IMA203:MATIM

Méthodes variationnelles:

said.ladjal@telecom-paris.fr

Introduction des méthodes variationnelles

- Pour corriger le défaut subi par une image on utilise une énergie qui sera faible pour une image « bonne » et grande pour une image « mauvaise »
- Il y a souvent deux termes dans une telle énergie:
    - Un terme « attache aux données » qui signifie que l'image reconstruite explique l'observation.
    - Un terme « régularité » qui privilégie les images qui ont des statistiques d'images naturelles.

Un exemple de construction d’une méthode variationnelle

- Nous allons rappeler le raisonnement qui mène à la restauration par norme du gradient au carré.
  - D’abord, nous allons voir pourquoi il est nécessaire de régulariser.
  - Ensuite nous rappelons une justification probabiliste de la construction d’une énergie à minimiser.

Modélisation

Le processus direct qui mène à la formation de l'image est modélisé par:

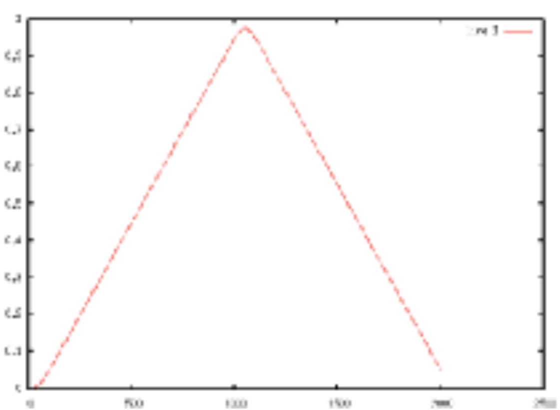
$$g = Af + b$$

L'opérateur A est linéaire et b représente le bruit.

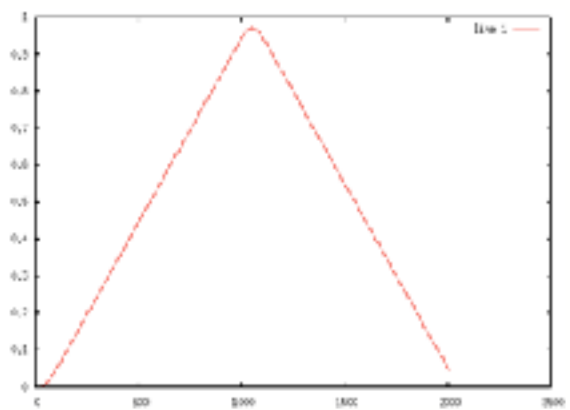
Problème inverse: Nécessité de la régularisation

- Une première idée consiste à appliquer l'opérateur inverse de A à l'équation précédente. Ici A est une petite convolution.

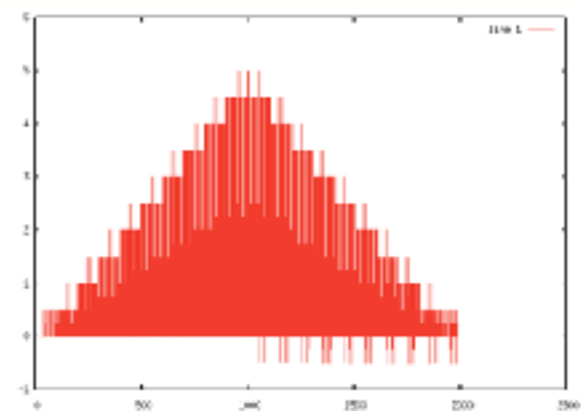
$$\tilde{f} = A^{-1} g$$



Original



Bruité



Inverse

Modélisation pour la suite

$$g = f * K + b$$

f : image parfaite

g : observation

K : Noyau de flou

b : bruit, supposé uniforme et indépendant du signal

Qu'est-ce que l'image parfaite?

Modélisation pour la suite: Continu/discret

- Le théorème de Shannon de bon échantillonnage permet un passage sans perte du discret au continu.
- Le caractère fini des images traitées implique de faire une hypothèse sur la continuation des images hors de leur support connu: On fera l'hypothèse de périodicité pour simplifier la présentation.

Modélisation pour la suite: Continu/discret

On peut donc, en négligeant le sous-échantillonnage, dire que le problème se présente comme:

$$
\begin{aligned}
&g = Af + b \\
&f \in \mathbb{R}^{N^2} : \ \text{image de taille } N \times N \text{ inconnue} \\
&g \in \mathbb{R}^{N^2} : \ \text{image de taille } N \times N \text{ connue} \\
&A : \ \text{matrice de convolution carrée de taille } N^2 \times N^2 \\
&b \in \mathbb{R}^{N^2} : \ \text{réalisation d'un bruit gaussien}
\end{aligned}
$$

L'hypothèse de périodicité implique que la matrice A est circulante
et diagonalisable dans la base de Fourier

Résolution: Wiener

- Le signal est supposé être un vecteur gaussien:

$$
\begin{aligned}
&X \text{ v.a. } \in \mathbb{R}^N \\
&\mathbb{P}(X) \sim e^{-\frac{1}{2} X^T C^{-1} X} \\
&C = E(X^T X) \text{ matrice de covariance déf. pos.} \\
&\mathbb{P}(B) \sim e^{-\frac{1}{2} \frac{\|B\|^2}{\sigma_b^2}}
\end{aligned}
$$

$Y = AX + B$

trouver $\tilde{X} = DY$ t.q.

$E(\|\tilde{X} - X\|^2)$ soit minimale

maximiser $e^{-\frac{1}{2}\tilde{X}^T C^{-1} \tilde{X}} e^{-\frac{1}{2}\frac{\|A\tilde{X} - Y\|^2}{\sigma_b^2}}$

i.e. minimiser $\frac{1}{\sigma_b^2}\|A\tilde{X} - Y\|^2 + \tilde{X}^T C^{-1} \tilde{X}$

Résolution: Wiener

- Le signal est supposé être un vecteur gaussien:

$$
\begin{aligned}
&X \text{ v.a. } \in \mathbb{R}^N \\
&\mathbb{P}(X) \sim e^{-\frac{1}{2} X^T C^{-1} X} \\
&C = E(X^T X) \text{ matrice de covariance déf. pos.} \\
&\mathbb{P}(B) \sim e^{-\frac{1}{2} \frac{\|B\|^2}{\sigma_b^2}}
\end{aligned}
$$

$$
\begin{aligned}
&Y = AX + B \\
&\text{trouver } \tilde{X} = DY \text{ t.q.} \\
&E(\|\tilde{X} - X\|^2) \text{ soit minimale}
\end{aligned}
$$

$$
\boxed{\tilde{X} = \underbrace{\left(A^T A + \sigma_b^2 C^{-1}\right)^{-1} A^T}_{\text{D}} Y}
$$

Conclusions provisoires

- Il n’est pas praticable d’inverser simplement le défaut qui affecte une image pour améliorer la qualité de celle-ci.
- Divers moyens (nous avons montré l’utilisation d’une hypothèse statistique) permettent de mettre en place une énergie dont le minimiseur est une image satisfaisante.

Méthodes de minimisation

- La méthode la plus naturelle pour minimiser une énergie est la descente de gradient:
  - Il s'agit de trouver une direction dans laquelle l'énergie décroit le plus vite.
  - Elle revient à dériver l'énergie suivant par rapport à sa variable (qui peut être en grande dimension).
  - Exemples: $E(f) = \|f - g\|^2 \quad \nabla E = (f - g)$

$$E(f) = \iint \|\nabla f\|^2 \quad \nabla E = \triangle f$$

Descente de gradient

- Le schéma général est le suivant:

$$f^{n+1} = f^n - \rho_n \nabla E(f^n)$$

Le pas de descente ρ_n peut être constant ou variable (voir plus loin pour un exemple)

La variation totale

- Comme nous l'avons vu les principaux termes de régularisation qui ont été utilisés sont des termes quadratiques.
- Ils présentent des avantages essentiellement calculatoirs. (Fourier, descente de gradient maîtrisée).
- Cependant, ils ont tendance à dégrader les bords de l'image en les floutant.

Définition

En continu:

$$TV(f) = \int \|\nabla f(x)\| dx$$

En discret:

$$TV(f) = \sum_{i,j} \sqrt{(f(i+1,j) - f(i,j))^2 + (f(i,j+1) - f(i,j))^2)}$$

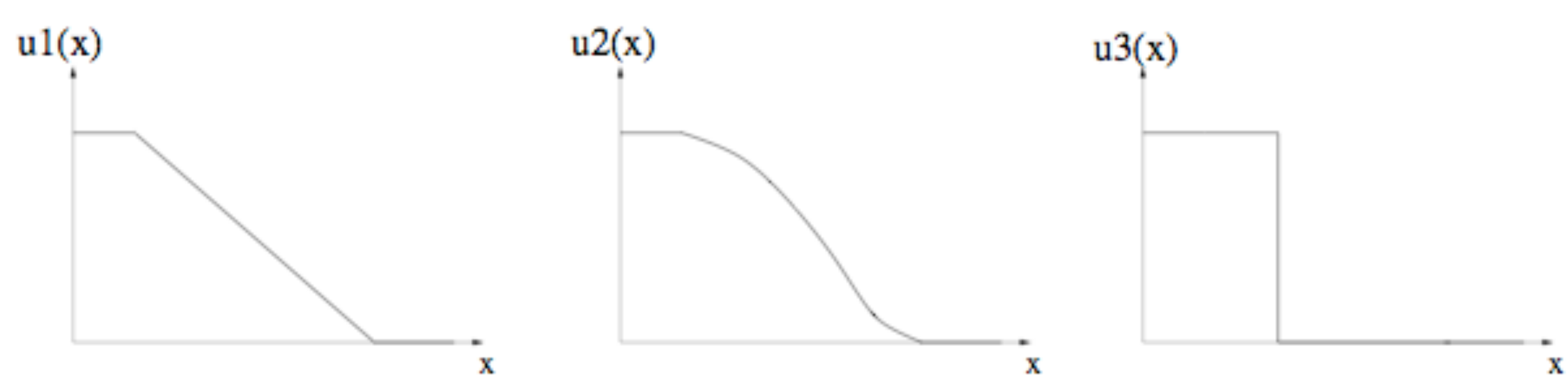
Le gradient de la variation totale

- Remarquer que celui-ci n'est pas bien défini dans les zones où l'image est de gradient nul

$$\nabla TV(f) = -\mathrm{div}\left(\frac{\nabla f}{\|\nabla f\|}\right)$$

Justification de l'utilisation de la variation totale comme fonctionnelle de régularisation

- Contrairement aux fonctionnelles quadratiques l'utilisation de la variation totale n'implique pas des solutions lisses.
- Parmi les trois fonctions ci-dessous, la variation totale est identique, contrairement à l'intégrale du carré de la dérivée.
- D'autres exemples suivront parmi les applications.



Ex: Zoom par minimisation de la VT

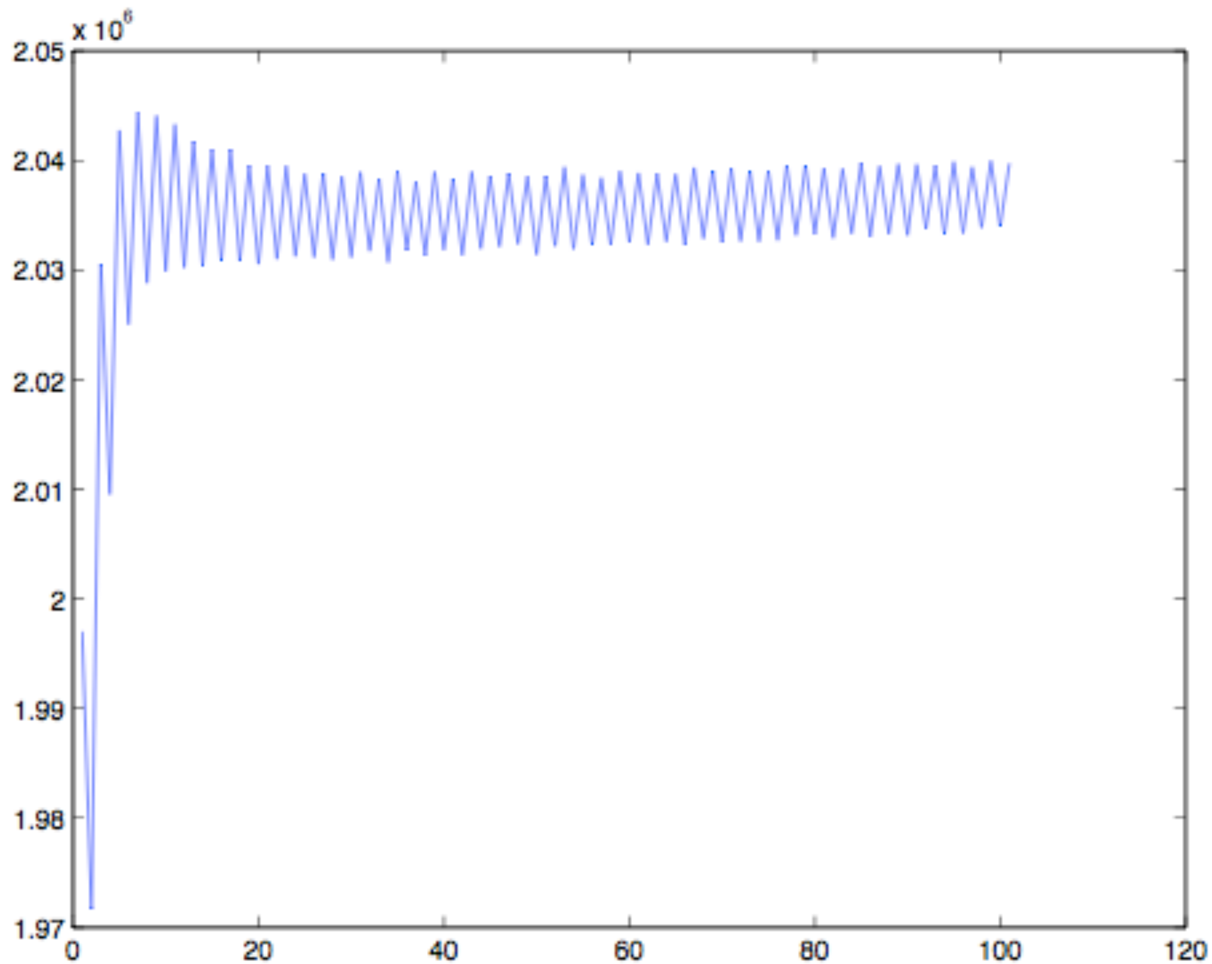
- François Malgouyres propose une méthode de zoom par minimisation de la variation totale.
- Elle se résout par descente de gradient sous contrainte.

$$u_0 = QSu$$

u0 est l'image originale qui doit être obtenue par sous-échantillonnage (Q) d'une version filtrée de u (Su).

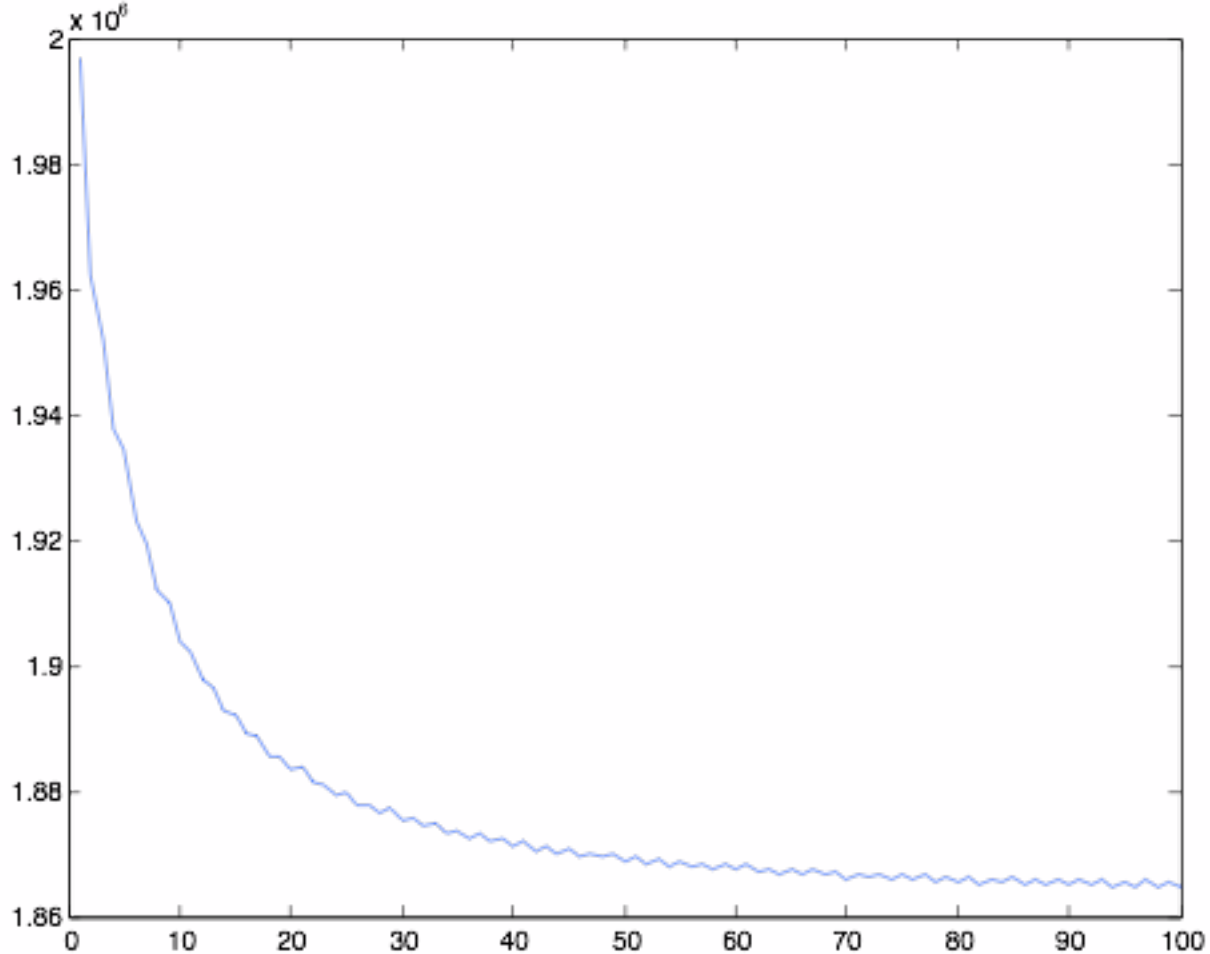
Problèmes numériques de minimisation de la VT par descente de gradient.

Pas=1



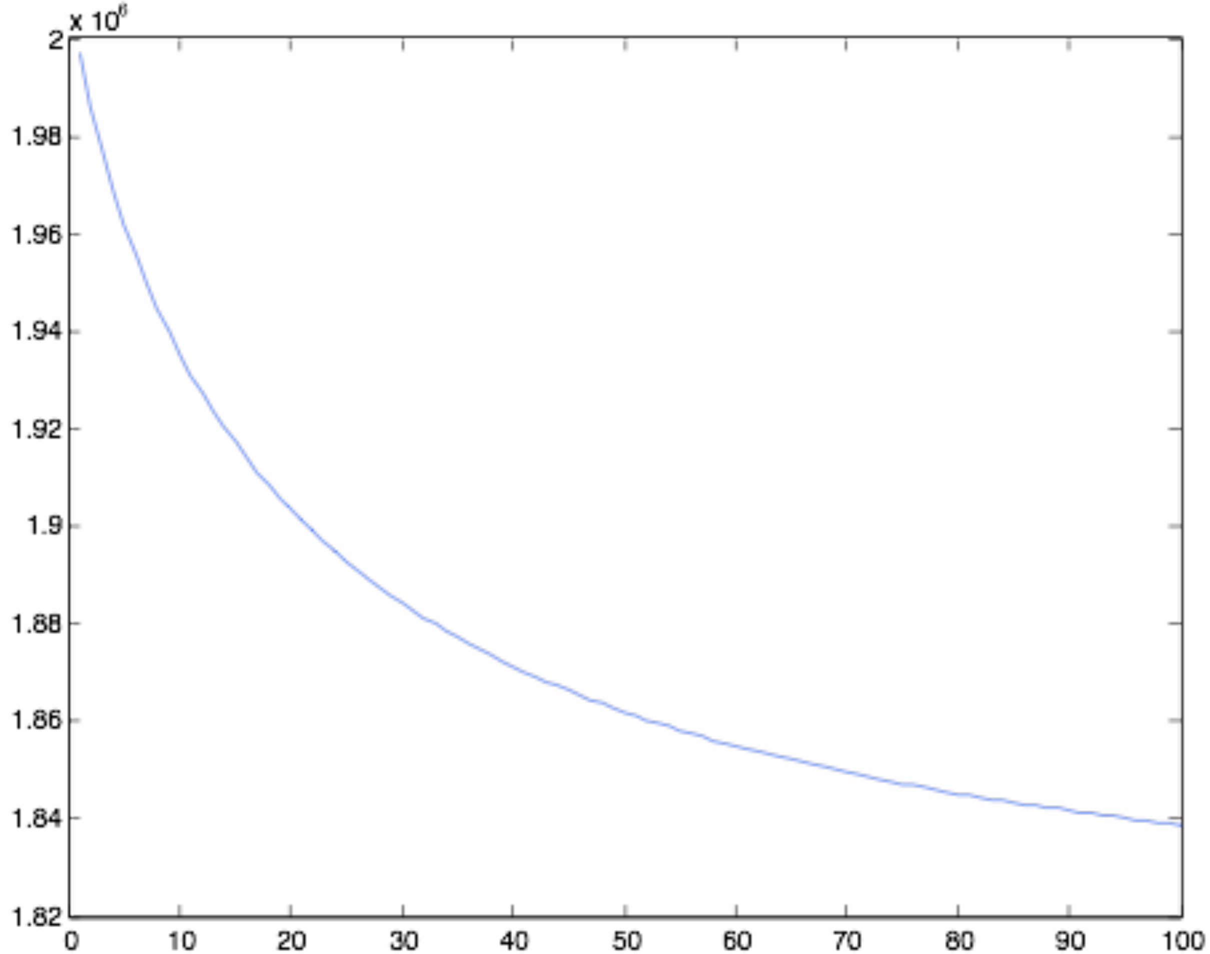
La variation totale augmente!!!

Pas=0.5



Cela semble bien se passer.

Pas=0.1



On pouvait faire mieux!

Débruitage par TV (ROF)

- Rudin, Osher et Fatemi ont proposé un schéma de minimisation de la variation totale sous contrainte de bruit gaussien connu.

Minimiser: $$\int ||\nabla u||$$

Sous contrainte: $$||u - u_0||_2^2 = \sigma^2$$

Cela aboutit à une équation d'Euler-Lagrange:

$$\frac{\partial u}{\partial t} = \mathrm{div}\left(\frac{\nabla u}{||\nabla u||}\right) - \lambda(u - u_0)$$

Les auteurs proposent une détermination de paramètre dynamiquement :

$$\lambda = -\frac{1}{2\sigma^2}\int ||\nabla u|| - \frac{1}{||\nabla u||} \times (\nabla u_0 . \nabla u)$$

Une méthode de minimisation de la variation totale:

Transformée de Fenchel: $$J^*(v) = \sup_u < u|v > -J(u)$$

$$J^* \text{ est convexe.}$$

$$J^{**} = J$$

Si: $J(\alpha u) = |\alpha| J(u)$ alors:

$$J^* = \begin{cases} 0 \text{ sur un fermé } K \\ +\infty \text{ ailleurs} \end{cases}$$

Sous-gradient

- Le sous-gradient remplace le gradient lorsqu'une fonctionnelle n'est pas différentiable.
- C'est un ensemble de vecteurs qui vérifient:

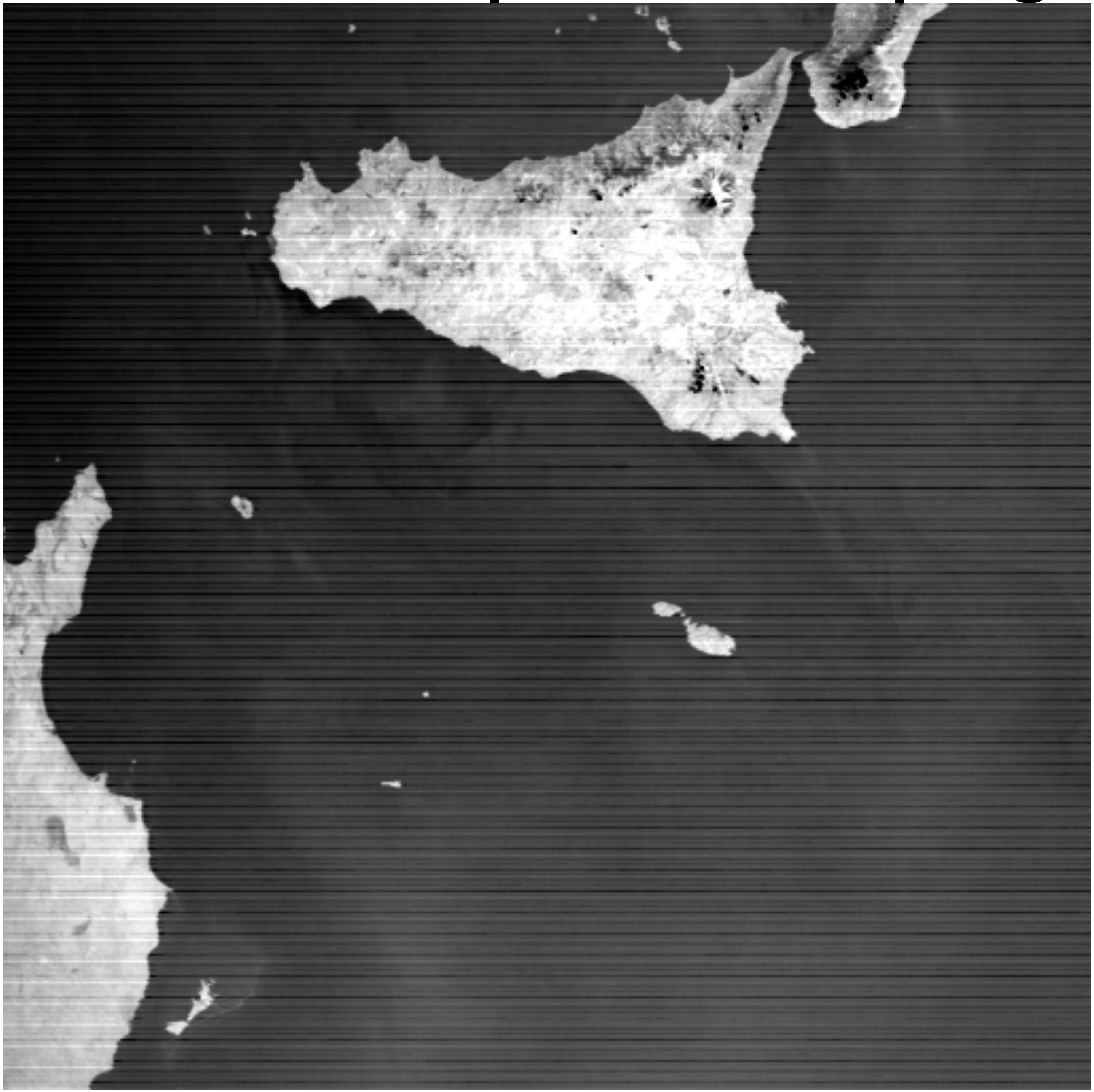
$$\forall u \in \partial J(x), \forall y, J(y) \geqslant J(x) + < u | y - x >$$

Lorsque J est différentiable le sous-gradient est réduit à un singleton: le gradient.

On a: $u \in \partial J(x) \Leftrightarrow x \in \partial J^*(u)$

On montre que minimiser ROF revient a minimiser une fonctionnelle équivalente a base de J*. Or J* est très simple et la minimisation d’une fonctionnelle dans laquelle elle intervient revient a une projection sur le ferme K.

Autre exemple: Destriping



On va modéliser:

$$g = a + b. \text{ en minimisant} TV_x(b) + TV_y(a)$$

Déconvolution par TV

Déconvolution par TV

Minimisation par « split »

$E(d, f) = TV(d) + \lambda \|g - Af\|^2 + \gamma \|d - \nabla f - b\|^2$
d: représente le gradient de f
b: c'est un accumulateur$b_{n+1} = b_n + \nabla f - d$
On minimise alternativement par rapport à d et f.

Voir: http://www.ipol.im/pub/art/2012/g-tvdc/