

Cours 3: Restauration d'images

Julien Lefèvre

Julien.lefevre@univ-amu.fr



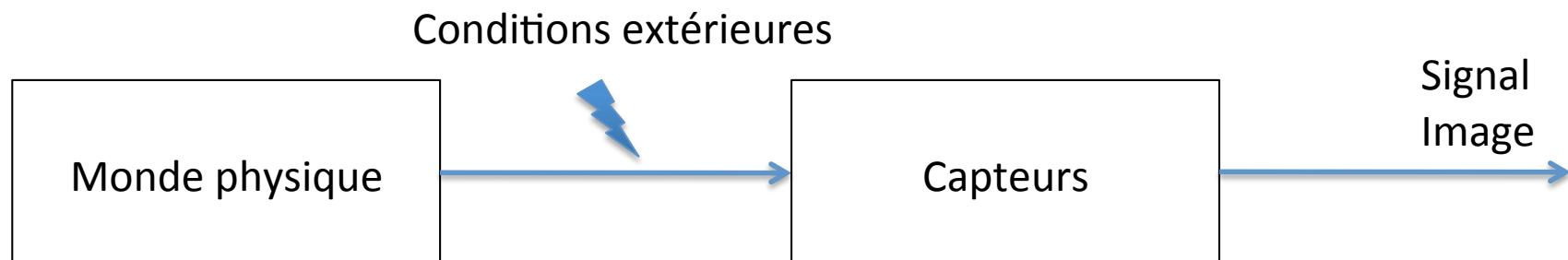
Laboratoire des sciences de l'Information et des Systèmes
UMR CNRS 7296

Plan du cours

- Introduction
- Premières approches
- Un cadre plus général
- Inpainting
- Pour aller plus loin...

Introduction

- Dans son processus d'acquisition, une image peut subir des dégradations
- Ces dégradations peuvent être dues :
 - aux capteurs
 - aux conditions de prise de vues

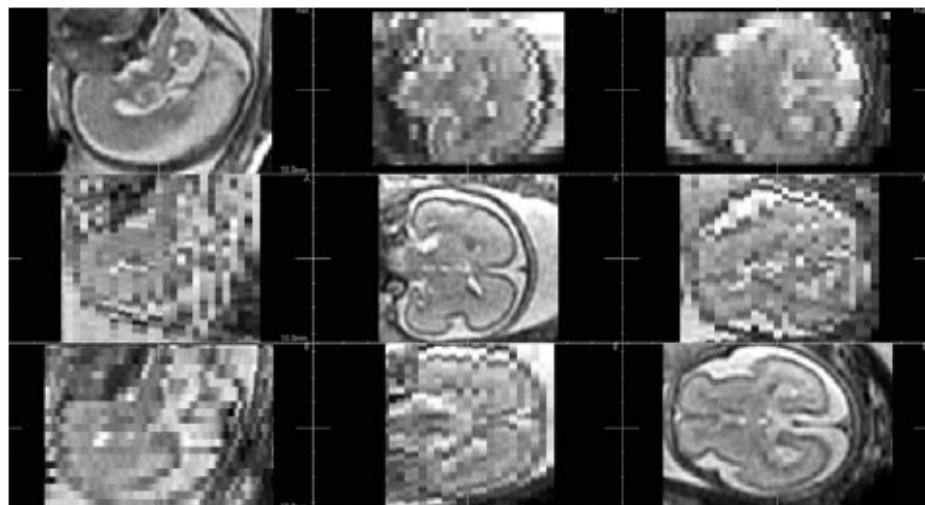


Introduction

- Défauts du capteur
 - Le flou du à l'optique et à l'intégration de la lumière sur chaque pixel
 - bruit de quantification
- Défauts extérieurs
 - Flou de bougé
 - Perturbations atmosphériques
 - Poussières sur l'objectif

Introduction

Artefacts de mouvements en IRM foetale



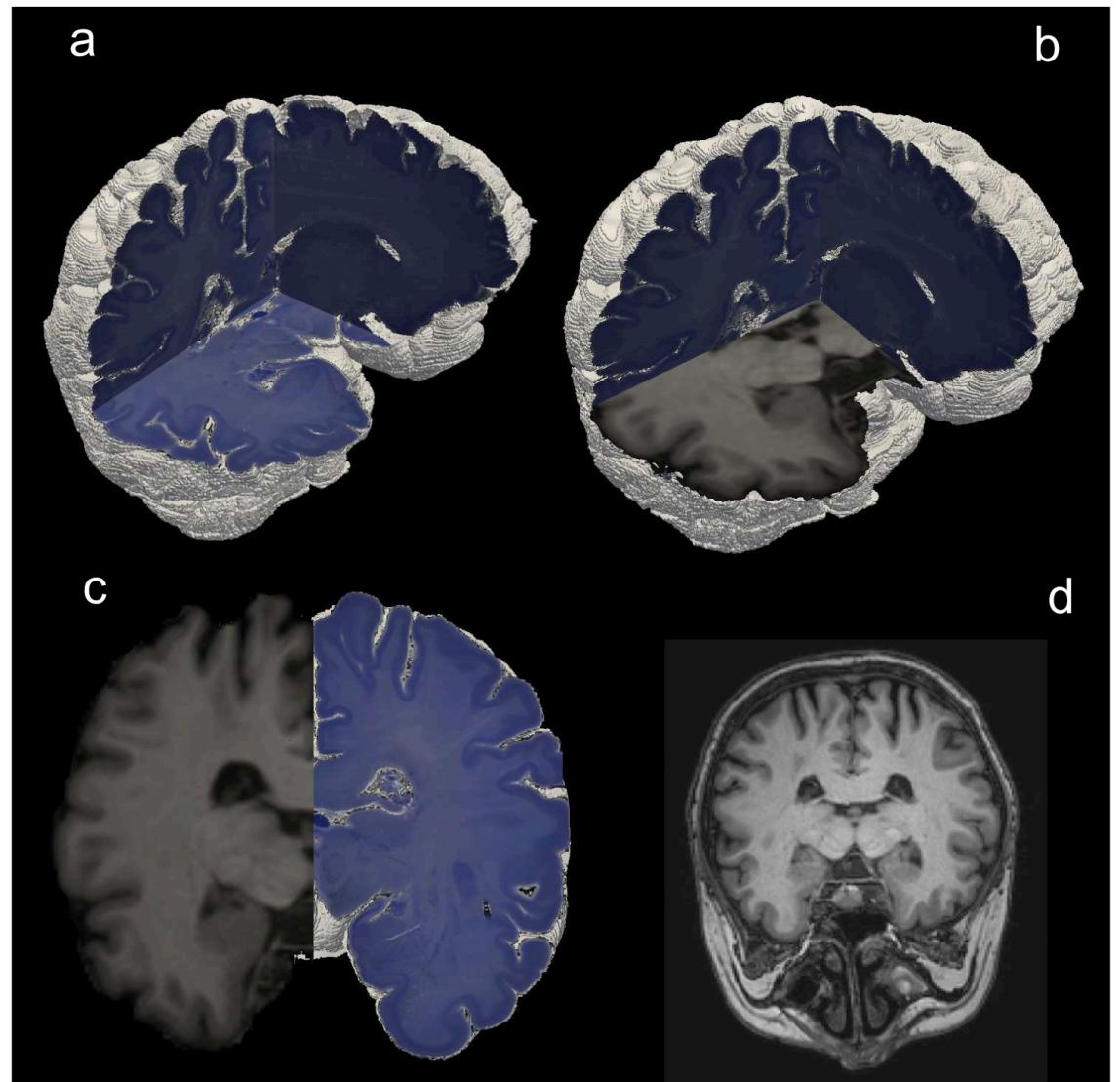
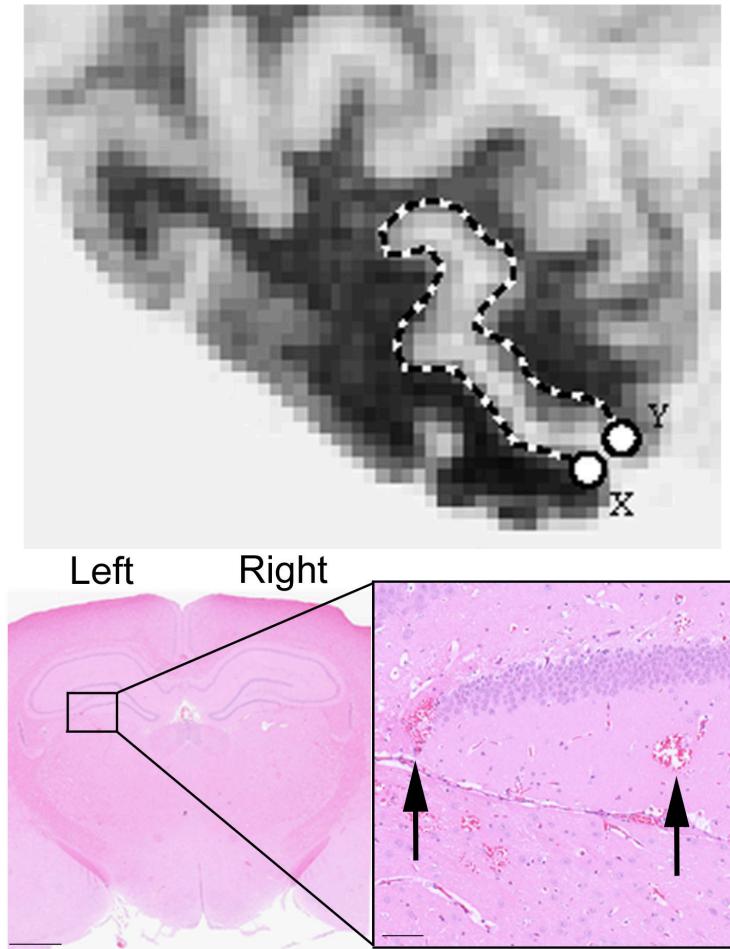
Super-résolution



Rousseau et al, 2006

Introduction

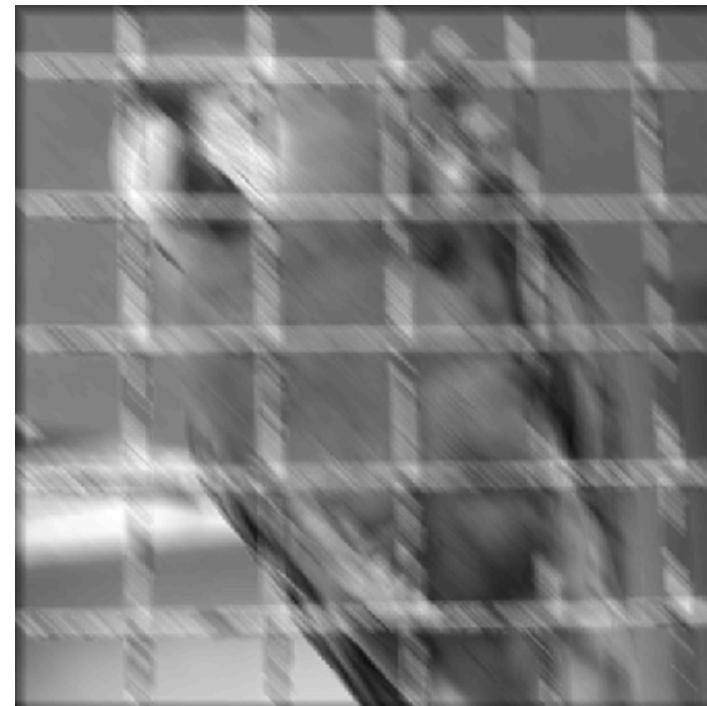
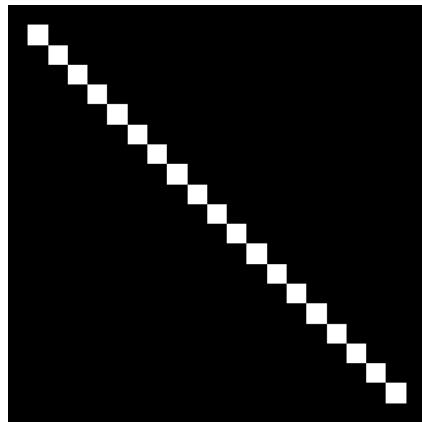
Echantillonage



Introduction

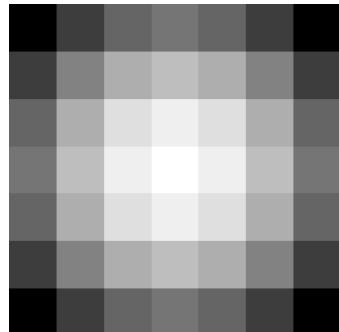
- Le flou se caractérise par une convolution (multiplication dans l'espace de Fourier)
- Flou de bougé

Exemple :



Introduction

- Le flou atmosphérique est modélisé par la convolution avec un filtre gaussien :



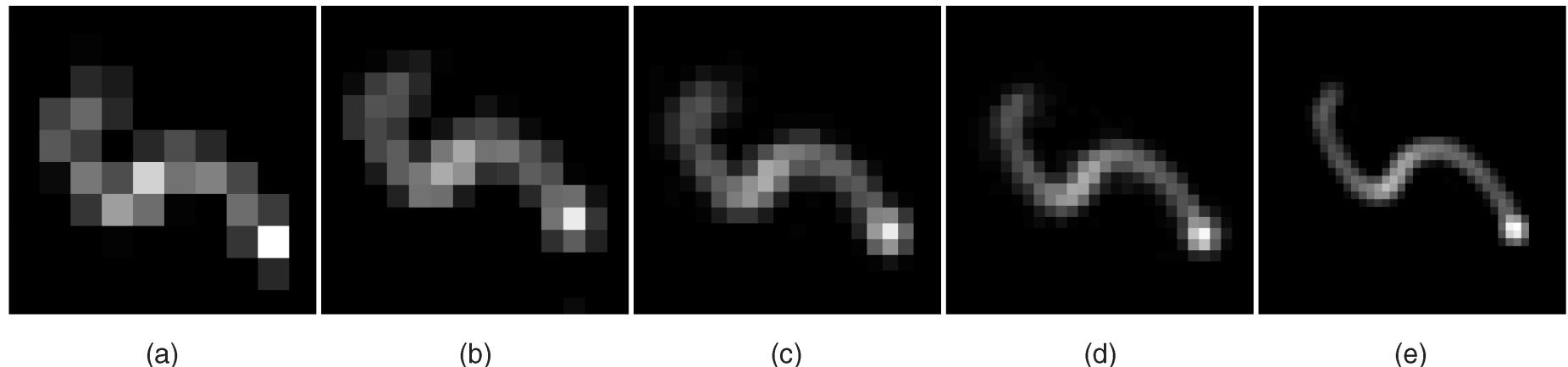
$\sigma=5$

$$\frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$



Introduction

- Noyaux de flou plus compliqués



Introduction

- Le bruit peut-être additif...

$$O(x,y) = I(x,y) + b(x,y)$$

Avec $b(x,y)$ variable aléatoire de moyenne égale à 0

- ... mais aussi multiplicatif

$$O(x,y) = I(x,y)b(x,y)$$

Avec $b(x,y)$ variable aléatoire de moyen

Exemple : bruit de speckle dans les échographies



Introduction

- Le bruit peut être :

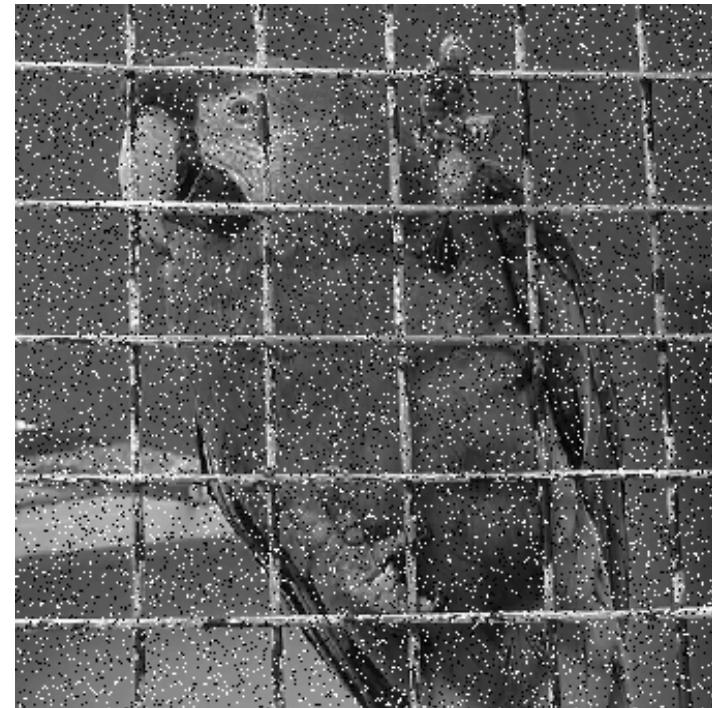
gaussien

$\mu=0, \sigma=5$



Impulsionnel (« poivre et sel »)

10%



Introduction

- Simulation du bruit gaussien (cf TP)
 - Une méthode fausse
 - Méthode des 12 variables aléatoires uniformes
$$S_n = U_1 + \dots + U_{12} - 6$$
 - Méthode Box-Muller

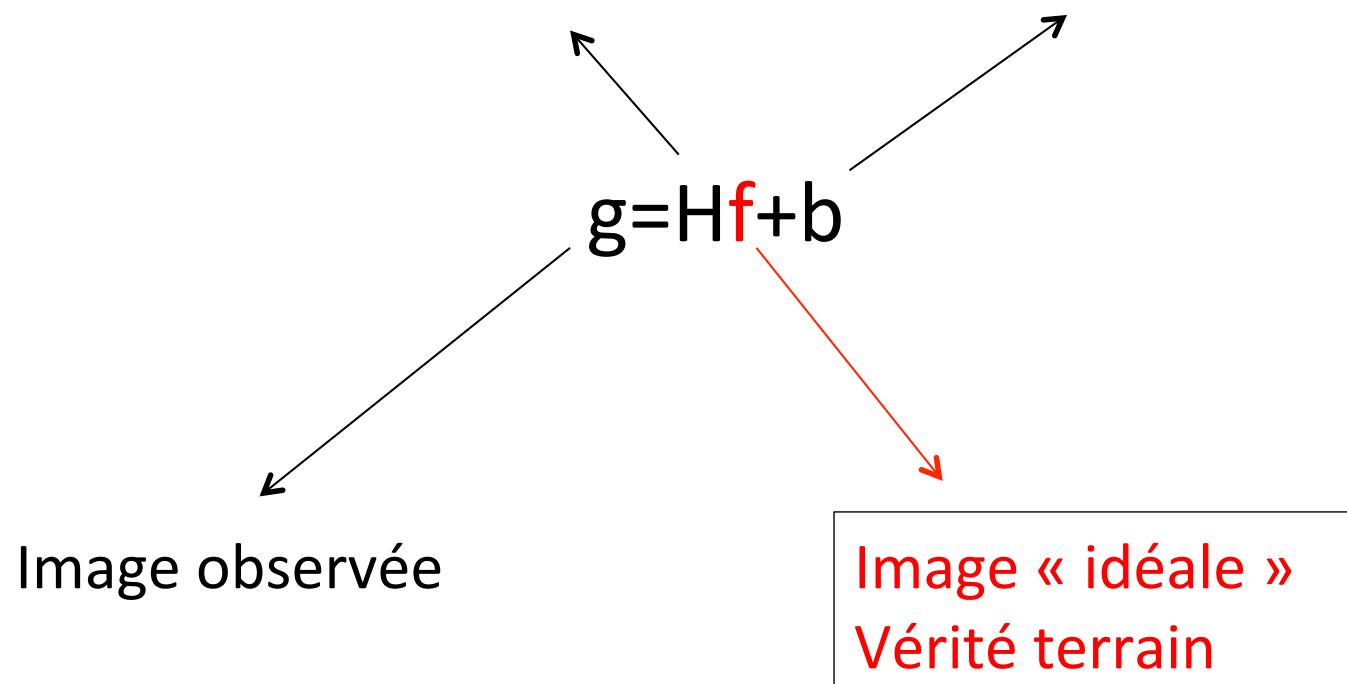
U et V deux lois uniformes sur $[0, 1]$ alors $\sqrt{-2 \log U} \cos 2\pi V$ et $\sqrt{-2 \log U} \sin 2\pi V$ suivent une loi normale centrée réduite.

Introduction

- Un modèle de formation d'une image

Opérateur traduisant la formation
de l'image à travers un capteur

Partie aléatoire (bruit)



Introduction

- Le modèle s'écrit en général à l'aide d'une convolution (filtre) : $g=H*f+b$
- Le problème consiste donc à retrouver f connaissant g
- Il s'agit d'un **problème inverse mal posé** : il y a en général une **infinité de solutions**

Premières approches

- Première approche

On peut réécrire la convolution comme un produit matriciel :

$$g = Hf + b$$

f, g et b sont des vecteurs de taille $(m \times n) \times 1$ (taille de l'image)

On cherche ensuite à trouver une approximation de i par la **méthode des moindres carrés.**

On veut minimiser l'erreur entre g et Hf (cf tableau)

Premières approches

- Première approche

La minimisation de $\|g-Hf\|^2$ par rapport à f s'obtient par différentiation.

On regarde les variations de l'expression au voisinage de f :

$$\|g-H(f+df)\|^2 = \|g-Hf\|^2 + 2\langle g-Hf, Hdf \rangle + \|Hdf\|^2$$

Le dernier terme est négligeable par rapport à $\|df\|$

Au voisinage d'un minimum (local) on a donc

$$\langle g-Hf, Hdf \rangle = 0 \text{ pour tout } df$$

Ce qui conduit à ${}^tH(g-Hf)=0$ et $f={}^t(HH)^{-1}{}^tHg$

Premières approches

- Première approche

En présence uniquement de flou on peut reconstruire l'image par une inversion matricielle



Premières approches

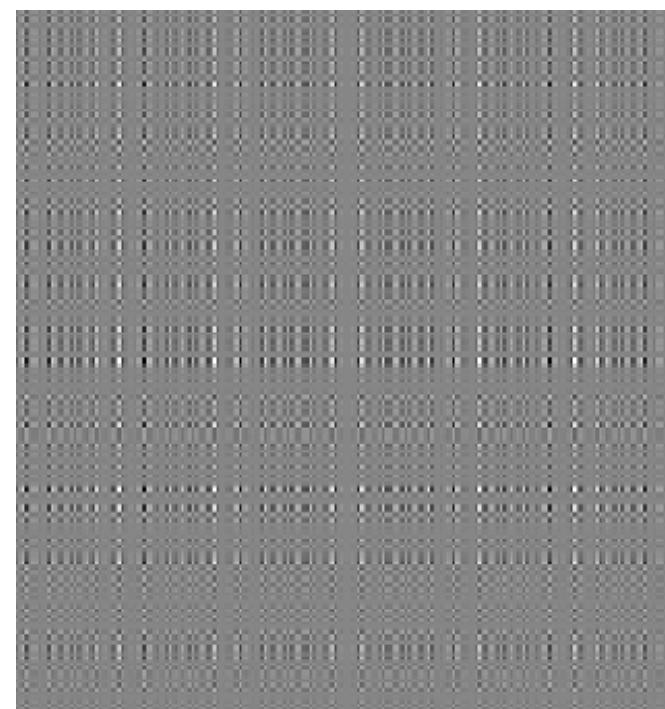
- Première approche

Cependant il faut connaître la matrice H. Une erreur sur son estimation peut totalement fausser la reconstruction

$\sigma=2$ pixels



$\sigma=2.1$ pixels

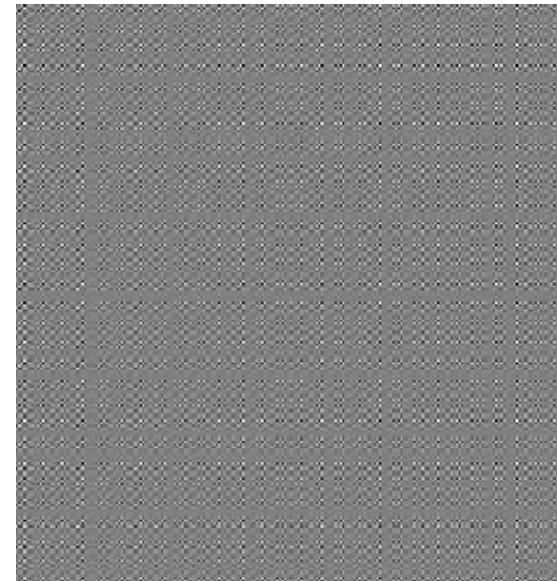


Premières approches

- Première approche

En pratique l'inversion matricielle est peu concluante en présence de bruit -> instabilités numériques

Bruit Gaussien $\mu=0$, $\sigma=20$



Un cadre plus général

- Retour sur le problème $g=Hf+b$
 - Problème mal posé, infinité de solutions
 - On peut restreindre l'espace des solutions à un ensemble d'images possédant certaines propriétés de régularité
 - Concrètement cela se traduit par la minimisation d'un critère global ou énergie :

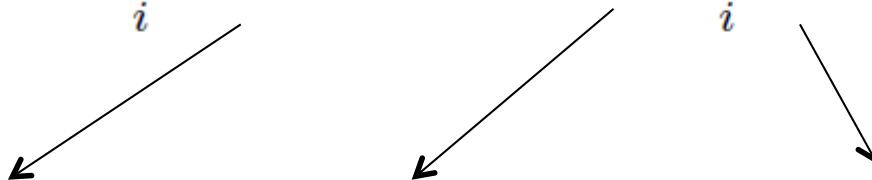
$$\text{Energie globale}(f) = \text{Attache aux données}(f, g) + \lambda \text{ Régularité}(f)$$

Un cadre plus général

- Régularisation de Tikhonov
 - Formulation discrète

$$E(f) = \|g - Hf\|^2 + \lambda \|Lf\|^2$$

$$E(f) = \sum_i (g_i - (Hf)_i)^2 + \lambda \sum_i (Lf)_i^2$$



Terme d'attache aux
données

Paramètre de
régularisation

Terme de régularité.
Typiquement on prend pour
L l'opérateur gradient

Un cadre plus général

- Minimisation de fonctions

Descente de gradient pour minimiser une fonction de n variables

On part de l'observation que la fonction décroît le plus dans la direction opposée à celle de son gradient.

- On part d'une initialisation x_0
- On construit successivement x_n de la façon suivante

$$x_{n+1} = x_n - \epsilon \nabla \Phi(x_n)$$

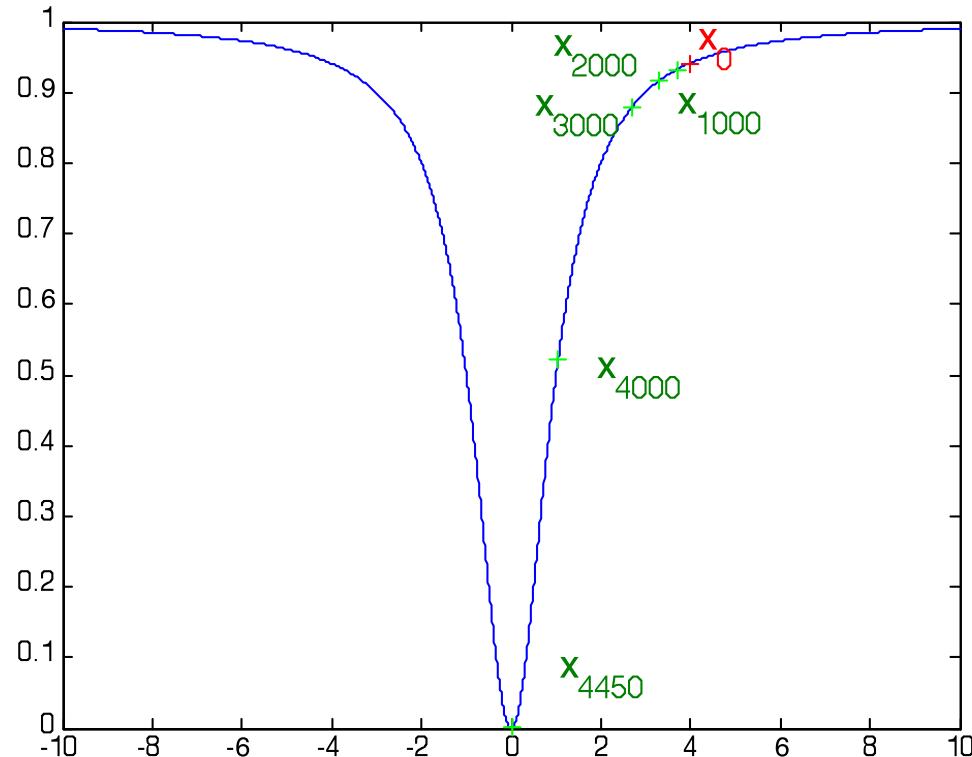
Avec ϵ petit

Un cadre plus général

- Minimisation de fonctions

Illustration de la descente de gradient

$$\Phi(x) = 1 - \frac{1}{1 + x^2}$$



$$\epsilon = 0.01$$

Critère d'arrêt

$$|x_{n+1} - x_n| < 1e-5$$

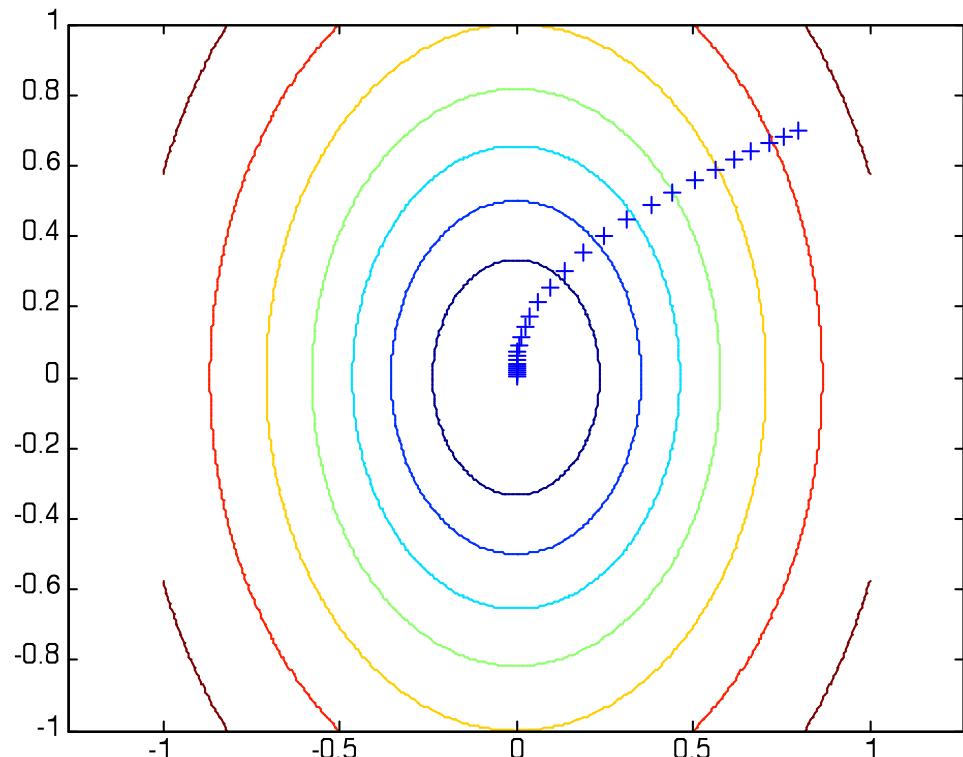
$$x_{fin} = 4.96e-4$$

Un cadre plus général

- Minimisation de fonctions

En Dimension 2

$$\Phi(x, y) = 1 - \frac{1}{1 + 2x^2 + y^2}$$



$$\epsilon = 0.1$$

54 itérations

Un cadre plus général

- Minimisation de l'énergie

L'énergie à minimiser est une fonction de $m \times n$ variables (intensités des pixels de l'image)

L'énergie est minimale en f quand elle est insensible aux faibles variations $\delta f = \varepsilon u$, avec u une image quelconque.



Il faut calculer le gradient de l'énergie

Un cadre plus général

- **Minimisation de l'énergie**

L'énergie à minimiser est une fonction de $m \times n$ variables (intensités des pixels de l'image)

L'énergie est minimale en f quand elle est insensible aux faibles variations $\delta f = \epsilon u$, avec u une fonction quelconque.

Si f est le minimum de E on a :

$$\lim_{\epsilon \rightarrow 0} \frac{E(f + \epsilon u) - E(f)}{\epsilon} = 0 \quad \text{pour tout } u$$

La dérivée de E au point f dans la direction u s'annule

Un cadre plus général

- Minimisation de l'énergie

Réécrivons l'énergie $E(f + \delta f)$ en fonction de $E(f)$.

$$E(f + \delta f) = \sum_i (g(i) - H(f + \delta f)(i))^2 + \lambda \sum_i \|L(f + \delta f)(i)\|^2$$

$$E(f + \delta f) = \sum_i (g(i) - Hf(i) - H\delta f(i))^2 + \lambda \sum_i \|Lf(i) + L\delta f(i)\|^2$$

$$\begin{aligned} E(f + \delta f) &= E(f) - 2 \sum_i H\delta f(i)(g(i) - Hf(i)) + 2\lambda \sum_i L\delta f(i) \cdot Lf(i) \\ &\quad + \sum_i (H\delta f(i))^2 + \sum_i \|L\delta f(i)\|^2 \end{aligned}$$

$$(a+b)^2 = a^2 + 2ab + b^2$$

De l'ordre de δf

De l'ordre de δf^2

Un cadre plus général

- Minimisation de l'énergie

Dans la partie qui nous intéresse

$$2 \sum_i H \delta f(i) (g(i) - Hf(i)) + 2\lambda \sum_i L \delta f(i) \cdot Lf(i)$$

On va pouvoir écrire chaque somme comme des produits scalaires entre (gros) vecteurs:

$$\langle H \delta f | g - Hf \rangle_1 = \sum_i H \delta f(i) (g(i) - Hf(i)) \quad \text{Sur les images}$$

$$\langle L \delta f | Lf \rangle_2 = \sum_i L \delta f(i) \cdot Lf(i) \quad \text{Sur les champs de vecteurs}$$

Un cadre plus général

- Minimisation de l'énergie

Un peu d'algèbre...

On utilise l'opérateur adjoint de H noté H^* qui doit vérifier

$$\langle Hu | v \rangle = \langle u | H^* v \rangle$$

Exemples :

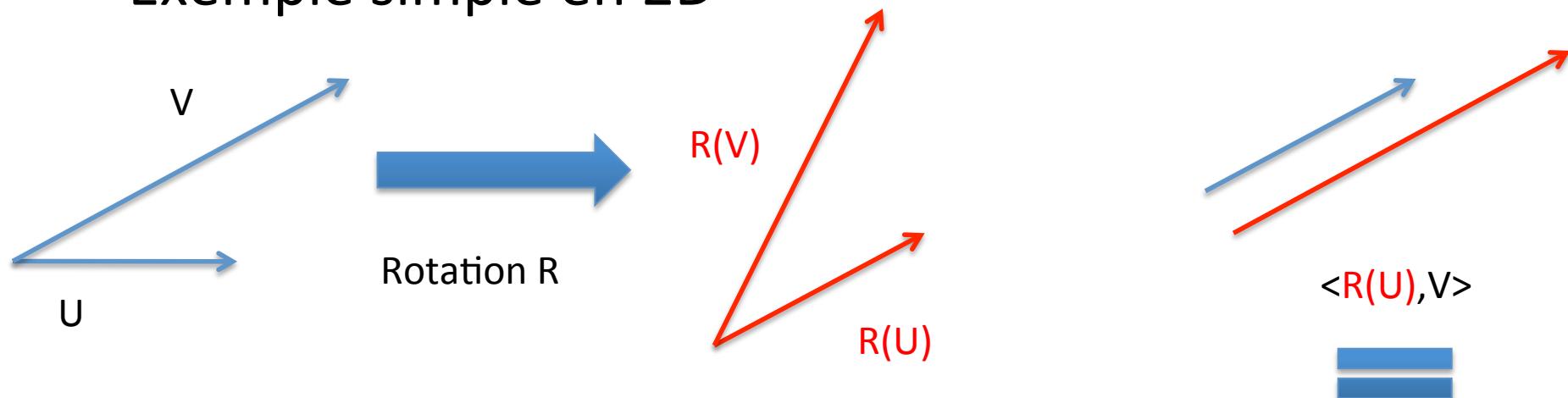
- l'identité est son propre adjoint,
- l'adjoint d'une rotation est la rotation d'angle opposée.

Plus généralement l'adjoint correspond à la **transposée de la matrice** associée à l'opérateur linéaire.

Un cadre plus général

- Aparté important: notions de dualité

Exemple simple en 2D



Quelle est la transformation R^* telle que

$$\langle R(U), V \rangle = \langle U, R^*(V) \rangle ?$$

$$\langle R^*(V), U \rangle$$



Un cadre plus général

- Minimisation de l'énergie

Alors on peut réécrire

$$E(f + \delta f) = E(f) - 2 < H^*(g - Hf), \delta f > + 2\lambda < L^*Lf, \delta f > \\ + O(||\delta f||^2)$$

Ainsi la condition sur la dérivée se traduit par

$$-H^*(g - Hf) + \lambda L^*Lf = 0$$

Un cadre plus général

- **Minimisation de l'énergie**

Pour minimiser l'énergie on utilise donc une **descente de gradient** :

- On part d'une image initiale f^0
- On construit une suite d'images f^n qui vérifie la relation

$$\frac{f^{n+1} - f^n}{\Delta t} = \underbrace{H^*(g - Hf^n) - \lambda L^* L f^n}_{\text{- Gradient de l'énergie}}$$

- Gradient de l'énergie

Un cadre plus général

- Minimisation de l'énergie

Cas du débruitage simple :

On cherche à minimiser

$$E(f) = \sum_i (g(i) - f(i))^2 + \lambda \sum_i \|\nabla f(i)\|^2$$

H est l'identité.

On cherche l'expression de L*Lf ou en d'autres termes, quel est l'adjoint du gradient ?

∇f : Image \rightarrow Champ vecteurs

$f(x,y) \rightarrow \nabla f(x,y) \in \mathbb{R}^2$

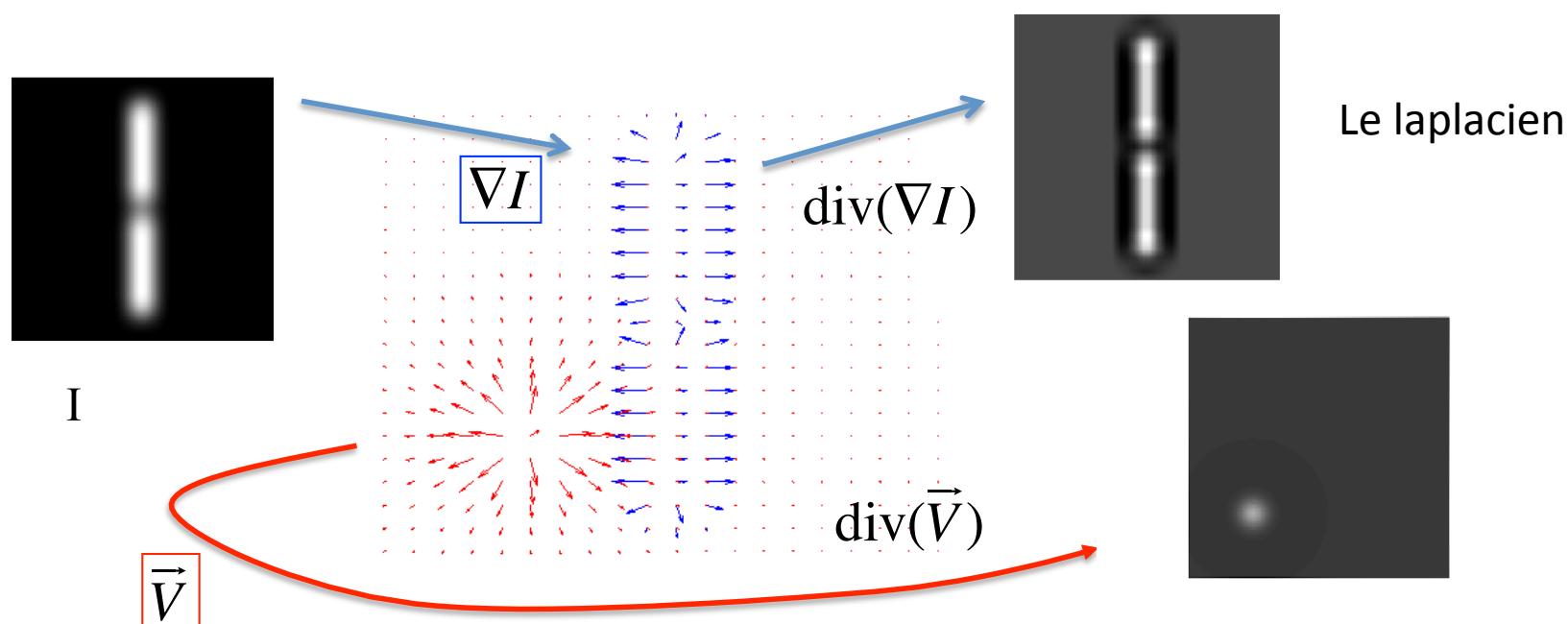
$\nabla^* f$: Champ vecteurs \rightarrow Image

$(V_x, V_y) \rightarrow \nabla^* f(V_x, V_y) \in \mathbb{R}$

Un cadre plus général

- Adjoint du gradient

C'est la divergence (cf cours équation de diffusion).



$$\sum I \text{div}(\vec{V}) = -\sum \nabla I \cdot \vec{V}$$

Un cadre plus général

- Minimisation de l'énergie

Il s'agit encore des formules de Green (intégration par parties)

$$\begin{aligned}\langle \nabla u | \nabla v \rangle_2 &= \sum_i \nabla u(i) \cdot \nabla v(i) \\ &= \langle u | \nabla^* \nabla v \rangle_2 = \langle u | -\operatorname{div} \nabla v \rangle_2 = \langle u | -\Delta v \rangle_2\end{aligned}$$

D'où l'EDP de débruitage

$$\frac{\partial f}{\partial t} = g - f + \lambda \operatorname{div}(\nabla f)$$

Débruitage simple

- Débruitage

$\sigma=20$



$\Delta t=0.1$, 1000 itérations,
 $\lambda=0.5$



$\Delta t=0.01$, 1000 itérations,
 $\lambda=5$



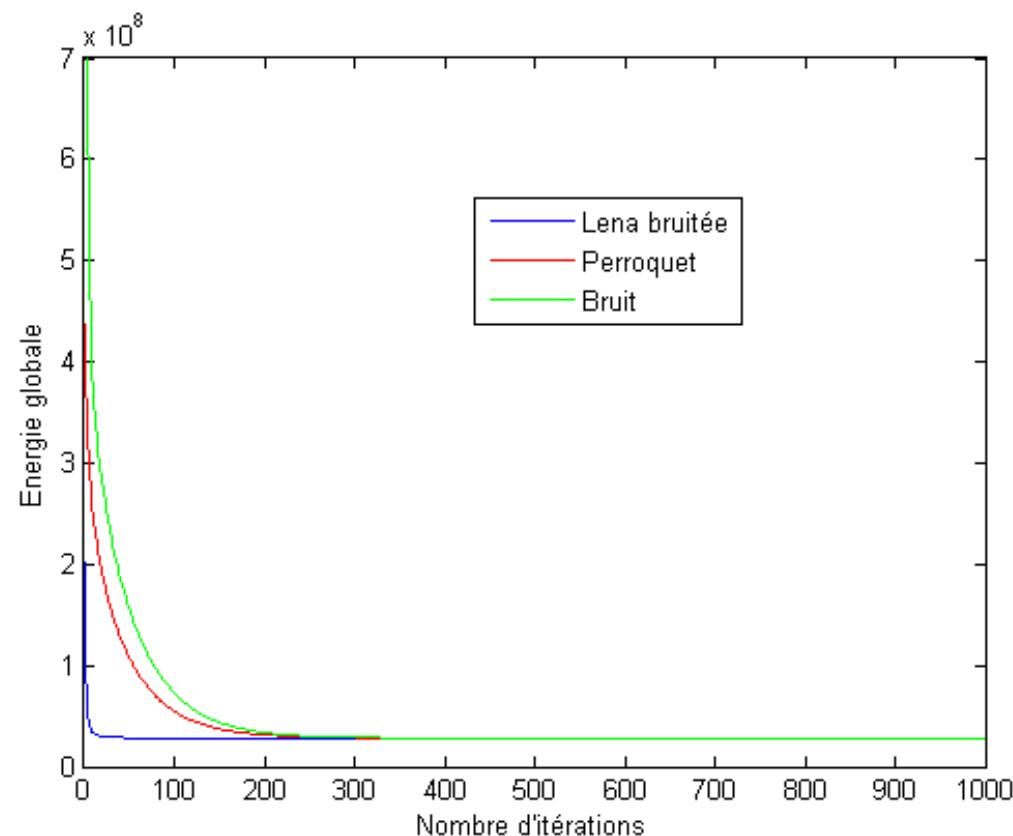
D



ce de la

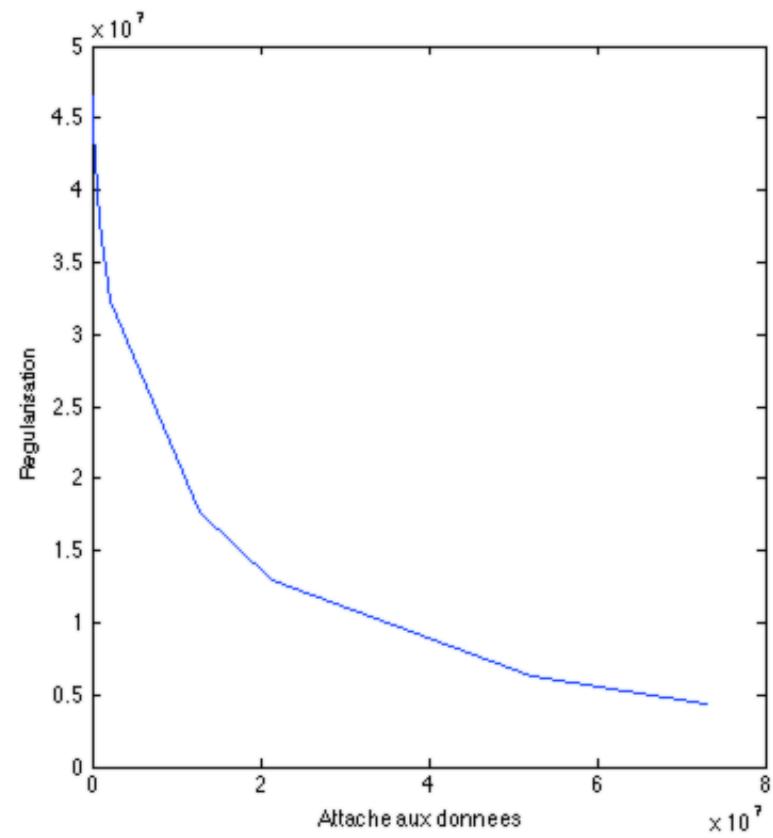
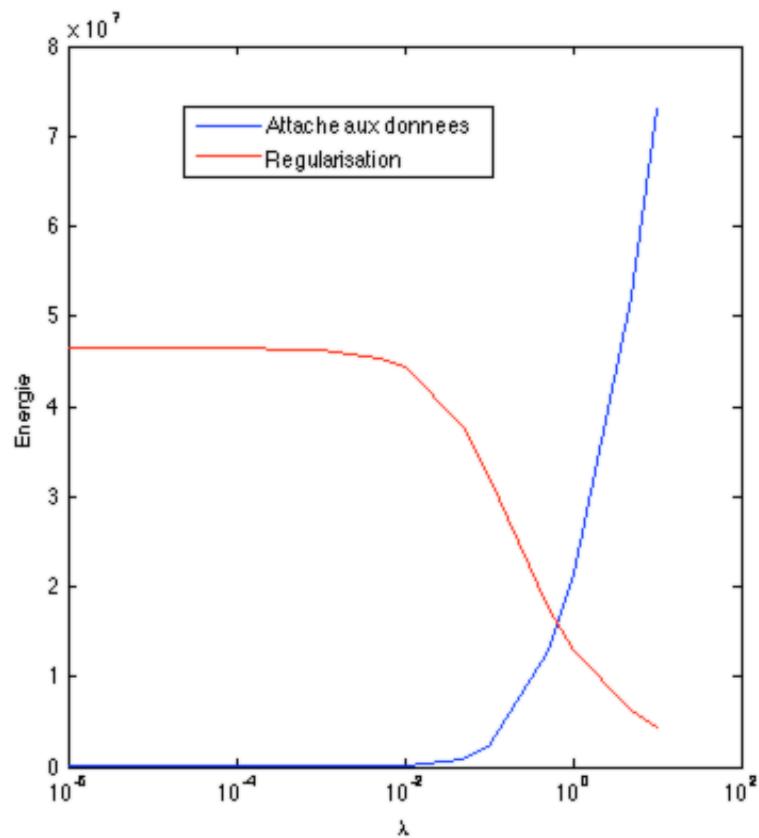
Débruitage simple

La condition initiale change la vitesse de convergence



Débruitage simple

Influence du paramètre de régularisation



Débruitage simple

- Rapport signal sur bruit

$$SNR = 10 \log_{10} \left(\frac{\sum_{i,j} I(i,j)^2}{\sum_{i,j} (I(i,j) - I_2(i,j))^2} \right)$$

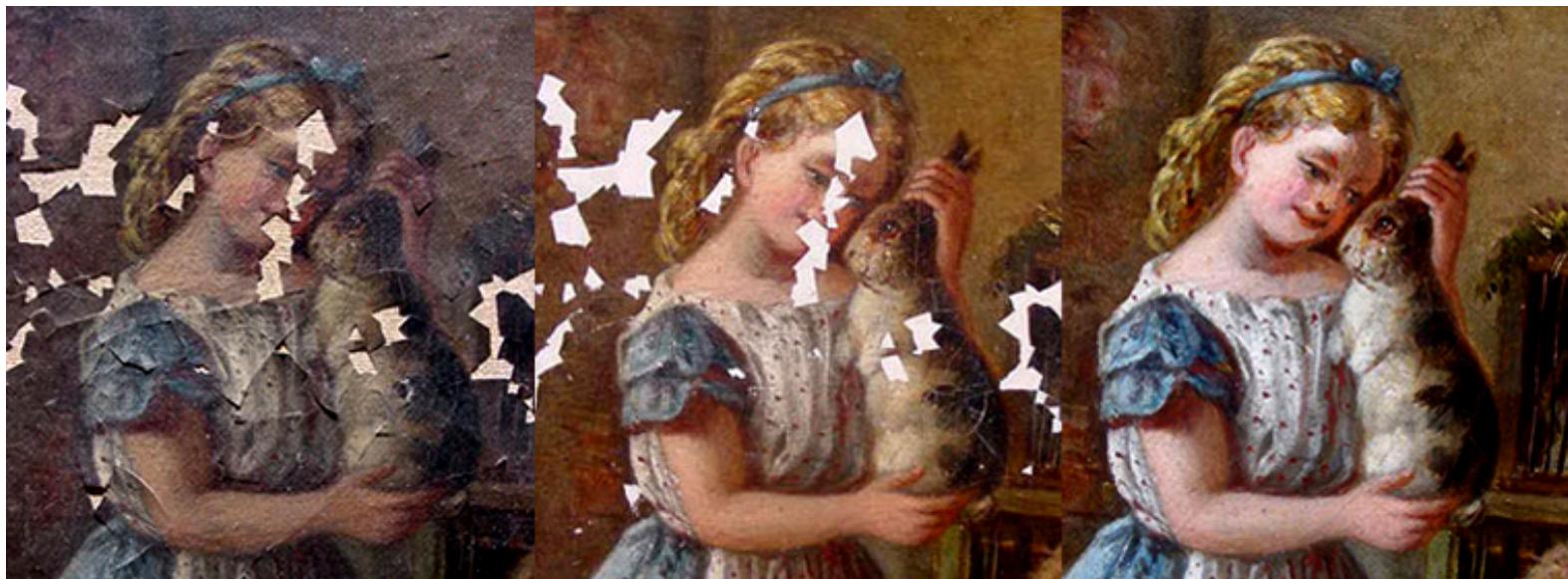
$$PSNR = 10 \log_{10} \left(\frac{MN \max_{i,j} I(i,j)^2}{\sum_{i,j} (I(i,j) - I_2(i,j))^2} \right)$$

Inpainting

- Position du problème

Il s'agit d'un problème aussi ancien que la création artistique elle-même.

Consiste en la retouche d'une image dégradée en certains endroits.



Inpainting

- Formalisation

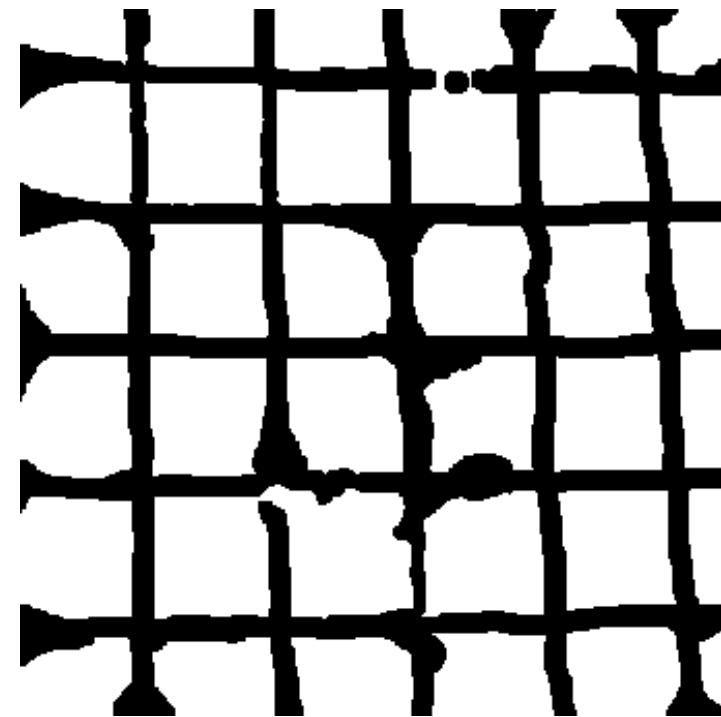
On définit un opérateur de masquage

Si on a perdu de l'information dans un domaine D on définit H_f par $H_f(x) = 0$ si $x \in D$, $H_f(x) = f(x)$ si $x \notin D$



Inpainting

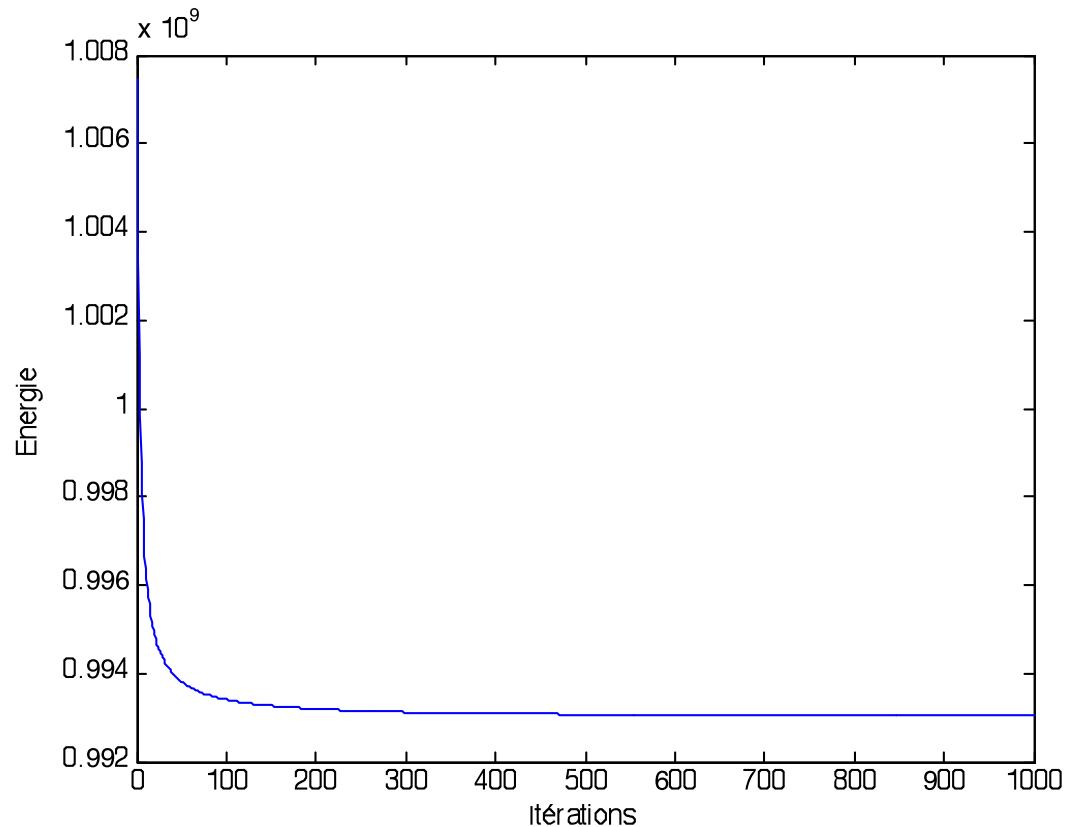
- Quelques résultats



Inpainting

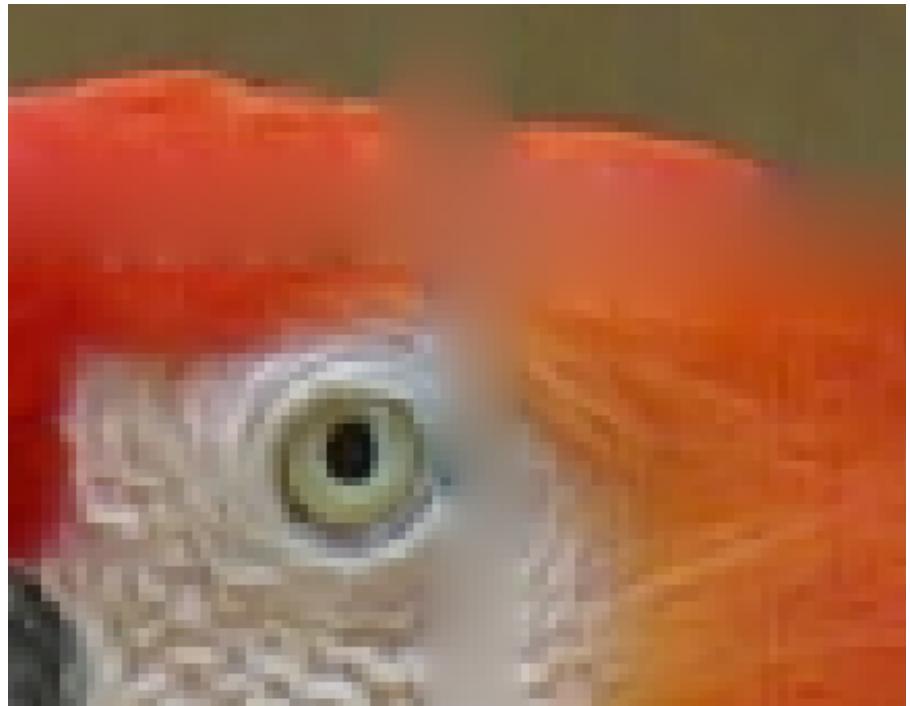
- Quelques résultats

$\Delta t=0.5$, 1000 itérations, $\lambda=0.2$



Inpainting

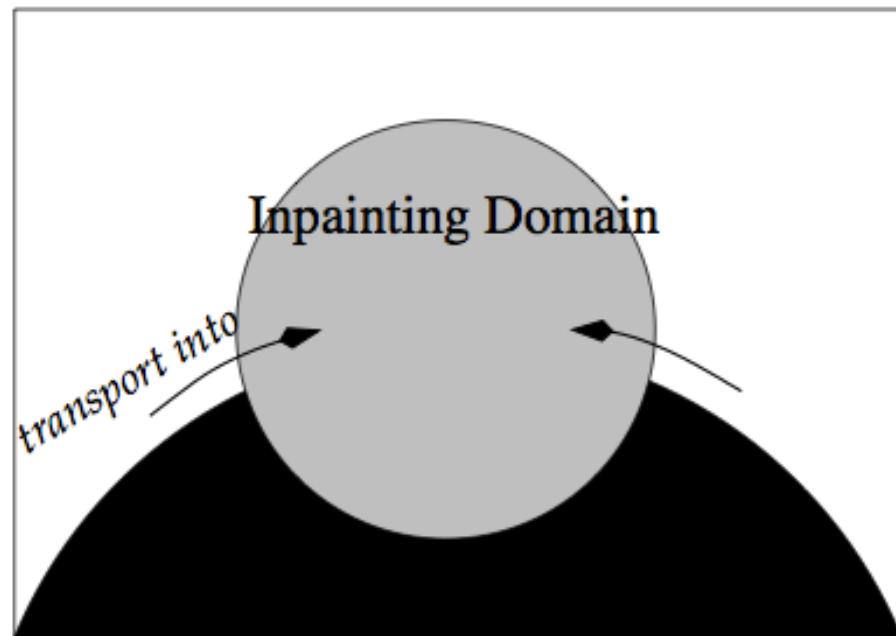
- Quelques résultats



Inpainting

- D'autres méthodes (plus efficaces)

Bertalmio et al, 2000 : principe de continuité des isophotes



Définition d'une mesure de régularité L (laplacien).

A l'équilibre les variations de cette mesure le long des isophotes doivent être nulles:

$$\nabla I^\perp \cdot \nabla(L(I)) = 0$$

Inpainting

- D'autres méthodes (plus efficaces)
Bertalmio et al, 2000 : principe de continuité des isophotes



Inpainting

- D'autres méthodes (plus efficaces)

Tschumperlé et al, 2005 : utilisation du tenseur de structure



Inpainting

- D'autres méthodes (plus efficaces)
 - Tschumperlé et al, 2005 : utilisation du tenseur de structure



Inpainting

- Méthodes de D. Tschumperlé



100 itérations, dt=10



1000 itérations, dt=10

Inpainting

- Méthodes de D. Tschumperlé



10 itérations, dt=10



100 itérations, dt=10

Pour aller plus loin...

- Choix du critère de régularisation

La norme utilisée pour régulariser a tendance à dégrader les contours car liée à une diffusion isotrope.

On peut utiliser plus généralement une fonction qui dépende de la norme du gradient de l'image

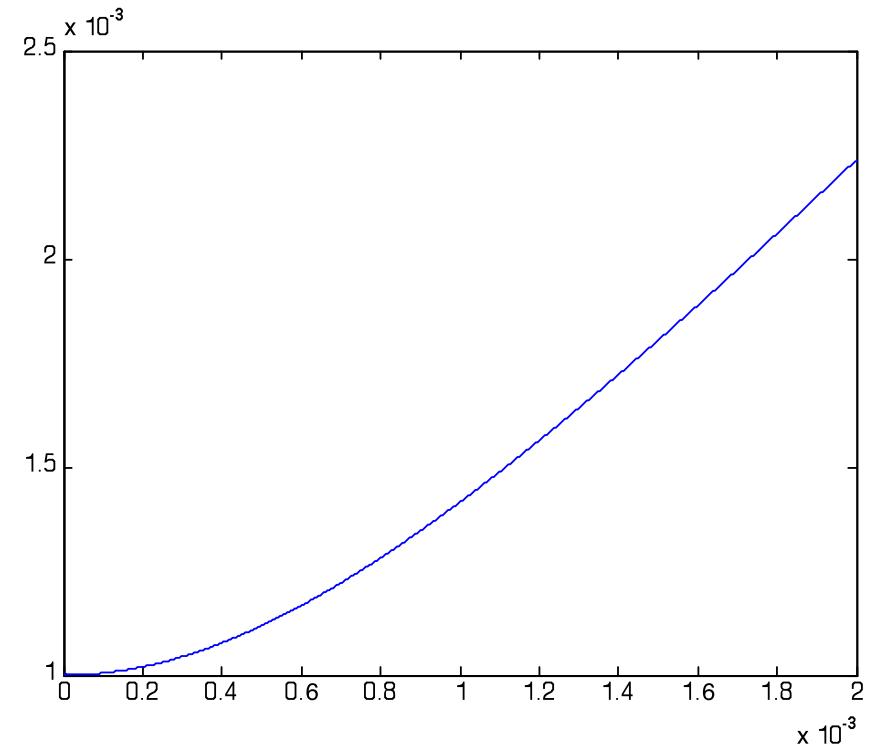
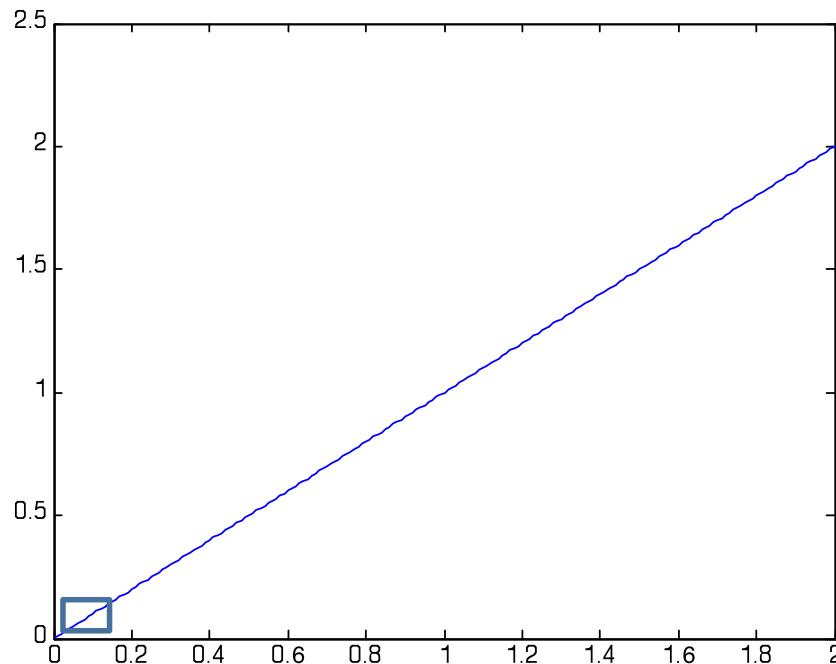
$\psi(x)=x$, régularisation par la variation totale

$\psi(x)=x^2$, diffusion isotrope

Pour aller plus loin

- Régularisation par la variation totale
Approximation de la variation totale par

$$\psi(x) = \sqrt{\epsilon + x^2}$$



Pour aller plus loin

- Régularisation par la variation totale

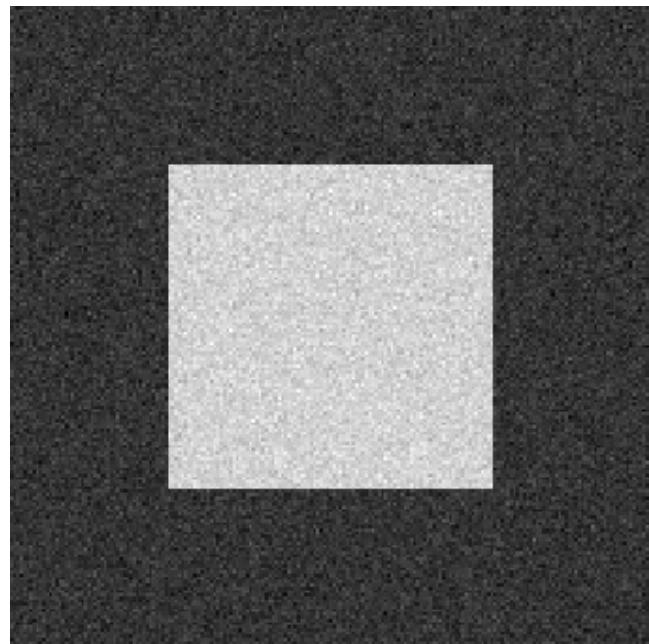
Algorithme de Chambolle (2004)

- 2 opérateurs duaux: gradient, divergence
- Calcul itératif comme pour une descente de gradient
(formule un peu difficile à établir)

Pour aller plus loin

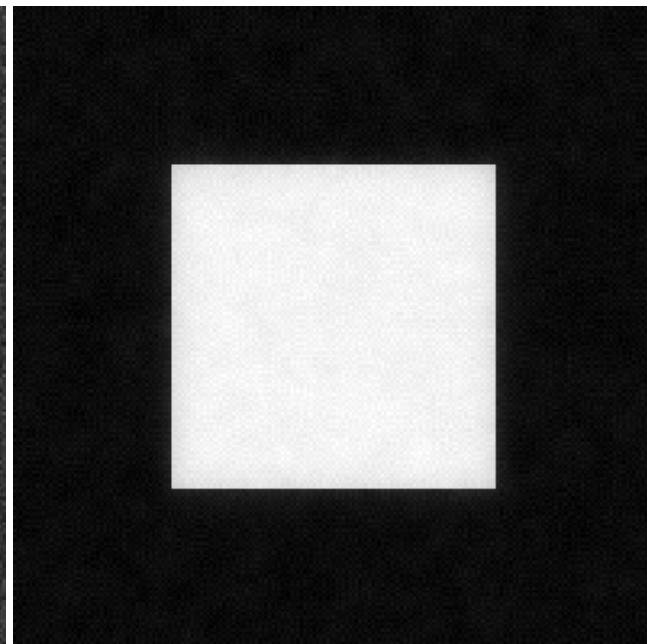
- Régularisation par la variation totale
Préservation des contours au cours du débruitage

Image bruitée



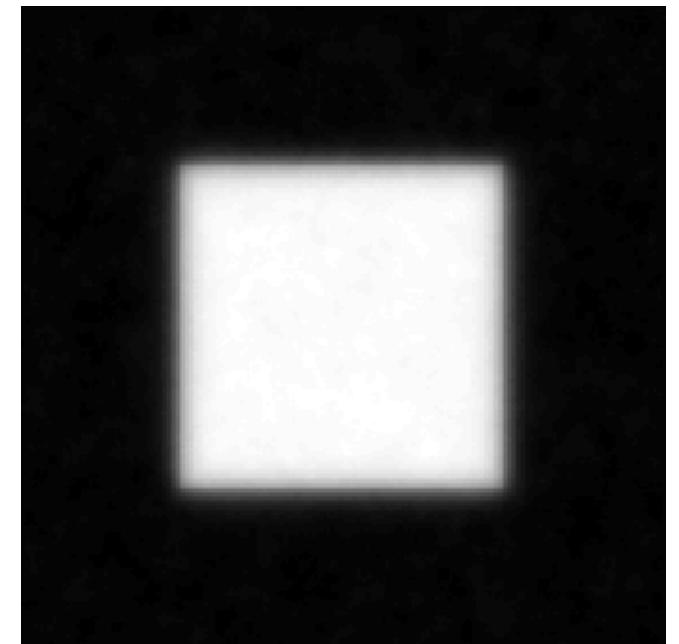
Débruitage non linéaire

$\Delta t=0.02$, 1000 itérations, $\lambda=100$



Débruitage linéaire

$\Delta t=0.01$, 1000 itérations, $\lambda=10$



Pour aller plus loin

- Régularisation par la variation totale
Profil d'une ligne

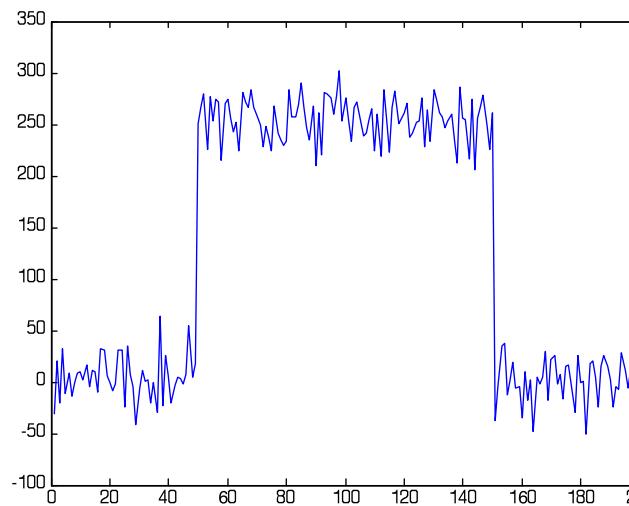
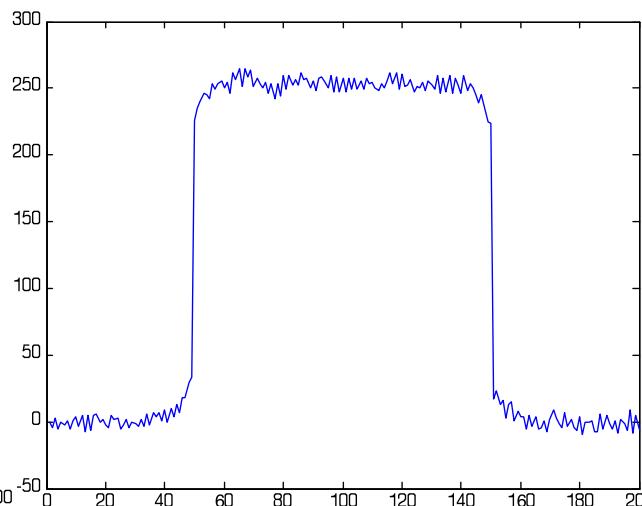
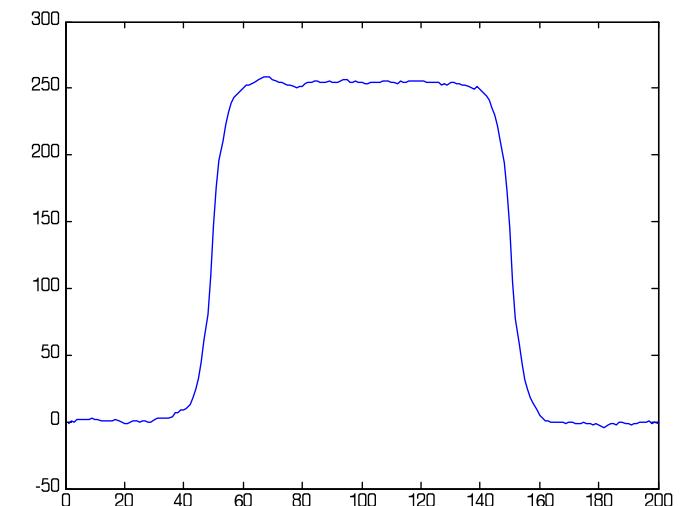


Image bruitée



Débruitage non linéaire



Débruitage linéaire

Pour aller plus loin

- Régularisation par la variation totale

$\Delta t=0.02$, 1000 itérations, $\lambda=20$



Pour aller plus loin

- Régularisation par la variation totale
 - Influence du paramètre de régularisation

$\lambda=1$



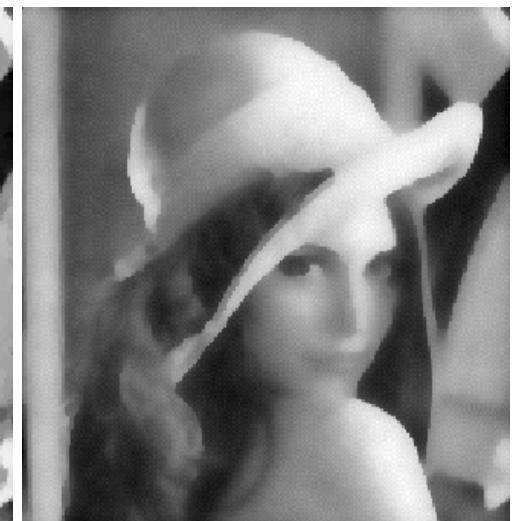
$\lambda=5$



$\lambda=20$



$\lambda=100$



Pour aller plus loin

- L'algorithme de Chambolle (2004)

Quelques notations:

Une image u de taille $N \times N$, $X = \mathbb{R}^{N \times N}$

Le gradient $(\nabla u)_{i,j} = ((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2)$

avec

$$(\nabla u)_{i,j}^1 = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } i < N, \\ 0 & \text{if } i = N, \end{cases}$$

$$(\nabla u)_{i,j}^2 = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } j < N, \\ 0 & \text{if } j = N, \end{cases}$$

Pour aller plus loin

- L'algorithme de Chambolle (2004)

Quelques notations:

La variation totale

$$J(u) = \sum_{1 \leq i, j \leq N} |(\nabla u)_{i,j}|,$$

La divergence

$$\begin{aligned} (\operatorname{div} p)_{ij} &= \begin{cases} p_{i,j}^1 - p_{i-1,j}^1 & \text{if } 1 < i < N, \\ p_{i,j}^1 & \text{if } i = 1, \\ -p_{i-1,j}^1 & \text{if } i = N, \end{cases} \\ &+ \begin{cases} p_{i,j}^2 - p_{i,j-1}^2 & \text{if } 1 < j < N, \\ p_{i,j}^2 & \text{if } j = 1, \\ -p_{i,j-1}^2 & \text{if } j = N, \end{cases} \end{aligned}$$

Pour aller plus loin

- L'algorithme de Chambolle (2004)

Il s'agit de résoudre

$$\min_{u \in X} \frac{\|u - g\|^2}{2\lambda} + J(u),$$

On peut montrer que $u = g - \pi_{\lambda K}(g)$.

Avec $\pi_{\lambda K}(g)$ qui peut se calculer comme la limite $\lambda \operatorname{div} p^n$

Et:

$$p^0 = 0$$

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau(\nabla(\operatorname{div} p^n - g/\lambda))_{i,j}}{1 + \tau |(\nabla(\operatorname{div} p^n - g/\lambda))_{i,j}|} \quad \tau \leq 1/8$$

Défloutage

- Défloutage

On cherche à minimiser

$$E(f) = \int_{Image} (g - Hf)^2 + \lambda \int_{Image} \|\nabla f\|^2$$

Or $H^* = H$, d'où l'EDP de défloutage

$$\frac{\partial f}{\partial t} = H(g - Hf) + \lambda \operatorname{div}(\nabla f)$$

Et sa discréétisation

$$\frac{f^{n+1} - f^n}{\Delta t} = H(g - Hf^n) + \lambda \Delta f^n$$

Défloutage

- Défloutage

$\Delta t=1$, 1000 itérations, $\lambda=0.1$



$\Delta t=1$, 1000 itérations, $\lambda=0.01$

Défloutage

- Défloutage + débruitage

$\Delta t=1$, 1000 itérations, $\lambda=0.1$

$\Delta t=0.5$, 1000 itérations, $\lambda=0.5$



Pour aller plus loin

- Défloutage avec un modèle de flou inconnu

IEEE TRANSACTIONS ON IMAGE PROCESSING, TO APPEAR.

Total Variation Blind Deconvolution

Tony F. Chan, C.K. Wong

Modèle de You et Kaveh (1996)

$$\min_{u,k} f(u, k) \equiv \min_{u,k} \frac{1}{2} \|k * u - z\|_{L^2(\Omega)}^2 + \alpha_1 \|u\|_{H^1}^2 + \alpha_2 \|k\|_{H^1}^2.$$

Ne permet pas d'obtenir un filtre discontinu.

Utilisation de la variation totale:

$$\begin{aligned} \min_{u,k} f(u, k) &\equiv \min_{u,k} \frac{1}{2} \|k * u - z\|_{L^2(\Omega)}^2 \\ &+ \alpha_1 \int_{\Omega} |\nabla u| dx + \alpha_2 \int_{\Omega} |\nabla k| dx. \end{aligned}$$

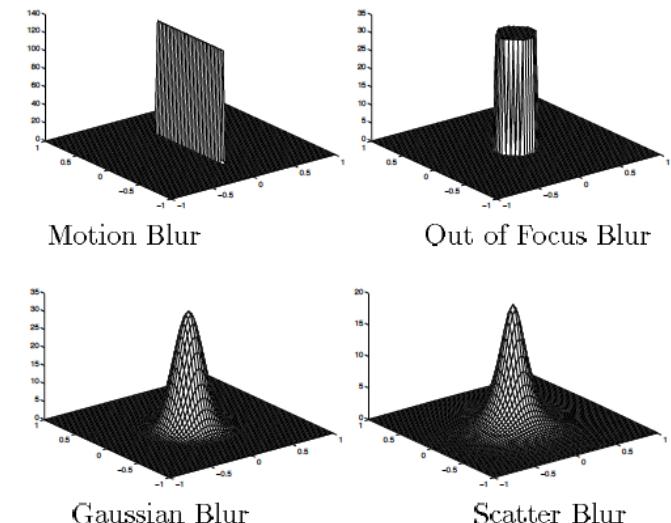
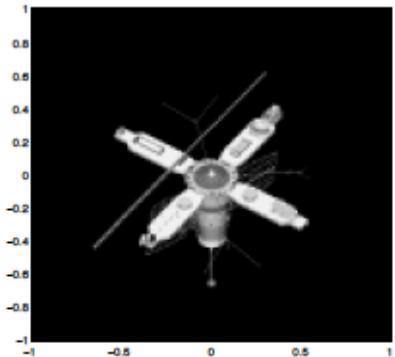
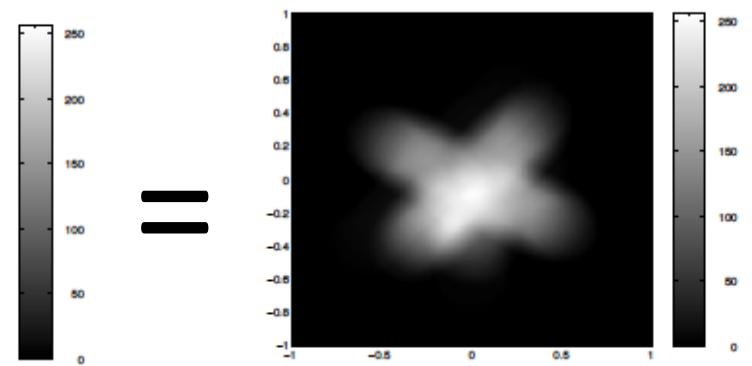
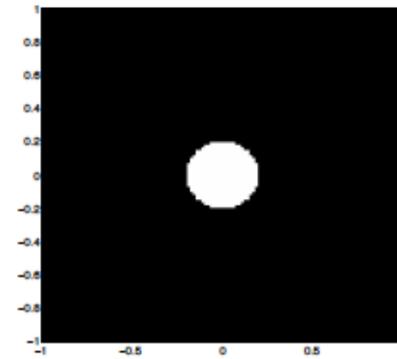


Figure 1. Some typical PSF

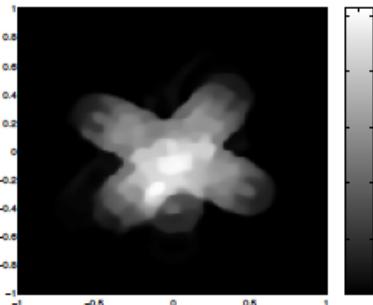
Pour aller plus loin



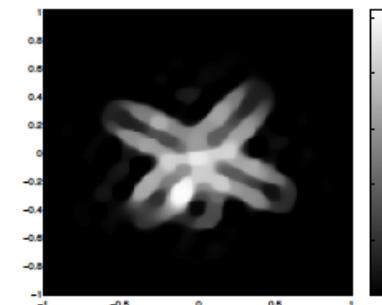
*



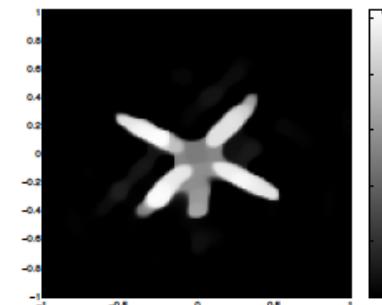
$\alpha_2 = 1e-7$



$1e-6$



$1e-5$



$1e-4$

