Dans notre cas, les transferts inter-grilles vont dépendre de la présence ou non de données valides. L'opération de restriction est effectuée à l'aide d'une méthode de *full weighting* : celle-ci consiste à remplacer 4 pixels par un seul dont la valeur résulte d'un calcul de moyenne. Dans notre approche, les points de l'image appartenant à D ne seront pas pris en compte dans les calculs de moyenne. L'opération d'interpolation, quant à elle, ne consiste qu'à initialiser les parties manquantes des grilles fines par une simple prolongation (*sample and hold*) des grilles grossières (Fig. 4).

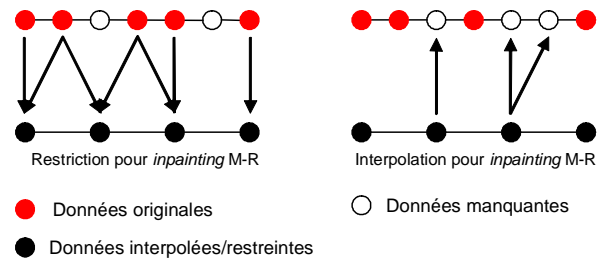


FIG. 4: Méthodes de restriction et d'interpolation

Notre algorithme fonctionne de la façon suivante :

- La fonction caractéristique χ_D du domaine D est sous-échantillonnée jusqu'à ce qu'elle n'indique plus aucune zone de *inpainting*, ou qu'elle ait atteint sa taille minimale. Nous notons L le nombre total de niveaux de résolution;
- La matrice \mathbf{I} est restreinte $L-1$ fois. Le niveau $L-1$ sera le plus petit niveau de résolution et $\mathbf{I}_{L-1} = (\mathbf{I})_{\downarrow 2^{L-1}}$ la représentation de l'image qui lui est associée;
- A moins qu'elle ne contienne aucune zone d'*inpainting*, l'image \mathbf{I}_{L-1} est restaurée à l'aide du modèle de l'équation (4);
- L'image \mathbf{I}_{L-1} restaurée est utilisée pour initialiser l'image de résolution supérieure $\mathbf{I}_{L-2} = (\mathbf{I}_{L-1})_{\uparrow 2}$. Chaque image \mathbf{I}_l sera initialisée à l'aide de l'image \mathbf{I}_{l+1} avant diffusion, jusqu'à $l = 0$.

4 Résultats de simulation

4.1 Masquage d'erreurs de transmission

La figure 5 nous montre un exemple de résultat obtenu par notre algorithme de reconstruction ($\lambda = 1$) sur l'image JPEG "Tiffany" (512×512 , couleur) pour une transmission 16-QAM sur un canal gaussien ($SNR = 12.0dB$, 4.6% de données perdues). Comme nous pouvons le constater, malgré la perte de blocs de pixels de dimensions 16×16 , l'algorithme a réussi à récupérer une bonne partie des structures corrompues. Des mesures de PSNR viennent par ailleurs confirmer la qualité du résultat, celui-ci passant de 22,69dB à 29,55dB.

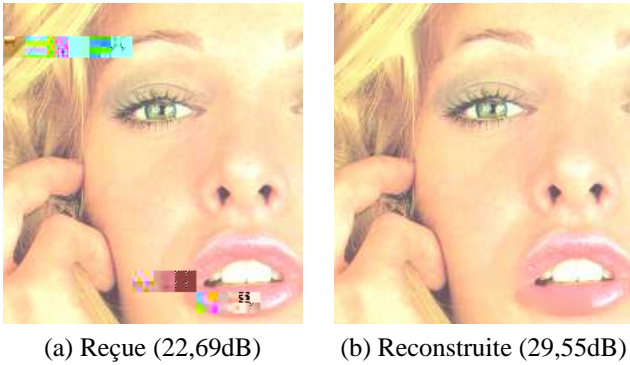


FIG. 5: Reconstruction de l'image "Tiffany" (plan rapproché)

4.2 Masquage d'erreurs de transmission et de quantification

Dans cette section, nous montrons que dans le cas d'un codage bas débit, connu pour engendrer des erreurs de quantification visibles à l'oeil humain, il est possible d'utiliser les propriétés de régularisation de notre modèle pour masquer de telles erreurs. Dans notre exemple, l'image "Peppers" (512×512 , couleur) a été encodée à l'aide d'un codeur JPEG à un débit de 0.4bpp . Afin de coupler *inpainting* et régularisation, la valeur du multiplicateur de Lagrange λ est passée de 1 to 0,3 au cours du processus de diffusion pleine-résolution (niveau 0). La figure 6 nous montre un résultat obtenu pour une transmission 16-QAM sur canal gaussien ($SNR = 11.0\text{dB}$, 6.1% de données perdues). Nous remarquons alors qu'au delà de la reconstruction des zones corrompues, l'orientation des flots de diffusion le long des lignes d'isophotes de l'image nous a permis de régulariser ses structures locales, et atténuer ainsi les effets de blocs et effets de Gibbs. Cette fois, le PSNR est passé de 20,96dB à 25,73dB.

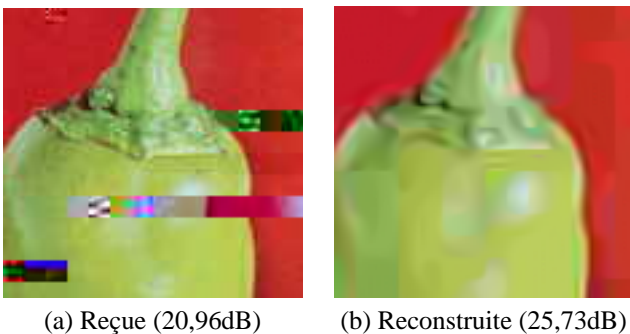


FIG. 6: Reconstruction et régularisation de l'image "Peppers" (plan rapproché)

5 Conclusion

Dans cet article, nous avons présenté une nouvelle méthode de masquage d'erreurs de transmission sur des images JPEG couleur par diffusion multi-résolution sous contraintes géométriques. Notre algorithme s'inspire d'un modèle d'*inpainting* par Variation Totale, qui nous permet d'élaborer un processus d'interpolation directionnelle à partir d'une base mathématique robuste. En interpolant les informations de voisinage d'une zone

corrompue de l'image par diffusion géométrique, nous avons pu en effet restaurer une bonne partie des structures qui la composent. De même, l'utilisation d'une approche multi-résolution nous permet non-seulement d'accélérer le processus de diffusion, mais aussi d'effectuer des connexions longue-distance, en reproduisant un comportement similaire à celui d'un observateur humain (c-à-d des structures globales aux plus fines). Au delà des erreurs de transmission, nous avons montré que notre modèle pouvait aussi être utilisé pour masquer d'éventuelles erreurs de quantification (artefacts de compression), grâce à ses propriétés de régularisation. Remarquons que cette méthode ne peut pas être utilisée pour reconstruire correctement de larges zones texturées. C'est pourquoi nos travaux futurs auront pour objectif de combiner celle-ci à des méthodes de synthèse de textures. De même, une extension au cas de séquences d'images à partir d'analyses de flots optiques, ainsi que l'implantation physique de cette méthode sur carte DSP, sont à présent à l'étude.

Références

- [1] V. DeBrunner, L. DeBrunner, L. Wang, and S. Radhakrishnan, "Error control and concealment for image transmission," *IEEE Communications Society Surveys and Tutorials*, vol. 3, no. 1, pp. 2–9, 2000.
- [2] B. Haar Romeny, *Geometry-Driven Diffusion in Computer Vision*, Kluwer Academic Publishers, 1994.
- [3] T.F. Chan and J. Shen, "Mathematical models for local nontexture inpaintings," *SIAM Journal on Applied Mathematics*, vol. 62, no. 3, pp. 1019–1043, 2002.
- [4] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. 2000, pp. 417–424, ACM Press/Addison-Wesley Publishing Co.
- [5] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [6] V. Caselles, J.M. Morel, and C. Sbert, "An axiomatic approach to image interpolation," *IEEE Transactions on Image Processing*, vol. 8, no. 3, pp. 376–386, March 1998.
- [7] F. Guichard and J.M. Morel, "Introduction to partial differential equations in image processing," in *IEEE International Conference on Image Processing*, Washington, DC, 1995, p. Tutorial Notes.
- [8] S. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.
- [9] D. Tschumperlé and R. Dériché, "Diffusion PDE's on vector valued images, local approach and geometric viewpoint," *IEEE Signal Processing Magazine - Special issue on Mathematical Methods in Imaging*, vol. 19, no. 5, pp. 15–25, 2002.
- [10] L. Alvarez, P-L. Lions, and J-M. Morel, "Image selective smoothing and edge detection by nonlinear diffusion. ii," *SIAM Journal on Numerical Analysis*, vol. 29, no. 3, pp. 845–866, 1992.