Restauration d'images floutées & bruitées par une variante originale de la variation totale

Enhancement of blurred & noisy images based on an original variant of the total variation

Khalid JALALZAI[1] Antonin CHAMBOLLE[2]

[1] CMAP, École Polytechnique, 91128 Palaiseau Cedex, France.
e-mail : khalid.jalalzai@polytechnique.edu
[2] CMAP, École Polytechnique, 91128 Palaiseau Cedex, France.
e-mail : antonin.chambolle@polytechnique.fr

Résumé

Dans cet article, nous introduisons une nouvelle variante de la variation totale (TV) *dont l'objectif est de simplifier la restauration d'images à base de* TV *lorsque celles-ci sont dégradées par un noyau qui se calcule facilement du côté Fourier (flou, transformée de Radon,...). L'idée est de remplacer simplement le terme* TV *par la norme* L^1 *d'un certain champ de vecteur, pour lequel l'optimisation est beaucoup plus facile. Cette approche nous permet ainsi d'utiliser un algorithme récent et rapide pour restaurer entre autres des images bruitées et floutées. Nous comparons notre approche avec la méthode classique basée sur la variation totale et montrons sa supériorité.*

Mots Clés

Restauration d'images, variation totale, défloutage, débruitage, méthode du gradient accélérée, problème inverse linéaire, fonction à variation bornée.

Abstract

In this paper, we introduce a new variant of the total variation (TV). *Its purpose is to simplify* TV*-based restoration when the image is degraded by some kernel which is easily computed in the Fourier domain (blur, Radon transform...). We actually replace the* TV *term by a mere* L^1 *norm of some field, for which the optimization is much easier. This approach permits us to use a recent and fast algorithm to enhance, in particular, blurred and noisy images. We also compare our approach with standard total variation based denoising and show that it performs better.*

Keywords

Image reconstruction, total variation, deconvolution, deblurring, accelerated gradient scheme, linear inverse problem, bounded variation function.

1 Introduction

En 1992, Rudin, Osher et Fatemi (ROF), dans leur article fondateur [13], ont introduit la variation totale pour régulariser les solutions de problèmes inverses en imagerie. Cette approche a porté ses fruits dans la restauration d'images puisque la minimisation de la variation totale permet de lisser les images sans pour autant détruire le bord des objets. Une approche possible pour s'attaquer à la minimisation du problème de ROF consiste à utiliser la méthode dite *Implicite-Explicite* ou encore *Forward-Backward splitting* (voir Combettes-Wajs [3] par exemple). Il s'agit en fait de

$$\text{minimiser } (\varphi + \psi)$$

où φ et ψ sont toutes les deux des fonctions convexes avec certaines propriétés de régularité. Généralement dans le domaine du traitement du signal, étant donné un signal u, $\varphi(u)$ est le terme de fidélité aux données et est égal à $\frac{1}{2}\|Au - g\|^2$ où g est un signal bruité qui a également subi une perturbation linéaire A. Le second terme ψ reflète une connaissance a priori sur le bruit par exemple. Lorsque $\psi = TV$ comme dans le problème de ROF, il est en général difficile de calculer (ou d'approcher) le minimiseur à savoir l'opérateur proximal du signal perturbé $\text{prox}_{TV}(g)$ (cf. Moreau [9] ou encore une fois Combettes et Wajs [3] pour plus de détails à ce sujet).

Par conséquent, nous proposons dans cet article une variante du problème de ROF où le terme ψ est tout simplement la norme L^1 d'un certain champ de vecteur p. Ce nouveau terme préserve les belles propriétés de la variation totale. Son intérêt réside dans le fait que son opérateur proximal est facile à calculer et de sucroît il ouvre la porte à des algorithmes de type *compressed sensing* (cf. Nesterov [12] ou Beck et Teboulle [2]).

Cependant notre idée est différente des méthodes de type *Lagrangien augmenté* (cf. Tai et Wu [15]) ou *Split Bregman* (cf. Goldstein and Osher [6]) où le champ de vecteur p doit satisfaire à la limite (parfois de manière approchée) la contrainte $p = \nabla u$, alors que dans notre approche p n'a aucune raison d'être un gradient.

2 Quelques notations

Dans la suite, une image u sera une matrice $n \times n$ à entrées réelles *i.e.* un élément de $X = \mathbb{R}^{n \times n}$. Pour simplifier les choses lorsque nous serons amenés à considérer la transformée de Fourier discrète de u, nous supposerons que l'image u est en fait périodique et définie pour tout $k \in \mathbb{Z}$ par $u_{i+kn,j+kn} = u_{i,j}$ avec $i, j \in \{1, ..., n\}$.
Pour définir la variation totale de l'image u nous devons tout d'abord introduire le gradient discret de u. Il s'agit du vecteur ∇u de $Y = X \times X$ défini par

$$(\nabla u)_{i,j} = \begin{pmatrix} u_{i+1,j} - u_{i,j} \\ u_{i,j+1} - u_{i,j} \end{pmatrix},$$

pour $i, j = 1, ..., n$.

Finalement, l'approximation la plus simple de la variation totale de $u \in X$ est définie par

$$TV(u) = \sum_{i,j} |(\nabla u)_{i,j}|$$

où $|\cdot|$ est simplement la norme euclidienne dans $\mathbb{R}^2$.

On en profite également pour introduire la divergence $\operatorname{div} p$ d'un élément p de Y et le laplacien Δv d'une image $v \in X$. Par analogie avec le cadre continu, on les définit de manière à ce qu'ils satifassent pour tout $u \in X$,

$$\langle \operatorname{div} p, u \rangle_X = -\langle p, \nabla u \rangle_Y \text{ et } \Delta v = \operatorname{div} \nabla v. \qquad (1)$$

3 L'approche classique basée sur TV

Étant donnée une image g bruitée et ayant été exposée à une perturbation linéaire A, la méthode de Rudin, Osher et Fatemi suggère de minimiser la quantité

$$F(u) = \frac{1}{2}\|Au - g\|^2 + \lambda TV(u) \qquad (2)$$

pour restaurer l'image g. Le paramètre positif λ permet ici de contrôler le niveau de régularisation souhaité.

En fait, le terme TV n'étant pas différentiable, on le remplace dans la pratique par une autre approximation de la variation totale :

$$TV_\varepsilon(u) = \sum_{i,j} \sqrt{\varepsilon^2 + |(\nabla u)_{i,j}|^2}$$

où ε est un réel strictement positif. On est donc amené à trouver l'unique u_ε qui minimise

$$F_\varepsilon(u) = \frac{1}{2}\|Au - g\|^2 + \lambda TV_\varepsilon(u).$$

On se trouve ici face à un problème d'optimisation convexe qui peut se résoudre simplement grâce à une descente de gradient. Il suffit pour cela de considérer une suite d'images (u_n) et un pas $h > 0$ suffisamment petit tels que

$$u_{n+1} = u_n - h\big[A^T(Au_n - g) + \lambda \nabla TV_\varepsilon(u_n)\big]$$

où

$$\big(\nabla TV_\varepsilon(u_n)\big)_{i,j} = -\operatorname{div}\left(\frac{(\nabla u_n)_{i,j}}{\sqrt{\varepsilon^2 + |(\nabla u_n)_{i,j}|^2}}\right) \qquad (3)$$

pour $i, j = 1, ..., n$. La suite (u_n) converge alors vers le minimiseur cherché. Reste à choisir u_0 : le plus judicieux étant de le prendre proche du minimiseur, $u_0 = g$ semble être un bon choix.

Malheureusement, le schéma naïf proposé ci-dessus est assez lent puisque sa complexité est en $O\left(\frac{1}{n}\right)$ ce qui signifie qu'il existe une certaine constante réelle C telle que

$$F_\varepsilon(u_n) - F_\varepsilon(u_\varepsilon) \leq \frac{C}{n}.$$

Une preuve de ce résultat classique peut être trouvée dans [10], [11] ou encore dans [2].

Dans [11], Nesterov a proposé un algorithme en $O(\frac{1}{n^2})$ qui résout le problème. Cette stratégie peut s'écrire sous la forme algorithmique qui suit :

$$\begin{aligned}
v_n &= \arg\min \Big\{ F_\varepsilon(u_n) + \langle v - u_n, \nabla F_\varepsilon(u_n) \rangle_X \\
&\qquad + \frac{L}{2}\|v - u_n\|^2,\ v \in X \Big\}, \\
w_n &= \arg\min \Big\{ \sum_{k=0}^{n} [F_\varepsilon(u_k) + \langle w - u_k, \nabla F_\varepsilon(u_k) \rangle_X] \\
&\qquad + \frac{1}{2}\|w - u_0\|^2,\ w \in X \Big\}, \\
u_{n+1} &= \frac{2}{k+3} w_n + \frac{k+1}{k+3} v_n,
\end{aligned}$$

où L est la constante de Lipschitz de ∇F_ε.

Cet algorithme combine efficacement descente de gradient classique (pour le calcul de v_n) et méthode du gradient conjugué (calcul de w_n). Chaque itération nécessite le calcul de $\gamma_n = \nabla F_\varepsilon(u_n)$ dont le premier terme se calcule facilement du côté Fourier si c'est le cas pour A. Quant au deuxième terme il s'agit simplement de (3). Nous renvoyons à Nesterov [10] pour plus de détails sur ces techniques. Dans [2], Beck et Teboulle proposent une variante simple et rapide de cet algorithme que l'on met à profit dans la suite.

4 Une variante de TV

Étant donnée une image $u \in X$, l'idée est de remplacer le terme TV dans (2) par

$$J(u) = \min_{\substack{p \in Y \\ \Pi p = \nabla u}} \|p\|_1$$

où d'une part, $\|p\|_1 = \sum_{i,j} \sqrt{(p^1_{i,j})^2 + (p^2_{i,j})^2}$ lorsque $p = (p^1, p^2) \in X \times X$ et d'autre part, Π est l'opérateur de projection sur les gradients défini par $\Pi p = \nabla \bar{v}$, avec $\bar{v}$ qui réalise le minimum

$$\min_{v \in X} \|\nabla v - p\|. \tag{4}$$

Ici $\|\cdot\|$ est la norme euclidienne de Y. Remarquons au passage que pour $u \in X$,

$$J(u) \leq TV(u) \leq TV_\varepsilon(u).$$

Il s'agit d'une conséquence immédiate de la définition. Dans la suite (cf. section 5), nous motiverons l'utilisation de cette fonctionnelle et en détaillerons quelques autres propriétés qui nous laissent à penser qu'elle se comporte de la même manière que TV.
Revenons donc à notre problème : la solution de (4) satisfait l'équation d'Euler-Lagrange

$$\nabla^*(\nabla u - p) = 0$$

ou, en utilisant les notations introduites en (1),

$$\Delta u = \operatorname{div} p,$$

(on rappelle que nos opérateurs ∇, div et Δ sont ici des opérateurs discrets avec des conditions au bord périodiques). Par conséquent,

$$J(u) = \min_{\substack{p \in Y \\ \operatorname{div} p = \Delta u}} \|p\|_1.$$

Donc le problème de Rudin, Osher et Fatemi exprimé avec cette nouvelle fonctionnelle consiste à minimiser

$$G(p) = \frac{1}{2}\|Au - g\|^2 + \lambda \|p\|_1$$

sur l'ensemble des (p, u) qui vérifient la contrainte $\Delta u = \operatorname{div} p$.

La minimisation de ce type de fonctionnelles présente depuis quelques années un attrait tout particulier dans le domaine de la compression de données et a fait l'objet de nombreux articles. En particulier, deux articles récents de Nesterov [12] et de Beck et Teboulle [2] s'intéressent à la minimisation de fonctions qui admettent une décomposition de la forme

$$\varphi + \psi$$

où φ est une fonction convexe continûment différentiable à gradient lipschitzien et ψ une fonction continue et convexe que l'on autorise à ne pas être différentiable mais est simple au sens où son opérateur proximal est facile à calculer (voir Combettes et Wajs [3] pour la définition). Ces caractéristiques siéent parfaitement aux deux termes de G et en conséquent nous noterons dorénavant

$$\varphi(p) = \frac{1}{2}\|A\Delta^{-1} \operatorname{div} p - g\|^2 \text{ et } \psi(p) = \|p\|_1.$$

Dans leur article, Beck et Teboulle suggèrent de construire une approximation (p_n) du minimiseur de G selon le schéma suivant :

$$\begin{aligned}
&q_1 = p_0 \in Y,\ t_1 = 1, \\
&p_n = \arg\min \Big\{ \varphi(q_n) + \langle p - q_n, \nabla\varphi(q_n)\rangle_Y \\
&\qquad\qquad + \frac{L'}{2}\|p - q_n\|^2 + \lambda\|p\|_1,\ p \in Y \Big\}, \\
&t_{n+1} = \frac{1 + \sqrt{1 + 4t_n^2}}{2}, \\
&q_{n+1} = p_n + \left(\frac{t_n - 1}{t_{n+1}}\right)(p_n - p_{n-1}),
\end{aligned}$$

où L' est la constante de Lipschitz de $\nabla\varphi$.

Notons que cet algorithme ne demande qu'un seul calcul de gradient à chaque itération, si on fait bien les choses. Quant à l'algorithme proposé par Nesterov dans [12], qui est en $O(\frac{1}{n^2})$ tout comme celui de Beck et Teboulle, et qui encore une fois est une astucieuse combinaison de descente de gradient et gradient conjugué, il demande malheureusement deux calculs de gradient à chaque itération ce qui ralentit notablement son exécution.

Dans le cas présent, on peut écrire plus concrètement

$$\begin{aligned}
p_n = \arg\min \Big\{ \sum_{i,j} \Big(& p_{i,j} \cdot (\gamma'_n)_{i,j} \\
& + \frac{L'}{2}|p_{i,j} - (q_n)_{i,j}|^2 + \lambda |p_{i,j}| \Big), p \in Y \Big\},
\end{aligned}$$

où

$$\begin{aligned}
\gamma'_n &= \nabla\varphi(q^n) \\
&= (A\Delta^{-1} \operatorname{div})^*(A\Delta^{-1} \operatorname{div} q_n - g),
\end{aligned}$$

qui se calcule bien du côté Fourier si c'est également le cas pour A (flou, transormée de Radon,...). La somme précédente se minimisant terme à terme, on obtient finalement

$$(p_n)_{i,j} = \max\left(0, |(x_n)_{i,j}| - \frac{\lambda}{L'}\right) \frac{(x_n)_{i,j}}{|(x_n)_{i,j}|}$$

avec $(x_n)_{i,j} = (q_n)_{i,j} - \frac{(\gamma'_n)_{i,j}}{L'}$.

Par ailleurs, dans le cas où A est simplement l'opérateur de convolution contre une gaussienne par exemple, $L' = \frac{1}{2}(1 - \cos(\frac{2\pi}{n}))^{-1}$.

5 Le cadre continu : motivations et perspectives

Nous profitons de cette section pour mentionner quelques résultats complémentaires sur la fonctionnelle J. L'ensemble de ces propriétés trouvent leurs preuves dans Jalalzai [7]. Pour commencer, fixons quelques notations propres à ce chapitre. Ainsi, dans la suite Ω désignera un ouvert de $\mathbb{R}^n$ à bord régulier et pour simplifier les choses nous nous plaçons dans un premier temps dans le contexte des fonctions u dont les dérivées au sens des distributions sont des fonctions intégrables notées Du, *i.e.* u est un élément de l'espace de Sobolev $H^1(\Omega)$.
La fonctionnelle que nous avons précédemment présentée est alors une discrétisation de

$$J(u) = \inf \left\{ \int_\Omega |\phi|,\ \phi \in L^2(\Omega)^n \text{ et } \Pi\phi = Du \right\}.$$

où Π est la projection orthogonale sur l'espace des gradients comme en section 4. Formellement, étant donnée une fonction $\phi \in L^2(\Omega)^n$, on pose $\Pi\phi = D\bar{v}$ où $\bar{v}$ minimise

$$\min_{v \in H^1(\Omega)} \|Dv - \phi\|_{L^2(\Omega)^n}.$$

Il est facile de voir qu'il existe une unique fonction $\psi \in L^2(\Omega)^n$ telle que l'on ait la décomposition de Helmholtz suivante :

$$\phi = \Pi\phi + \psi \quad \text{ où } \int_\Omega \nabla v \cdot \psi = 0 \text{ pour tout } v \in C^1(\bar{\Omega}).$$

(nous nous référons à Dautray-Lions [4] ou Temam [14] pour plus de détails à ce sujet).
L'un dans l'autre, nous avons prouvé que

$$J(u) = \inf \left\{ \int_\Omega |Du + \psi|,\ \psi \in L^2(\Omega)^n \text{ et } \int_\Omega \nabla v \cdot \psi = 0\ \forall v \in C^1(\bar{\Omega}) \right\}.$$

Néanmoins, cette nouvelle formulation de J conserve un sens même lorsque u est simplement une fonction à variation bornée dans Ω (notée $u \in BV(\Omega)$) ce qui signifie que $TV(u) < +\infty$ ou en d'autres termes que sa dérivée au sens des distributions Du est maintenant dans $\mathcal{M}_b(\Omega)^n$, l'ensemble des mesures de Radon vectorielles sur Ω. Dorénavant, nous laisserons également ψ parcourir l'espace $\mathcal{M}_b(\Omega)^n$. Pour tout ce qui concerne la théorie des fonctions à variation bornée mais également tout ce qui touche à la théorie de la mesure, nous renvoyons à l'ouvrage d'Ambrosio, Fusco et Pallara [1] ou bien encore à celui de Giusti [5]. Tout ceci motive une nouvelle définition de J lorsque $u \in BV(\Omega)$, à savoir :

$$J(u) = \inf \left\{ \int_\Omega |Du + \psi|,\ \psi \in \mathcal{M}_b(\Omega)^n \text{ et } \int_\Omega \nabla v \cdot \psi = 0\ \forall v \in C^1(\bar{\Omega}) \right\}.$$

Nous tenons à faire remarquer au passage que $J(u)$ est bien défini lorsque $u \in BV(\Omega)$ puisque l'on a

$$J(u) \leq TV(u) = \int_\Omega |Du|. \tag{5}$$

Grâce à un argument de dualité convexe, il est possible de montrer que sous certaines hypothèses supplémentaires sur Ω, on a

$$J(u) = \sup \left\{ \int_\Omega \nabla w \cdot Du,\ w \in C^1(\Omega) \text{ et } \|\nabla w\|_\infty \leq 1 \right\}.$$

Cette formulation duale permet de démontrer le résultat suivant :
Théorème : *Soit Ω un ensemble ouvert de $\mathbb{R}^n$ et $u = \chi_E$ la fonction caractéristique d'un ensemble de périmètre fini $E \subset \Omega$, ou bien même soit $u \in BV(\Omega)$ avec une dérivée Du concentrée sur l'ensemble de saut. Alors,*

$$J(u) = \int_\Omega |Du|.$$

La preuve utilise principalement le fait que les ensembles rectifiables admettent des hyperplans tangents en un sens faible. Ce théorème légitime l'utilisation de J dans le contexte du traitement d'image (et du signal plus généralement) puisque cette fonctionnelle coïncide dans le cas des images de type "cartoon" avec le terme TV.

Par ailleurs, le résultat précédent motive la définition d'une variante de J, notons là J', où le supremum du théorème précédent n'est pris que sur les fonctions w à supports compacts *i.e.*

$$\begin{aligned} J'(u) &= \sup \left\{ \int_\Omega \nabla w \cdot Du,\ w \in C_c^2(\Omega) \text{ et } \|\nabla w\|_\infty \leq 1 \right\} \\ &= \sup \left\{ \int_\Omega w \Delta u,\ w \in C_c^2(\Omega) \text{ et } \|\nabla w\|_\infty \leq 1 \right\}, \end{aligned}$$

par Gauss-Green. J' s'annule donc sur les fonctions harmoniques (*i.e.* telles que $\Delta u = 0$) ce qui est un avantage considérable sur TV qui ne s'annule que sur les fonctions constantes. Ceci signifie que si le signal g perturbé considéré en section 3 est harmonique, il peut être reconstitué parfaitement. Ceci laisse à penser que J' mais également J peuvent éviter le fameux *staircasing effect* puisqu'ils reconstiuent bien mieux les formes "lisses". Comme nous allons le voir dans la suite, nous retrouvons ce résultat dans les simulations sur ordinateur.

6 Simulations numériques

Dans cette section, nous comparons les deux approches basées sur les fonctionnelles TV et J en utilisant respectivement les deux algorithmes présentés. Dans notre implémentation, l'algorithme de Beck et Teboulle fait deux fois moins d'itérations par unité de temps puisqu'il fait intervenir quatre transformations de Fourier par boucle (calcul

de γ'_n) alors que l'algorithme de Nesterov n'en fait intervenir que deux (calcul de γ_n). Il est certainement possible de faire mieux notamment dans le cas où l'on considère J puisque le calcul fait massivement appel aux transformées de Fourier et est par conséquent facilement parallélisable.
De plus, comme attendu, les premiers résultats numériques montrent que la fonctionnelle J semble éviter le fameux *staircasing effect* (voir l'article de Louchet et Moisan [8]) produit par la minimisation de la variation totale et qui donne des images dont la configuration locale n'est pas naturelle.
Pour tous ces tests, nous avons utilisé un ordinateur personnel muni d'un processeur Core2 Duo cadencé à 2 Ghz et chaque programma a été exécuté pendant exactement 20 secondes sous MATLAB. Par ailleurs, le paramètre de régularisation λ est toujours maintenu égal à 1.

6.1 Premier exemple

Dans un premier temps, on porte notre attention sur la photo de Lenna de taille 256×256. Celle-ci a subi un flou gaussien de variance $\sigma_{\text{flou}} = 1.5$ suivi de l'addition d'un bruit gaussien de moyenne nulle et de variance $\sigma_{\text{bruit}} = 4$. L'image originale étant représentée en figure 1. Nous avons alors implémenté et exécuté les algorithmes de Beck et Teboulle et celui de Nesterov pour tenter de restaurer cette image. Les résultats sont présentés en figure 3 et 4.

FIG. 1 – Photo originale de Lenna

FIG. 2 – $\sigma_{\text{flou}} = 1.5$, $\sigma_{\text{bruit}} = 4$

FIG. 3 – Lenna restaurée par J, 600 itérations

FIG. 4 – Lenna restaurée par TV, 1000 itérations

6.2 Second exemple

Le deuxième exemple, compare la capacité de ces deux algorithmes à restaurer un texte numérisé. L'image considérée est un poème de taille 256×256 et a subi les mêmes perturbations que ci-avant avec cette fois-ci $\sigma_{\text{flou}} = 1$ et $\sigma_{\text{bruit}} = 4$.

FIG. 5 – La première strophe d'un célèbre poème de Verlaine

FIG. 6 – σ $_{\text{flou}} = 1$, $\sigma_{\text{bruit}} = 4$

FIG. 7 – Poème restauré par J, 600 itérations

FIG. 8 – Poème restauré par TV, 1200 itérations

Références

[1] Ambrosio, L., Fusco, N., Pallara, D. : Functions of Bounded Variation and Free Discontinuity Problems. Oxford University Press, 2000.

[2] Beck, A., Teboulle, M. : A fast iterative shrinkage-thresholding algorithm for linear inverse problems. Accepted by SIAM Journal on Imaging Sciences, 2008.

[3] Combettes, P. L., Wajs, V. R. : Signal recovery by proximal forward-backward splitting. SIAM Journal on Multiscale Modeling and Simulation, vol. 4, no. 4, pp. 1168-1200, 2005.

[4] Dautray, R., Lions, J.-L. : Mathematical Analysis VI and Numerical Methods for Science and Technology, Volume 6, Evolution Problems II. Springer, 1993.

[5] Giusti, E. : Minimal Surfaces and Functions of Bounded Variation. Birkhäuser, 1984.

[6] Goldstein, T., Osher, S. : The Split Bregman Method for L1 Regularized Problems. UCLA CAAM Report 08-29, 2008.

[7] Jalalzai, K. : Étude des propriétés d'une variante de la variation totale. Mémoire de Master, 2008.

[8] Louchet, C., Moisan, L. : Total variation denoising using posterior expectation. Disponible sur la base de donnée hal à l'adresse `http://hal.archives-ouvertes.fr`, 2008.

[9] Moreau, J.-J. : Fonctions convexes duales et points proximaux dans un espace hilbertien. C. R. Acad. Sci. Paris Sér. A Math. 255, pp. 2897-2899, 1962.

[10] Nesterov, Y. : Introductory lectures on convex optimization. Kluwer Academic Publishers, 2004.

[11] Nesterov, Y. : Smooth minimization of non-smooth functions. Mathematical Programming (A), pp. 127-152, 2005.

[12] Nesterov, Y. : Gradient methods for minimizing composite objective function. CORE Report, 2007.

[13] Rudin, L.I., Osher, S., Fatemi, E. : Nonlinear total variation based noise removal algorithms. Physica D, vol. 60, pp. 259-268, 1992.

[14] Temam, R. : Navier-Stokes Equations Theory and Numerical Analysis. AMS Bookstore, 2001.

[15] Tai, X.-C., Wu, C. : Augmented Lagrangian Method, Dual Methods and Split Bregman Iteration for ROF Model. UCLA CAAM Report 09-05, 2009.