

Restauration d'images par approche variationnelle et EDP

Laure Blanc-Féraud

DR CNRS

Laboratoire I3S

INRIA Sophia Antipolis

Ariana

CONTENU global du cours

◆ Introduction

- Formation des images en satellitaire, échantillonnage
- Déconvolution : ex de pb mal posé
- Régularisation linéaire

$$\|g - Hu\|_2^2 + \lambda \|\nabla u\|_2^2$$

◆ Minimisation de critère : approche variationnelle

- Espaces fonctionnels pour la minimisation
- Existence, unicité de solution
- Introduction au calcul des variations,
- Introduction aux EDP en traitement d' image

$$\rightarrow \begin{cases} H^*(Hu - g) - \lambda \Delta u = 0 \\ \frac{\partial u}{\partial n} \Big|_{\partial\Omega} = 0 \end{cases}$$

$$\rightarrow \begin{cases} \frac{\partial u}{\partial t} = F(x, \nabla u, \nabla^2 u) \\ u(x, y, t = 0) = g(x, y) \end{cases}$$

CONTENU (suite)

◆ Régularisation non linéaire

➤ Régularisation L^2/L^1

➤ Espace BV, Variation Totale

➤ Solutions dans l'espace BV, dualité, algorithmes

$$\|g - Hu\|_2^2 + \lambda \int_{\Omega} \varphi(|\nabla u|) dx$$

$$\|g - Hu\|_2^2 + \lambda \int_{\Omega} |Du|$$

◆ Segmentation d'image, approche variationnelle

➤ Fonctionnelle de Mumford et Shah

➤ Notion de Γ -convergence : lien restauration/segmentation

◆ Et les représentations parcimonieuses ? $\|g - Hu\|_2^2 + \lambda \sum_{j,k} 2^{-\frac{j}{2}} |\langle u, \psi_{j,k} \rangle|$

➤ Régularisation dans le domaine de la transformée en ondelettes

➤ Solution dans un espace de Besov (et algorithmes)

➤ Modèles de régularisation / parcimonie

Construction d'une image

$$g = H(u) \oplus n$$

- ◆ g : une observation = des quantités physiques : optique, radar, laser, IR, champ magnétique, rayons X, ultrasons, émissions de photons...
- ◆ H : opérateur qui relie l'observation à la quantité que l'on cherche à imager u , à travers l'appareil de mesure et éventuellement un processus de reconstruction.
- ◆ u : image que l'on cherche à obtenir
- ◆ n : partie aléatoire du processus de reconstruction (bruit)

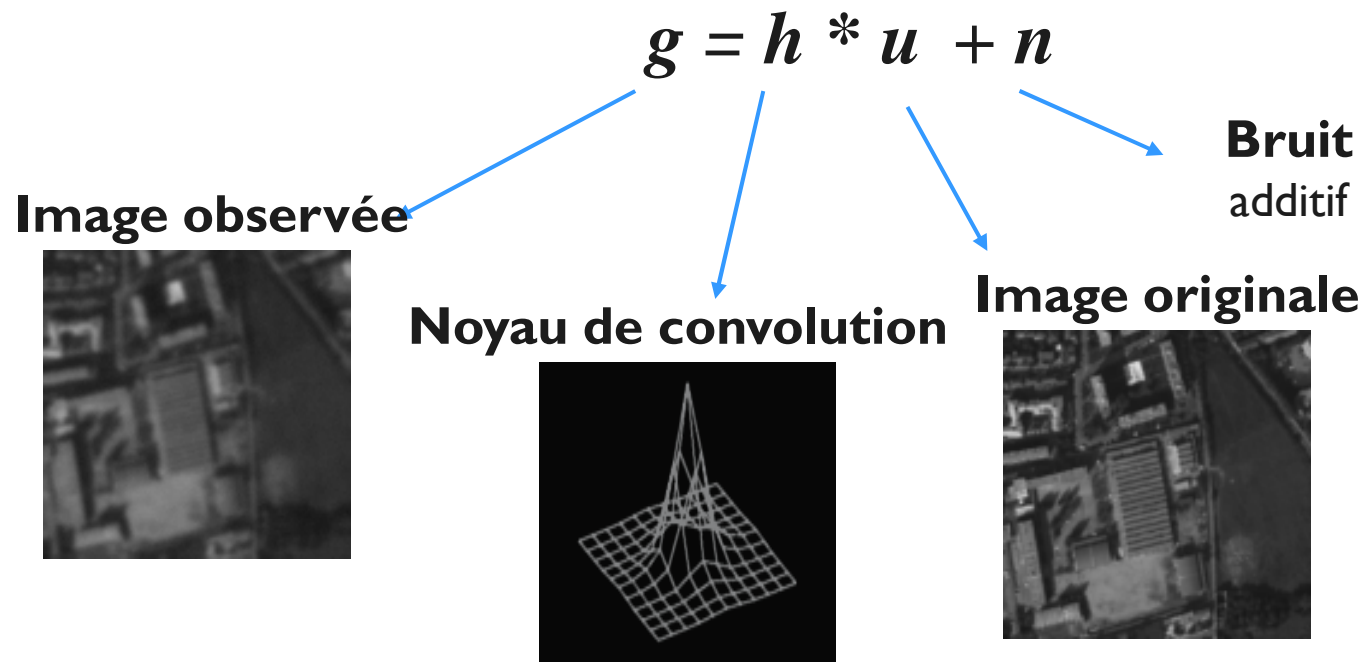
Exemples

◆ Linéaire : $H(u) = H.u$

■ Restauration d'une image optique :

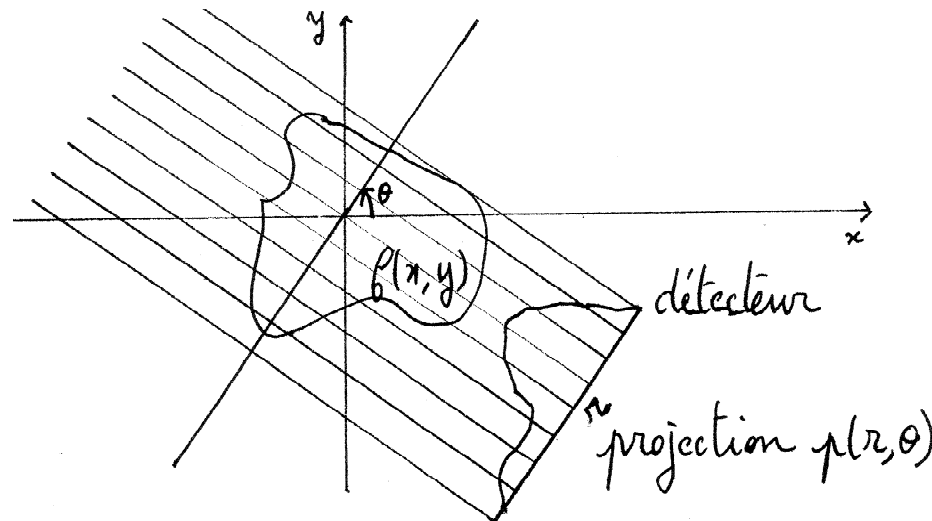
H est une convolution : $H.u = h * u$

H modélise les dégradations de l'appareil optique (diffraction, défocalisation, bougé, intégration sur les capteurs...)



Exemples

- ◆ Linéaires : $H(u) = H.u$
 - **Reconstruction tomographique**, appliquée dans les domaines de la médecine, des géosciences, ...
 - Reconstruire le volume d'un objet (le corps humain dans le cas de l'imagerie médicale), à partir d'une série de mesures effectuées à l'extérieur de l'objet.
Exemple de la tomographie à rayons X



$H =$ matrice de Radon = projection 1D selon l'angle θ

Exemples

- ◆ Linéaires : $H(u) = H.u$

- Super résolution, **échantillonnage irrégulier**

On observe une image g irrégulièrement échantillonnée. u est la même image sur une grille régulière.

$$g = \Delta_{\Lambda}(u) + n \quad \text{où} \quad \Delta_{\Lambda}(\cdot) = \sum_{\lambda_k \in \Lambda} \delta(\cdot - \lambda_k)$$

λ_k sont les positions des échantillons irrégulièrement répartis. Peut se modéliser par transformée de Fourier discrète

$$g = SFu + n$$

F Transformée de Fourier directe régulière

S Transformée de Fourier inverse irrégulière

$$S : \mathbf{C}^N \rightarrow \mathbf{R}^M$$

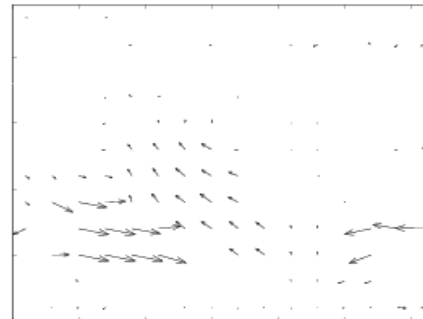
$$a \rightarrow g = (g_k)_{k=1, \dots, M}, \quad g_k = \sum_{l=1}^N a_l \exp\left[\frac{2i\pi}{N} \langle \lambda_k, l \rangle\right]$$

Exemples

- ◆ Non Linéaires :

- Flot optique

on observe $u(x,t)$ et $u(x,t+1)$ et on cherche le mouvement apparent v tel que $u(x+v,t) = u(x,t+1)$.



Linéarisation par dérivation : équation du flot optique

$$\nabla u(x,t+1) \cdot v(x,t) + I_t(x,t) = 0$$

- Diffraction inverse microonde, contrôle non destructif : équations de Maxwell ...

◆ **Problème direct** $g = H(u) \oplus n$

C'est l'équation de construction de l'image, la modélisation mathématique des phénomènes physiques de l'acquisition.

Définit g à partir de H , u et n .

◆ **Problème inverse**

A partir des données observées g , et de la modélisation du problème direct, trouver u en supposant H connu et les paramètres de loi de n connus.

→ H n'est souvent que partiellement connu,

→ Il faut souvent estimer les paramètres du bruit au préalable ou en même temps.

Restauration d'image : déconvolution, débruitage

- ◆ Modèle simple :
 - Convolution,
 - Mais difficile : problème mal posé.
- ◆ Deux exemples d'acquisition
 - Imagerie satellitaire
 - Imagerie biologique par microscopie confocale
- ◆ Et toujours présent : la diffraction

Observation de la Terre

- ◆ Illustration sur un exemple en imagerie optique satellitaire



SPOT5 (Stockholm) France+, 2,5m



Ikonos (San Juan) USA, 1m



Quickbird(Beijing) USA, 61cm

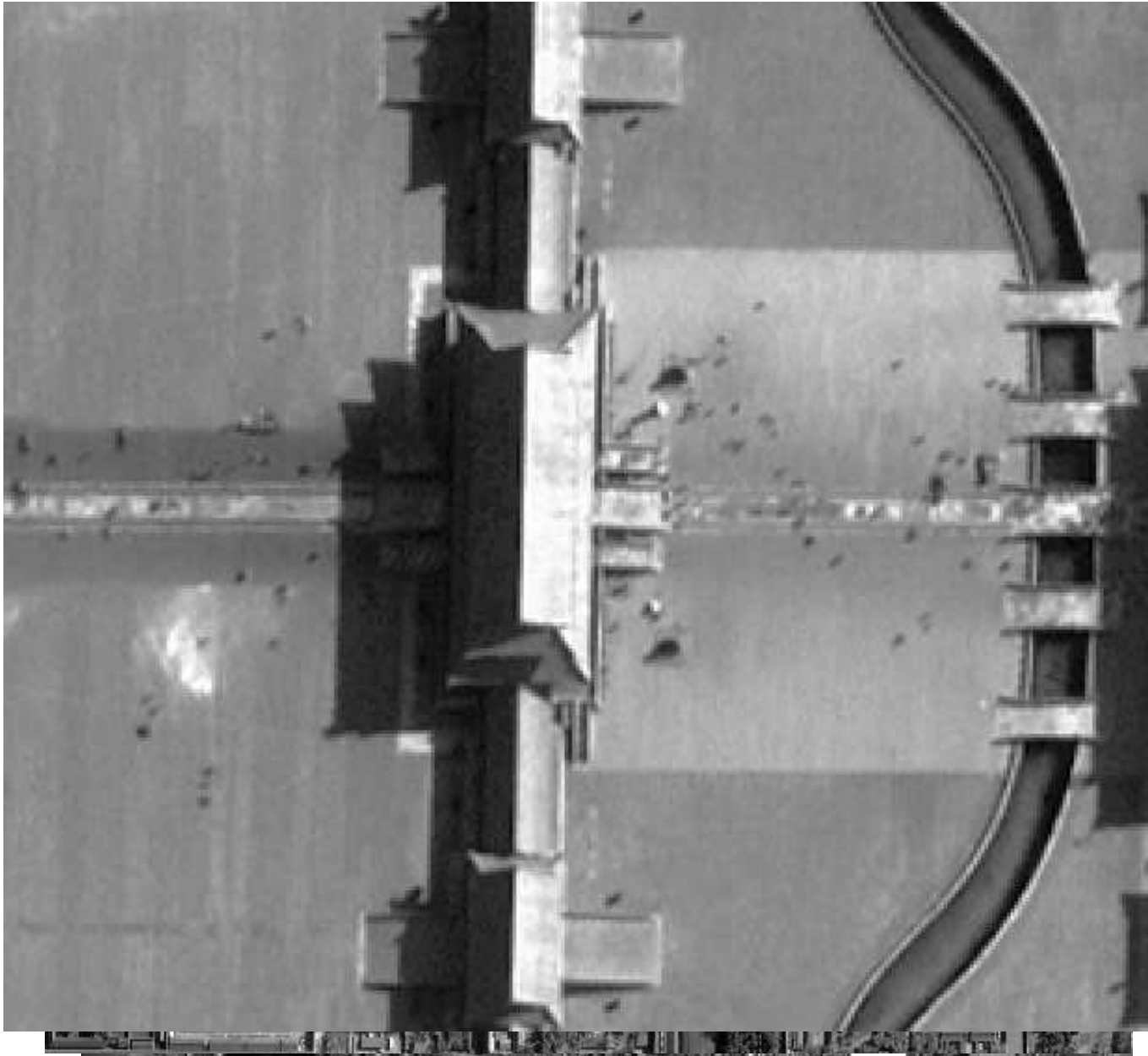


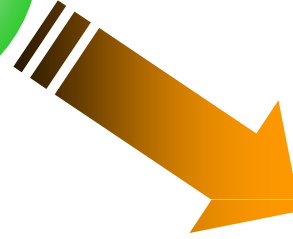
Image aérienne (Amiens)



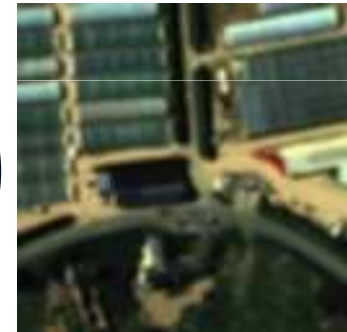
Observation



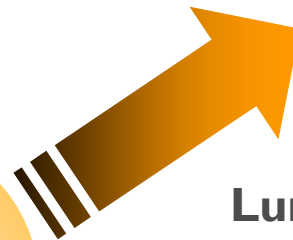
**Modèle
d'observation**



**Modèle
d'image
observée**



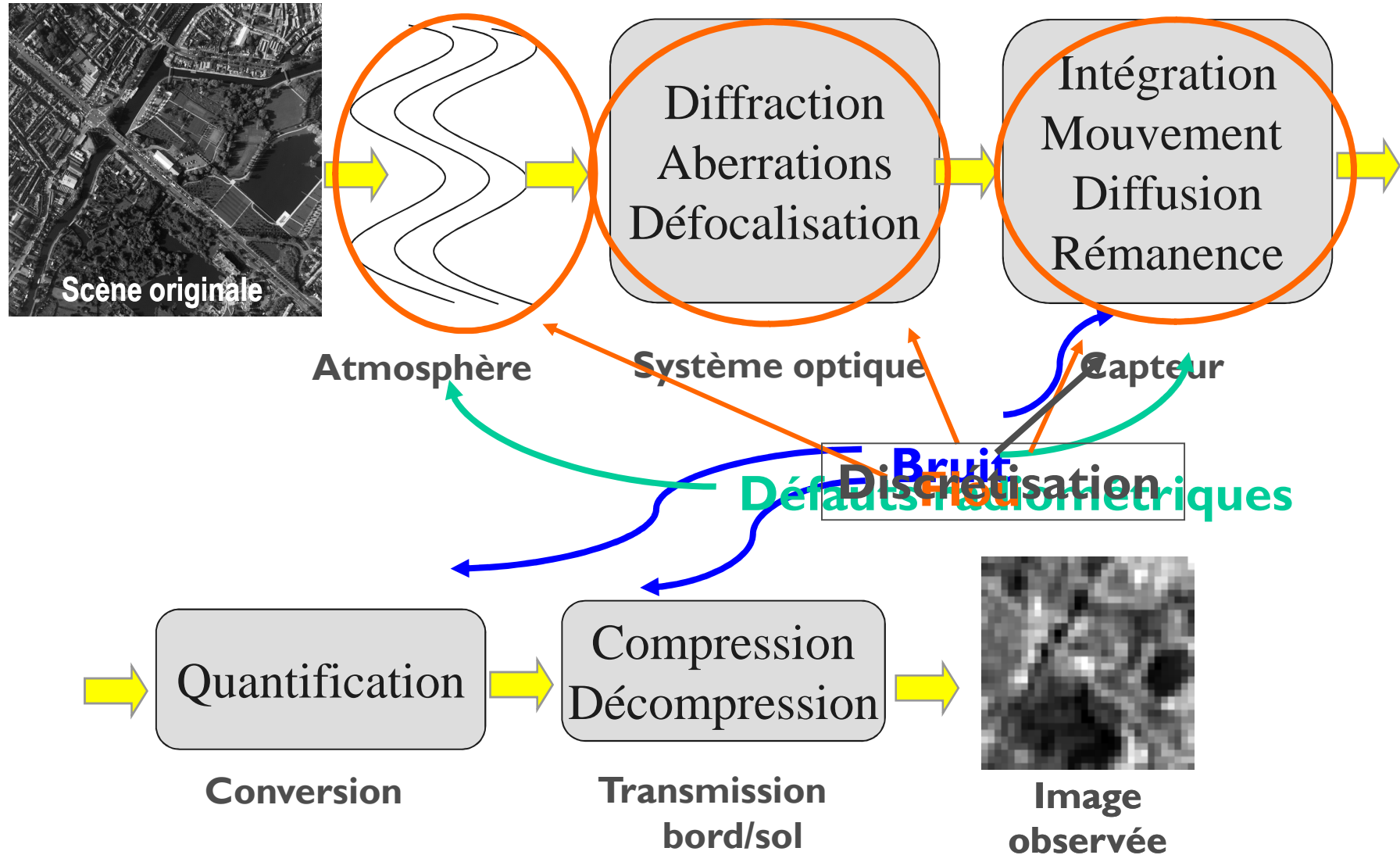
**scène
paysage
vérité terrain**



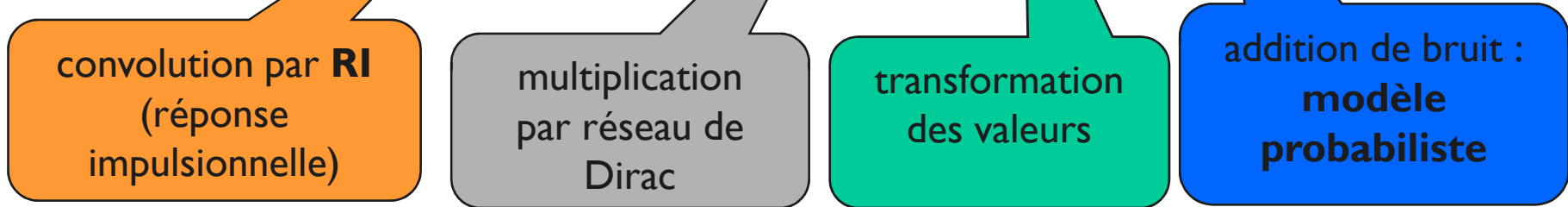
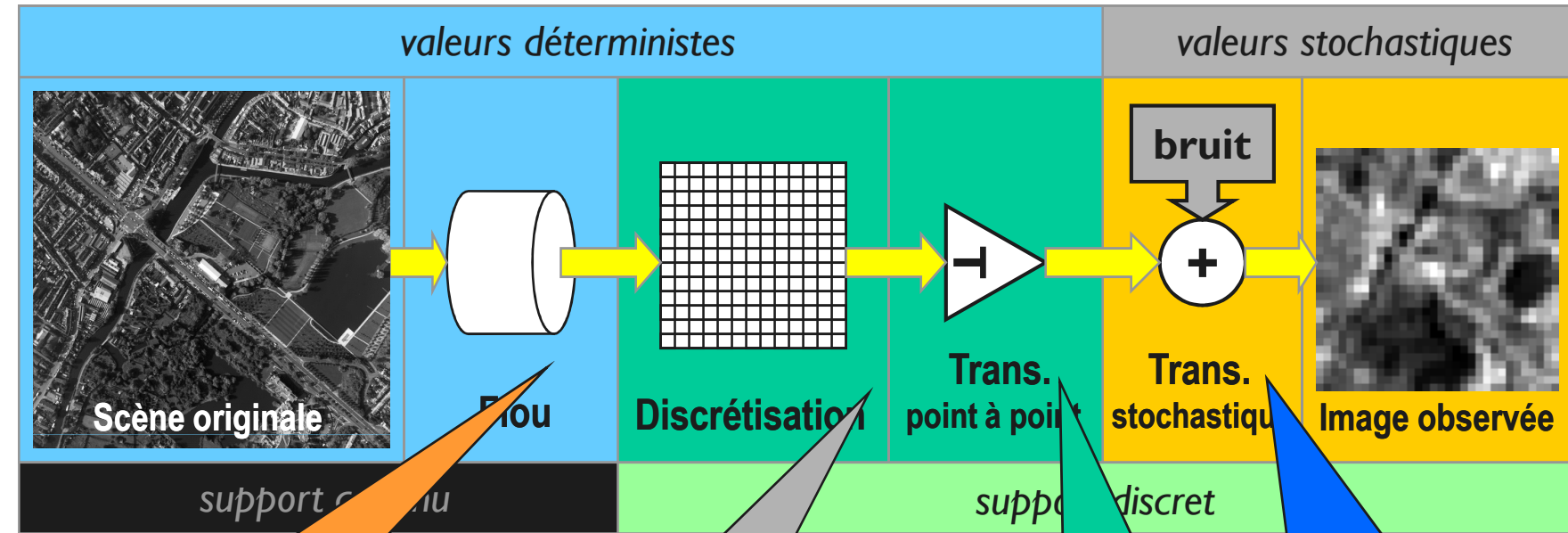
Luminance de la scène
dans la bande spectrale
considérée et sous l'angle
d'observation choisi

$$g : \Omega \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$$

Chaîne des dégradations



Modèle d'observation



$$g_{i,j} = T_{i,j} \left((RI * u)_{d(i,j)} \right) + N_{i,j}$$

Dispersion spatiale: RI

On appelle **réponse impulsionnelle** (RI) ou fonction de dispersion la réponse du système imageur à une distribution d'intensité ponctuelle (distribution de Dirac) au niveau de la scène. La RI s'exprime en continu dans le repère image. Elle est positive et normalisée.

$$RI : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^+$$

$$x \rightarrow RI(x) \quad \text{et} \quad \int_{\Omega} RI(x) dx = 1$$

En toute généralité RI dépend de la position et de la distribution d'intensité dans la scène

Dispersion spatiale: FTM

Hypothèse de stationnarité + indépendance par rapport à la scène :

Flou = produit de convolution

$$g_{x,y} = (RI * u)_{x,y} = \int_{\Omega} RI(x-s, y-t)u(s,t)dsdt$$

Dans l'espace de Fourier

$$F(g)_{u,v} = (FTM)_{u,v} (FTP)_{u,v} \cdot F(u)_{u,v}$$

FTM = Fonction de Transfert de Modulation, $FTM = |F(PSF)|$

FTP = Fonction de Transfert de Phase

De la scène à l'image formée au niveau du capteur :

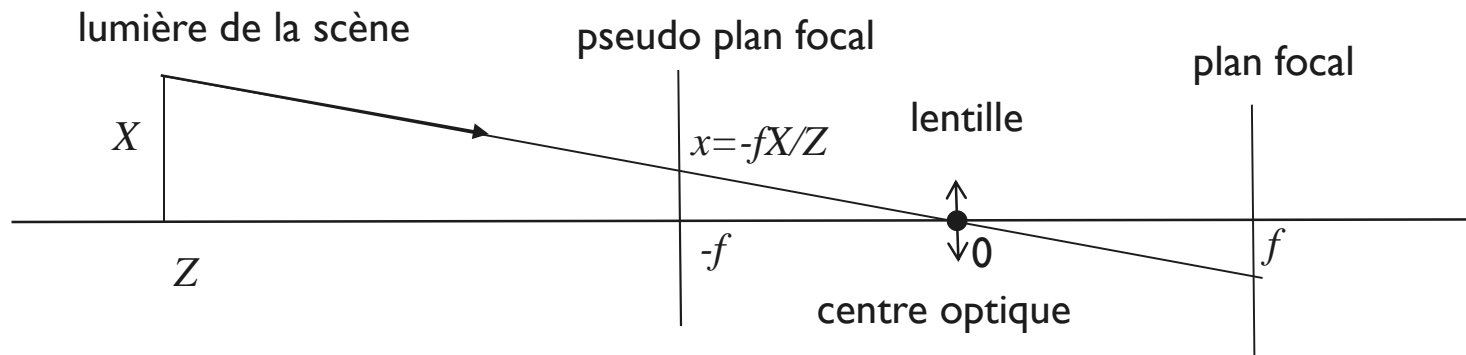
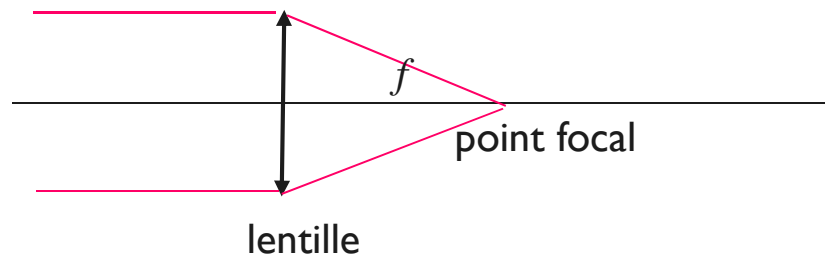
Atmosphère, système optique, capteur

Flou système optique: diffraction

- ◆ **Diffraction** : due à l'ouverture limitée du système optique.
- ◆ Théorème de diffraction de Fraunhofer: l'image physique formée à travers une ouverture plane O est donnée par

$$v = K_{diffrac} * u \quad K_{diffrac}(x) = \frac{1}{2f^2} \left| \int_o \exp\left(\frac{2i\pi\langle x, \xi \rangle}{\lambda f}\right) d\xi \right|^2$$

où u est l'image idéale, λ la longueur d'onde (en lumière visible $\lambda=0.6\mu m$), f la longueur focale :



Flou système optique: diffraction (suite)

- ◆ Lorsque O est un disque de diamètre D la PSF diffraction est

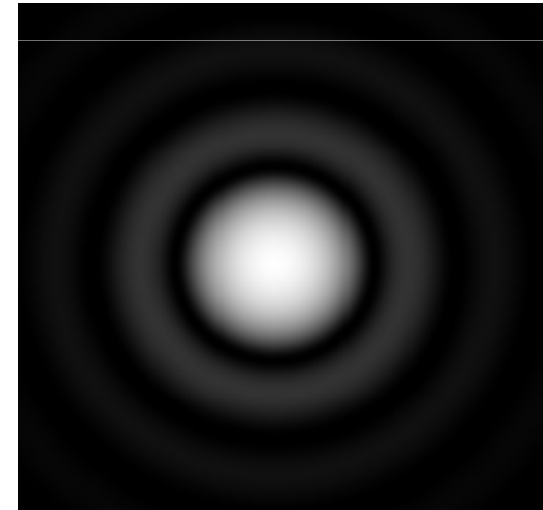
$$K_{diffrac}(x) = C \left(\frac{2J_1(r)}{r} \right)^2 \quad \text{où} \quad r = \frac{\pi D |x|}{\lambda f}, \quad C = \frac{\pi^2 D^4}{32 f^2} \quad \text{et} \quad J_1(t) = \frac{1}{\pi} \int_0^\pi \cos(\theta - t \sin \theta) d\theta$$

J_1 est la fonction de Bessel de première espèce à l'ordre 1.

- ◆ L'image de $K_{diffrac}$ est le disque (motif) d'Airy → tâche circulaire brillante et des anneaux autour en atténuation.

Le rayon de la tâche d'Airy donne une idée de la dimension du détail le plus petit qui puisse être reproduit par une optique idéale. Le premier zéro de J_1 est en $t \approx 1.22\pi$, le rayon de la tâche d'Airy est

$$r_A = 1.22 \frac{\lambda f}{D}$$



Flou système optique

- ◆ **Diffraction**
- ◆ **Défocalisation** : la focalisation n'est pas sur le plan des objets d'intérêt.
- ◆ **Aberration** : sphériques, géométriques, chromatiques...

Capteur

◆ Intégration

- Chaque capteur est un intégrateur. En l'absence de mouvement, en supposant une matrice de capteurs, chacun est modélisé par une cellule rectangulaire de taille $p_x \times p_y$: zone photosensible, et de répartition sur une zone rectangulaire de pas p_{ex} et p_{ey} : taille physique des pixels.

$$u_{k,l} = \int_{[-p_x, p_x] \times [-p_y, p_y]} u(kp_{ex} + x, lp_{ey} + y) dx dy$$

- Encore une fois on retrouve un modèle de convolution. La PSF liée à l'intégration sur le capteur est

$$PSF_{int} = 1_{[-p_x, p_x] \times [-p_y, p_y]} \quad \text{et} \quad u_{k,l} = (PSF_{int} * u)(kp_{ex}, lp_{ey})$$

- La FTM associée est

$$(FTM_{int})_{u,v} = \frac{\sin\left(\frac{\pi u p_x}{p_{ex}}\right)}{\frac{\pi u p_x}{p_{ex}}} \frac{\sin\left(\frac{\pi v p_y}{p_{ey}}\right)}{\frac{\pi v p_y}{p_{ey}}}$$

- Remarque : la réponse du capteur n'est en général pas linéaire, en particulier pour les petites et les grandes valeurs. On observe donc $\phi(u_{k,l})$.

◆ Facteurs électroniques : Diffusion des charges , Rémanence

Dispersion spatiale

- ◆ La PSF est donnée (au minimum) par

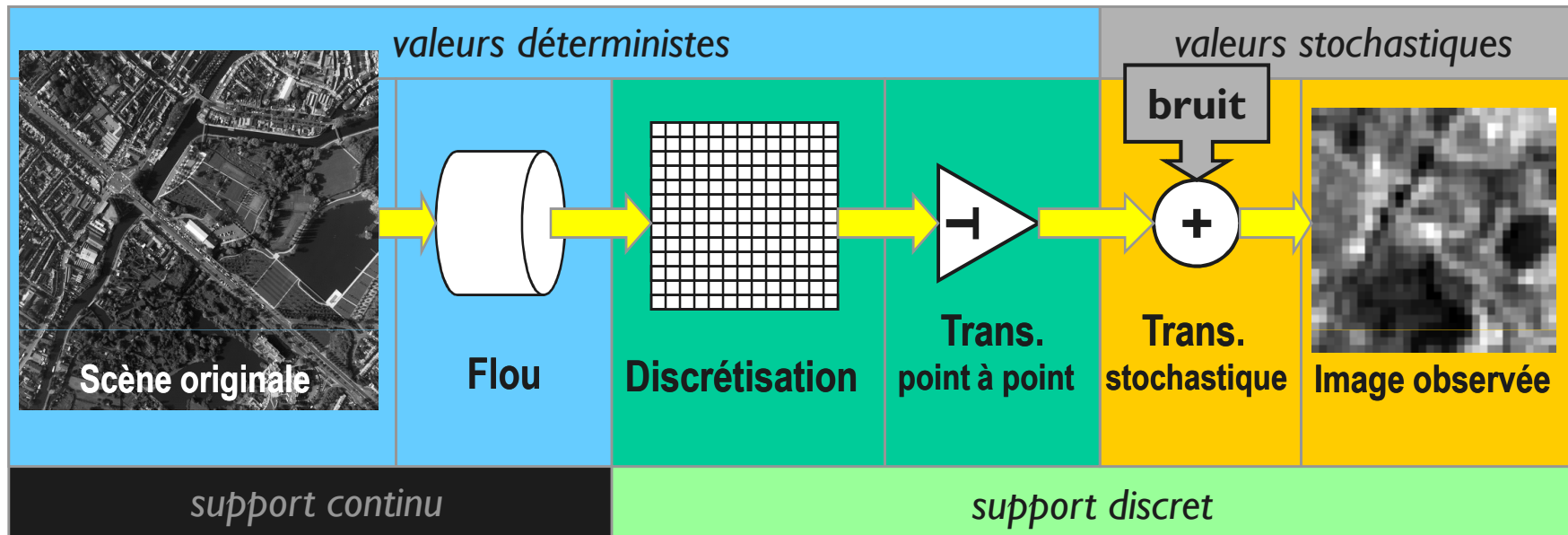
$$PSF = PSF_{optique} * PSF_{capteur}$$

ce qui correspond à la FTM

$$FTM = FTM_{optique} \cdot FTM_{capteur}$$

- ◆ Utilisation de modèles simplifiés, avec peu de paramètres, ou approximation gaussienne

Modèle d'observation



- ◆ Remarques sur la discrétisation en 2D

Discrétisation : échantillonnage

◆ Echantillonnage idéal:

Le processus d'échantillonnage est décrit grâce au théorème d'échantillonnage de Shannon, qui stipule qu'une fonction à bande limitée peut être reconstruite exactement à partir d'échantillons discrets pris sur une grille appropriée.

Théorème (Shannon): Soit $u \in S$ et $\tau > 0$.

$$\text{si } \text{supp}(\hat{u}) \subset \left[-\frac{\pi}{\tau}, \frac{\pi}{\tau} \right]^2 \text{ alors}$$
$$\forall (x, y) \in \mathbb{R}^2, \quad u(x, y) = \sum_{k, l \in \mathbb{Z}^2} u(k, l) \text{sinc}\left(\frac{x}{\tau} - k\right) \cdot \text{sinc}\left(\frac{y}{\tau} - l\right) \quad (6)$$

$\text{supp}(\cdot)$ est le support défini par $\text{supp}(\hat{u}) = \{\zeta \mid \hat{u}(\zeta) \neq 0\}$

Cela veut dire que $\zeta_{\max} \leq \frac{\pi}{\tau}$ soit $2\pi f_{\max} \leq \frac{\pi}{\tau}$ soit $\tau \leq \frac{1}{2f_{\max}}$

Le pas d'échantillonnage est plus petit que $1/2f_{\max}$.

Ce théorème peut être généralisé à des fonctions u moins régulières.

Echantillonnage

La remarque fondamentale est la suivante : discrétiser en espace (de pas τ) revient à périodiser en fréquence (de période $2\pi/\tau$) et inversement.

Echantillonner et **périodiser** sont des fonctions **duales** pour la TF.

- ◆ En effet soit $\delta_{x,y} \in S'$ la **distribution de Dirac** définie par

$$\langle \delta_{x,y}, \varphi \rangle = \varphi(x, y) \quad \forall \varphi \in S, \quad \varphi \text{ fonction test}$$

et S l'espace de Schwartz des fonctions C^∞ à décroissance rapide.

- ◆ On définit le **peigne de Dirac** : $\Pi_\tau = \sum_{k,l \in \mathbb{Z}^2} \delta_{(k\tau, l\tau)}$

- ◆ Sa transformée de Fourier est donnée en fonction du peigne de Dirac espacé de $2\pi/\tau$:

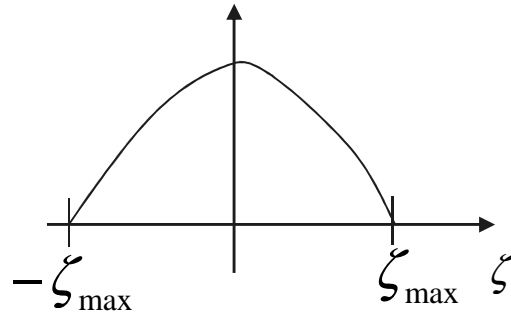
$$\hat{\Pi}_\tau = \left(\frac{2\pi}{\tau} \right)^2 \Pi_{\frac{2\pi}{\tau}}$$

- ◆ la discrétisation de u : $u \rightarrow u \cdot \Pi_\tau$ est dans Fourier $2\pi/\tau$ périodique

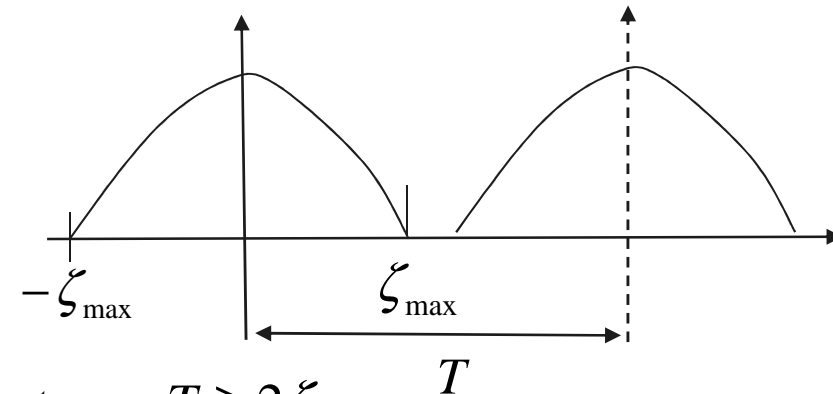
$$\hat{u} \rightarrow \frac{1}{(2\pi)^2} \hat{u} * \hat{\Pi}_\tau = \frac{1}{\tau^2} \hat{u} * \Pi_{\frac{2\pi}{\tau}} = \frac{1}{\tau^2} \sum_{k,l} \hat{u} \left(s - \frac{2\pi k}{\tau}, t - \frac{2\pi l}{\tau} \right)$$

Echantillonnage

En 1D



Après périodisation de période T :

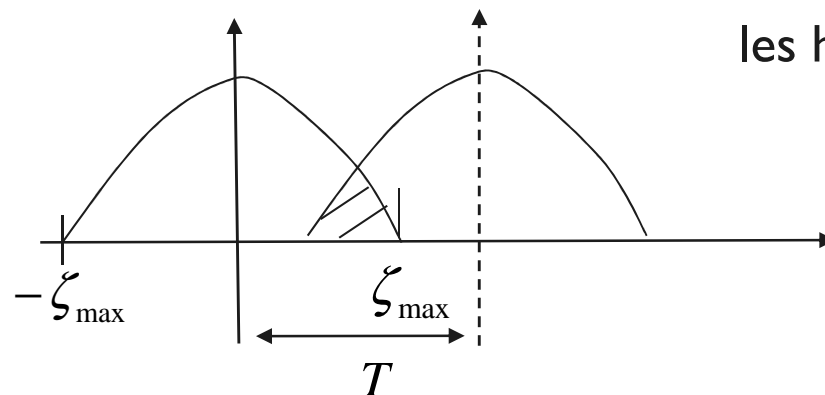


Pour ne pas avoir de repliement, il faut que $T \geq 2\zeta_{\max}$

donc que $\frac{2\pi}{\tau} \geq 2 \cdot 2\pi f_{\max}$ soit $\tau \leq \frac{1}{2f_{\max}}$

Repliement de spectre – aliasing : Si la période d'échantillonnage est trop grande $\tau' > 1/2f_{\max}$ alors ce sous échantillonnage entraîne de l'aliasing

les hautes fréquences sont perdues



Echantillonnage/ringing

Interprétation: pour que la formule de reconstruction soit exacte il faut

1/ que le signal soit à bande limitée. Ceci pose problème en image car les images ont des discontinuités le long des contours des objets, donc des hautes fréquences. D'après le théorème de Shannon, pour ne pas avoir de repliement, il faut couper toutes les hautes fréquences de l'image géométrique sur le plan focal et considérer

$$\hat{v} = \hat{u} \cdot 1_{\left[\frac{-\pi}{\tau}, \frac{\pi}{\tau} \right]}$$

Ce n'est pas une bonne manière de faire car Fourier inverse conduit à une convolution en x et y avec une fonction sinc, ce qui entraîne un phénomène de ringing au niveau des contours : contours parallèles qui décroissent, espacés de 2τ .

2/ un ensemble infini d'échantillons pour reconstruire exactement le signal ce qui n'est pas le cas en pratique. La fonction sinc est à décroissance lente ce qui entraîne des erreurs importantes avec un nombre fini d'échantillons.

Point 1 : un bon capteur est flou

Un bon capteur d'image doit avoir une FTM faible à la fréquence de coupure.

→ Un bon capteur d'image renvoie des images floues !

Autrement dit : il est toujours possible de gagner en résolution numériquement, par déconvolution de l'image par la fonction d'acquisition.

→ Il y a toujours de la diffraction

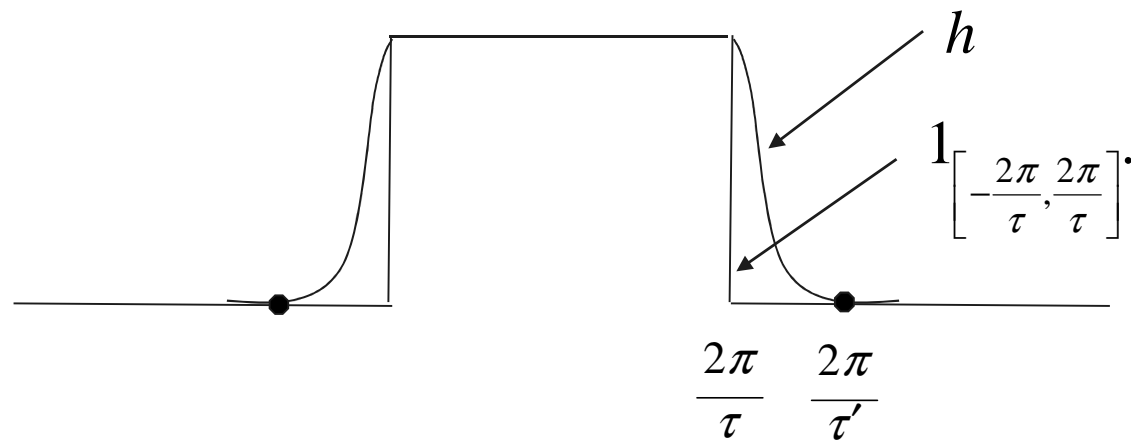
→ Heureusement qu'il y en a pour échantillonner

→ Si on sait la modéliser et l'estimer, on peut toujours restaurer numériquement l'image.

Point 2 : Echantillonnage/reconstruction

On peut sur-échantillonner avec $\tau' < \tau$ et prendre une fenêtre lissée

$$\text{sur } \left[-\frac{2\pi}{\tau'}, \frac{2\pi}{\tau'} \right] \text{ t.q. } h = \begin{cases} 1 \text{ sur } \left[-\frac{2\pi}{\tau}, \frac{2\pi}{\tau} \right] \\ 0 \text{ sur } \mathbb{R} \setminus \left[-\frac{2\pi}{\tau'}, \frac{2\pi}{\tau'} \right] \end{cases} \text{ et } 0 < h < 1 \text{ ailleurs}$$



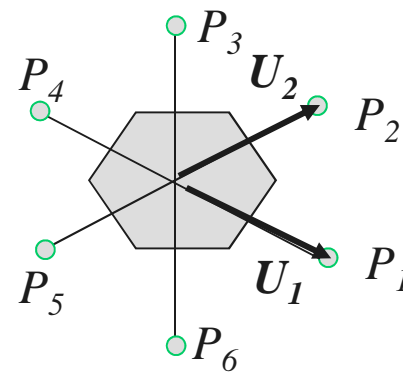
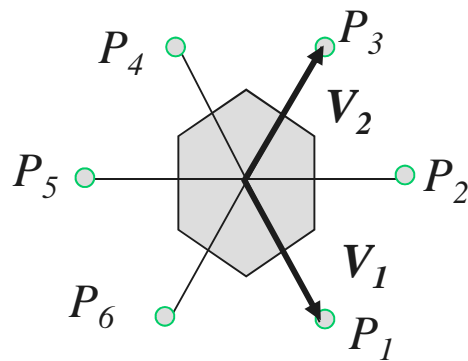
h est beaucoup plus régulière que la fonction porte et aura une décroissance dans Fourier beaucoup plus rapide.

Discrétisation 2D: réseau quelconque

- ◆ Bernard Rougé du CNES a introduit une notion importante qui consiste à considérer le triplet : instrument, échantillonnage, bruit pour optimiser l'échantillonnage en fonction de la FTM et du bruit.
- ◆ Pour cela il faut introduire la notion de cellule réciproque : support fréquentiel lié à l'échantillonnage.

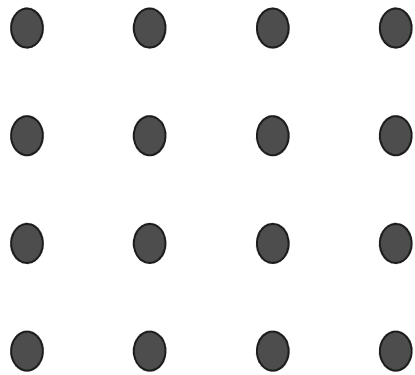
On suppose qu'on a un réseau régulier de points en géométrie quelconque. Ce réseau est engendré par deux vecteurs non colinéaires V_1 et V_2 . Soit la matrice $V = [V_1, V_2]$. La matrice U réciproque dans l'espace spectral est donnée par U telle que $Id = U^t V$. (Id est la matrice identité 2×2), U^t est la transposée de U (lié à la transformée de Fourier du réseau de diracs).

Cellule directe: Les vecteurs V_1 et V_2 engendrent par combinaison linéaires à coefficients entiers un ensemble de points du plan. Nous considérons les 6 points les plus proches de l'origine P_1, P_2, \dots, P_6 . La frontière de la cellule directe est constituée des médiatrices des segments OP_i . On construit de même à partir des vecteurs réciproques U_1, U_2 la cellule réciproque.

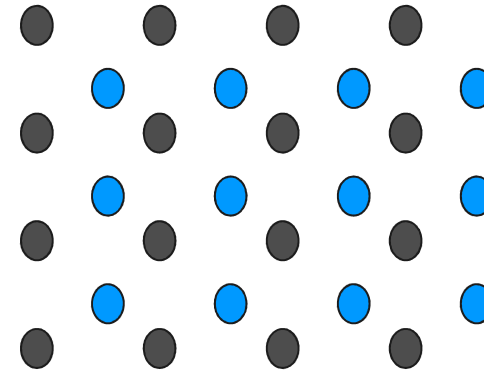


Réseaux orthogonal et quinconce

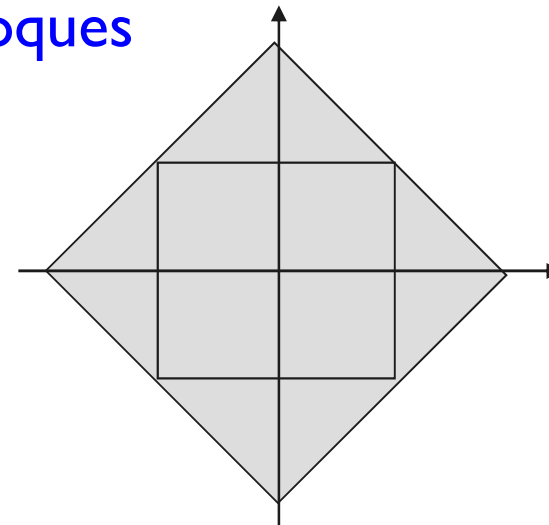
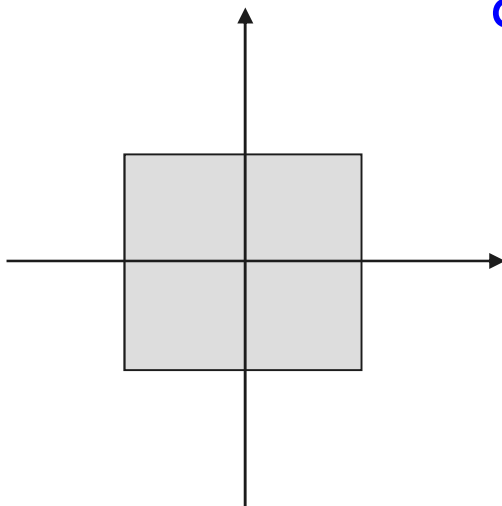
Réseau orthogonal



Réseau quinconce

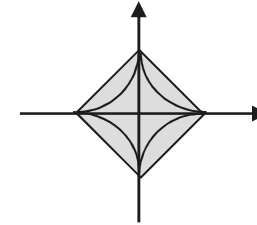


cellules réciproques

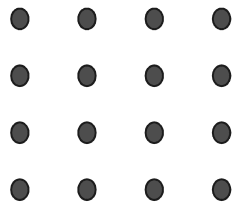


Echantillonnage optimisé

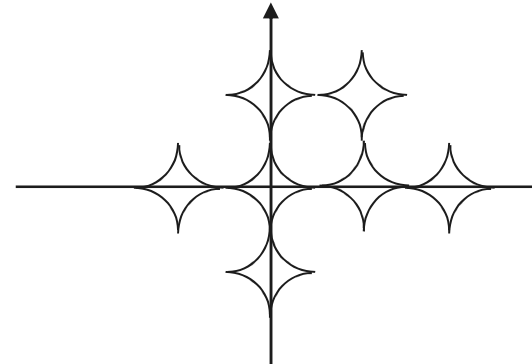
Si le spectre de la fonction à échantillonner ne contient pas de hautes fréquences en x, y en même temps, donc est à support tel que :



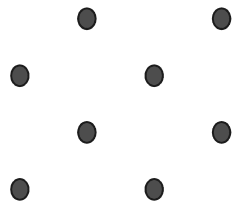
Si la grille est rectangulaire:



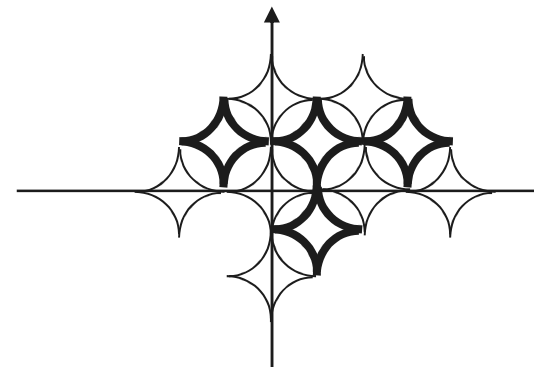
le spectre est



Si la grille est quinconce
(2 fois moins de points)



le spectre est



Triplet FTM, échantillonnage, bruit

- ◆ **Cellule réciproque** : support fréquentiel lié à l'échantillonnage.
- ◆ **Support utile** de la FTM : domaine fréquentiel où le signal d'amplitude A est tel que $A.FTM > k \sigma$
- ◆ Echantillonnage est optimal si la cellule réciproque est adaptée au support utile de la FTM

[Rougé, Théorie de la chaîne image et restauration d'image optique à bruit final borné, mémoire d'habilitation à diriger des recherches, 1997]

→ **triplet instrument, échantillonnage, bruit** : (FTM, R, σ)

- ◆ triplet idéal : la **cible**. On ne cherche pas l'image continue (la scène) mais une approximation échantillonnée de cette image.

Triplet idéal

- ◆ Bruit : bruit quantification.
- ◆ FTM idéale : correspond à l'échantillonnage idéal
- ◆ Difficulté liée au principe d'incertitude d'Heisenberg (précision temps%fréquence). Impossibilité d'avoir un support borné en espace et en fréquence.
- ◆ Optimisation de la fonction d'échantillonnage pour avoir la compacité spatiale et fréquentielle simultanément (concentration d'énergie). On n'utilise plus la fonction porte (sinc) mais une fonction à support borné en spatial et infini en fréquentiel mais telle que 94% à 99% de l'énergie soit dans la cellule réciproque.
- Fonctions **sphéroïdales aplaties** : « **prolate** » (cercle inscrit, cercle circonscrit).

[A. Lannes, E. Antérieux & K Bouyoucef « Fourier interpolation and reconstruction via Shannon-type techniques » *Journal of modern optics*, janv. 1996]

Discrétisation

$$u: \Omega \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$$

◆ Continu

- ◆ On cherche à reconstruire
- ◆ h = opérateur de flou discret

$$u: (\mathbb{N} \times \mathbb{N})$$

discret

$$(\mathbf{r}_I * u)_d = u$$

$$g = h * (\mathbf{r}_I * u)_d = (\mathbf{r}_I * u)_d$$

- ◆ d'où

$$FTM \cdot FTM_I \hat{u} = FTM_0 \hat{u}$$

- ◆ On a et

$$FTM = \frac{FTM_0}{FTM_I}$$

$$h = \left(F^{-1}(FTM) \right)_d$$

Bruits

- ◆ **Bruit quantique** : accumulation électrons, comptage de photons. Bruit de statistique de Poisson.
- ◆ **Bruit thermique et bruit de lecture** : bruit gaussien.
- ◆ Bruit de quantification : uniforme, variance faible relativement aux autres bruit.
- ◆ Bruit de compression : coloré, corrélé, non stationnaire. Assimilé à du bruit gaussien en première approximation.
- ◆ Bruit de transmission : perte de paquets... Difficile à prendre en compte, de même que le bruit de compression. On suppose transmission parfaite.

Hypothèses (réalistes) : Indépendance des bruits entre eux et indépendance du bruit d'un pixel à l'autre (bruit blanc) et stationnarité de la loi (même loi en chaque pixel).

Bruit de **Poisson** + bruit **gaussien** → approximation par un **bruit additif gaussien blanc** de moyenne nulle et de variance dépendant de l'intensité $u_{i,j}$ au pixel (i,j) considéré

$$P(n/u) = \prod_{i,j} \mathcal{N}_2 \left(0, (A + B u_{i,j}) \text{Id} \right)$$

Bruit de loi stationnaire normale, de variance non stationnaire.

A fort taux de compte (scène réelle en illumination normale), la loi de Poisson tend vers la loi gaussienne et le bruit est gaussien blanc $\mathcal{N}(0, \sigma^2 \text{Id})$

Rappel

- ◆ Soit n le bruit. C'est une variable aléatoire multidimensionnelle sur un champ de pixels (i,j) $i,j=1,\dots,N$. Sa matrice d'autocorrélation est la matrice symétrique

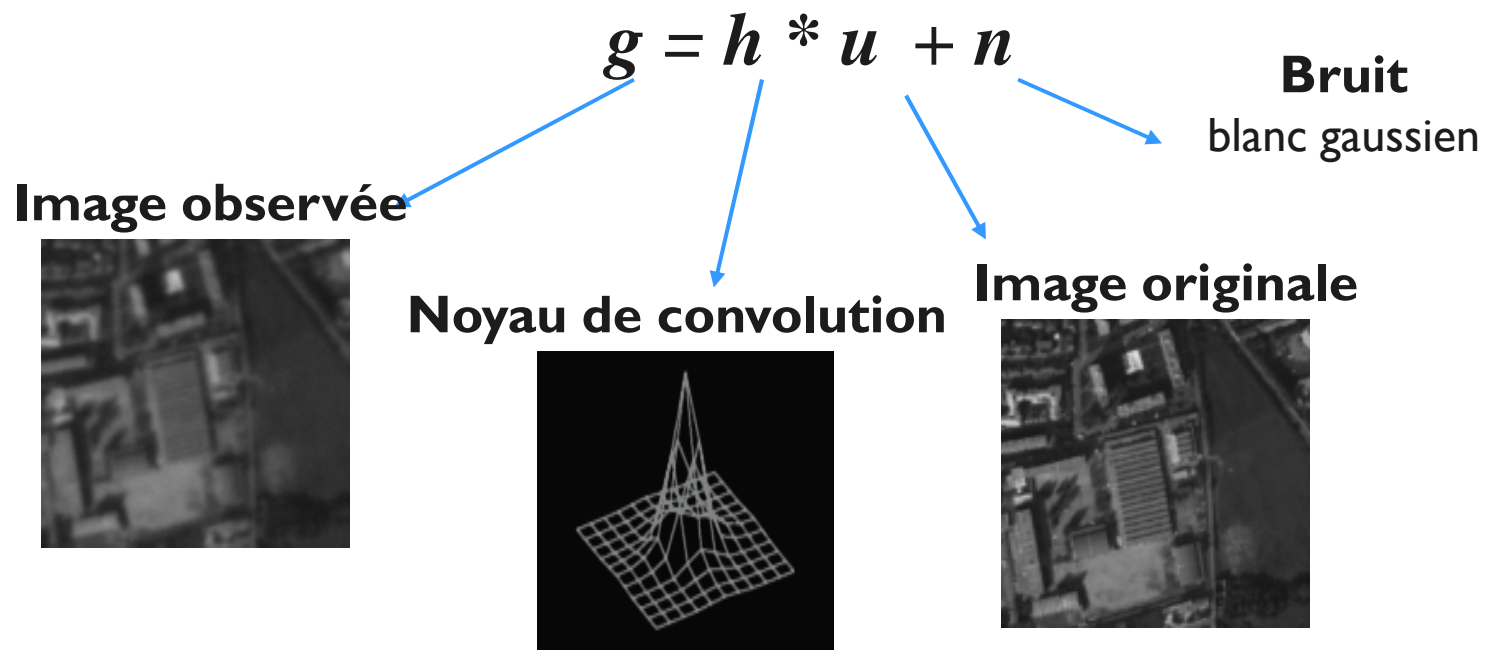
$$E(nn^t) = \begin{pmatrix} E(n_1n_1) & E(n_1n_2) & \cdot & \cdot & E(n_1n_N) \\ E(n_1n_2) & E(n_2n_2) & & & \\ & & \cdot & & \\ & & & \cdot & \\ E(n_1n_N) & & & & E(n_Nn_N) \end{pmatrix}$$

- ◆ Si v est une variable aléatoire scalaire, de densité de probabilité $p_v(v)$ et à valeurs dans \mathbb{R} . $E(v)$ désigne l'espérance mathématique de cette variable aléatoire, et est définie par $E(v) = \int_{\mathbb{R}} x p_v(x) dx$
- ◆ n est un bruit blanc si les variables aléatoires sont décorrélées d'un pixel à l'autre, c'est-à-dire que la matrice d'autocorrélation est diagonale. Si de plus le bruit est stationnaire, alors la matrice d'autocorrélation s'écrit $E(nn^t) = \sigma^2 \text{Id}$
- ◆ Si la loi du bruit est normale $\mathcal{N}(m, \Lambda)$, la densité de probabilité jointe des variables sur tous les pixels est

$$P_n(n) = \frac{1}{[2\pi \det(\Lambda)]^{\frac{1}{2}}} \exp - \frac{(n-m)^t \Lambda^{-1} (n-m)}{2}$$

Dégradation (modèle simple)

Les images observées sont **dégradées** :



Notations variables discrètes/continues

$\Omega \subset \mathbb{R}^2$ sous-ensemble ouvert borné
variables continues : $u(x)$

$$u : \Omega \rightarrow \mathbb{R}$$

$x \rightarrow u(x)$ niveau de gris au point $x=(x_1, x_2)$

$\Omega \subset \mathbb{N}^2$ sous-ensemble borné de points discrets
variables discrètes : pixel i, j

$$u_{i,j} = u(i\Delta x, j\Delta y), i, j = 0 \dots N$$

g : image observée, dégradée de u

Restauration

- ◆ Retrouver u , ayant observé $g = Hu + n$
 - Statistiques (moyenne et variance) connues pour le bruit
- ◆ Si l'opérateur H est connu : Restauration = déconvolution.
- ◆ si $H = id$: débruitage
- ◆ Si l'opérateur H est imparfaitement connu : Restauration = déconvolution aveugle ou myope

Restauration

- ◆ Restauration : débruitage, déconvolution.
- ◆ Inverser $g = h * u + n$ est un **problème inverse mal-posé**.
- ◆ Problème **bien posé** au sens d 'Hadamard (1923): la solution
 - existe
 - est unique
 - dépend continûment des données.
- ◆ Formulation matrice vecteur : $g = Hu + n$
- ◆ H peut être inversible, mais on a toujours instabilité de la solution par rapport aux données

Dans le plan de Fourier

- ◆ La convolution discrète circulaire est transformée en un produit simple dans Fourier (conditions de bords périodiques):

$$g = h * u + n \quad \rightarrow \quad \hat{g} = \hat{h} \cdot \hat{u} + \hat{n}$$

- ◆ écriture matrice/vecteur

$$g = \mathbf{H}u + n \quad \rightarrow \quad \hat{g} = \text{diag}\{\lambda_k\} \hat{u} + \hat{n}$$

- ◆ Remarque : en pratique on utilise plutôt des conditions aux bords symétriques, donc on utilise plutôt une DCT.
- ◆ Rappel : transformée discrète de $u_{i,j}$, $i, j \in \{0, \dots, N-1\}$

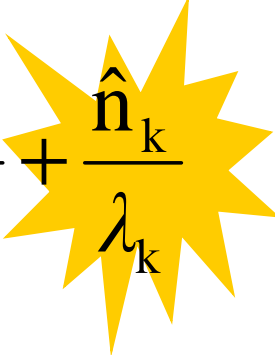
$$\hat{u}_{k,l} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} u_{i,j} e^{\frac{-2i\pi(ik+jl)}{N}} \quad \text{et} \quad u_{i,j} = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \hat{u}_{k,l} e^{\frac{2i\pi(ik+jl)}{N}}$$

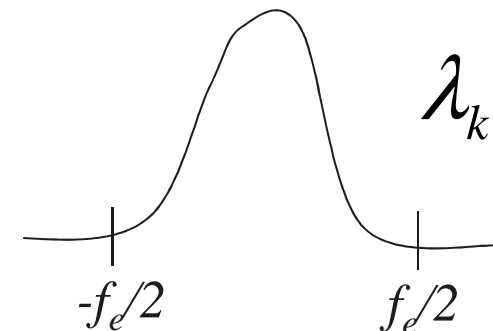
Analyse dans Fourier

- ◆ Les valeurs propres λ_k sont données par la transformée de Fourier de la PSF (fonction de flou) :
- ◆ Forme gaussienne, atténuation des hautes fréquences.

$$\hat{g} = \text{diag}\{\lambda_k\} \hat{u} + \hat{n}$$

- ◆ $\exists k, \lambda_k = 0$, le problème en u a une infinité de solutions
- ◆ $\forall k, \lambda_k \neq 0$, le problème en u a une unique solution, mais instable.

$$\hat{u}_k = \frac{\hat{g}_k}{\lambda_k} + \frac{\hat{n}_k}{\lambda_k}$$




Inversion

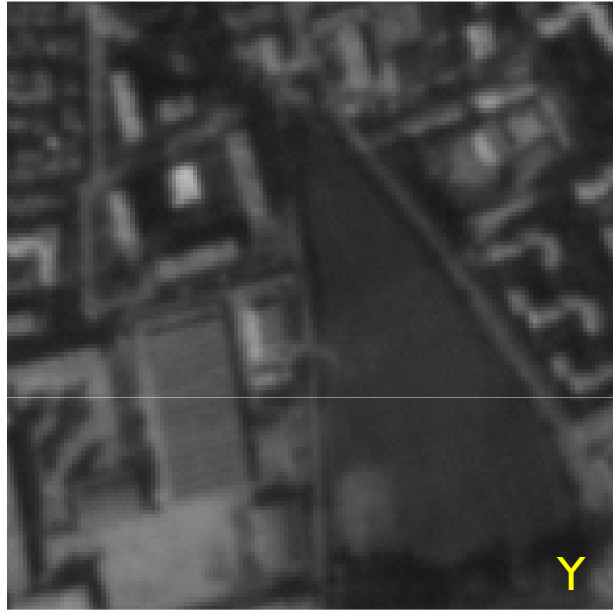


Image floue

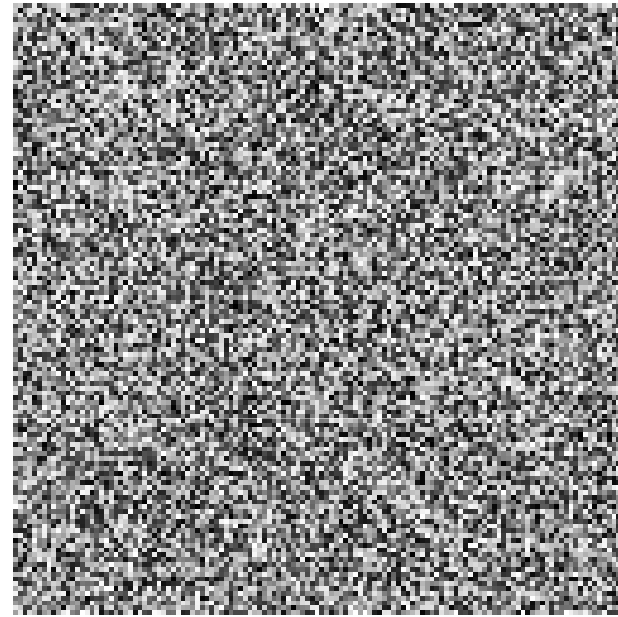
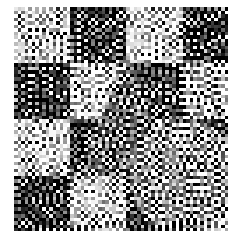
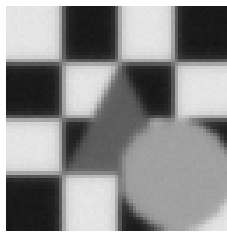
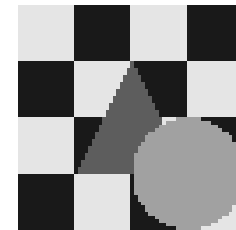


Image déconvoluée
sans régularisation



\neq



Conditionnement de matrice

- ◆ Si H est régulière, le conditionnement de H est : $\text{cond}(H) = \|H\| \cdot \|H^{-1}\|$

où la norme matricielle est $\|H\| = \sup_{x \neq 0} \left(\frac{\|Hx\|}{\|x\|} \right)$

- ◆ Avec la norme euclidienne et pour H régulière $\text{cond}(H) = \frac{\lambda_{\max}}{\lambda_{\min}}$

$$g = Hu \quad \text{donc} \quad \|g\| \leq \|H\| \cdot \|u\| \quad \text{d'où} \quad \frac{1}{\|u\|} \leq \|H\| \cdot \frac{1}{\|g\|}$$

petite perturbation $\delta g \rightarrow g + \delta g = H(u + \delta u)$

$$\delta g = H\delta u \quad \|\delta u\| \leq \|H^{-1}\| \cdot \|\delta g\| \quad \text{finalement} \quad \frac{\|\delta u\|}{\|u\|} \leq \|H\| \cdot \|H^{-1}\| \cdot \frac{\|\delta g\|}{\|g\|}$$

$$\text{soit} \quad \frac{\|\delta u\|}{\|u\|} \leq \text{cond}(H) \cdot \frac{\|\delta g\|}{\|g\|}$$

Moindres carrés et maximum de vraisemblance

- ◆ La solution des moindres carrés est celle qui résout le problème :

$$\min_u \|g - Hu\|^2$$

- ◆ Dans l'approche stochastique, cette solution est équivalente au **maximum de vraisemblance** dans le cas d'un bruit gaussien.

G , U et N sont considérées comme étant des champs de variables aléatoires. Le bruit N est blanc gaussien de loi $\mathcal{N}(0, \sigma^2 Id)$, donc

$$P_n(\mathbf{n}) = \frac{1}{2\pi\sigma} \exp\left(-\frac{\mathbf{n}^t \mathbf{n}}{2\sigma^2}\right) \quad \text{où} \quad \mathbf{n}^t \mathbf{n} = \|\mathbf{n}\|_2^2 = \sum_{i,j=1}^N n_{i,j}^2$$

La vraisemblance est définie par la probabilité conditionnelle des observations sachant u .

On cherche l'image u qui maximise la probabilité d'avoir observé g , soit

$$\max_u P(g/u)$$

$$\text{avec} \quad P(g/u) = P(\mathbf{n} = g - Hu) = \frac{1}{2\pi\sigma} \exp\left\{-\frac{\|g - Hu\|^2}{2\sigma^2}\right\}$$

Voir cours Jérôme Idier

Moindres carrés

- ◆ Minimum d'une fonction convexe. Condition premier ordre nécessaire et suffisante pour trouver le minimum

$$\frac{\partial}{\partial u} \|g - Hu\|^2 = 0$$

- ◆ Solution dans le domaine spatial $u_{est} = (H^*H)^{-1}H^*g$

- Calcul dérivée en 1D
$$\frac{\partial}{\partial u_i} \sum_l \left[g(l) - \sum_{k=-K}^K h(k)u(l-k) \right]^2$$

Terme non nul quand $i=l-k$ soit $l-i=k$.

$$\sum_l 2 h(l-i) \left[g(l) - \sum_{k=-K}^K h(k)u(l-k) \right] = 2 h^* * (g - h * u)$$

Moindres carrés

- ◆ Solution dans le domaine spatial

$$\mathbf{u}_{est} = (\mathbf{H}^* \mathbf{H})^{-1} \mathbf{H}^* \mathbf{g}$$

- ◆ Solution dans le domaine fréquentiel

$$\hat{\mathbf{u}}_{est} = \frac{FTM^*}{|FTM|^2} \hat{\mathbf{g}}$$

Méthodes itératives

- ◆ Descente de gradient

$$J(\mathbf{u}) = \|\mathbf{g} - \mathbf{H}\mathbf{u}\|^2, \quad \nabla J(\mathbf{u}) = \mathbf{H}^* (\mathbf{H}\mathbf{u} - \mathbf{g})$$

⇔ Schéma dynamique $\rightarrow \mathbf{u}(t) \quad \frac{\partial \mathbf{u}}{\partial t} = -\nabla J(\mathbf{u}) \approx \frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\delta t}$

- ◆ (Algorithme de Landweber)

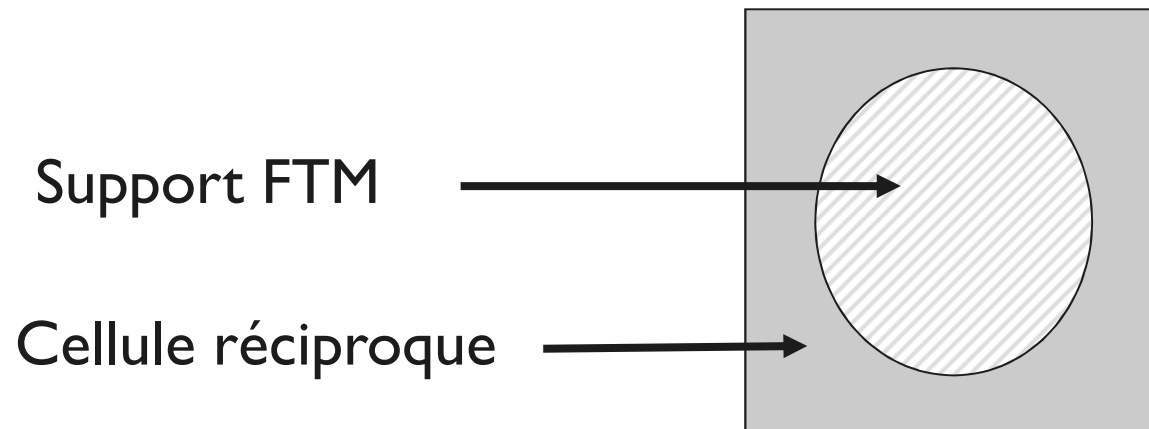
$$\mathbf{u}^{n+1} = \mathbf{u}^n + \alpha \mathbf{H}^* (\mathbf{g} - \mathbf{H}\mathbf{u}^n) \quad \text{avec } 0 < \alpha < 2$$

- ◆ Accélération par gradient conjugué
- ◆ Régularisation par arrêt des itérations

Méthodes avec contraintes

$$\mathbf{u}^{n+1} = P_C \left[\mathbf{u}^n + \alpha \mathbf{H}^* (\mathbf{g} - \mathbf{H}\mathbf{u}^n) \right]$$

- P_C : opérateur de **projection**
- Ex.: contrainte de **positivité** $C = \{ \mathbf{u} / u_{i,j} > 0 \}$
- On peut aussi introduire des contraintes dans le domaine **fréquentiel**. Projecteur dans le domaine fréquentiel par exemple si la cellule réciproque (liée à l'échantillonnage) est mal adaptée à la fonction instrument (FTM):



Filtre de Wiener

$$\min_u E_{g/u} \left[\|u_{vraie} - u\|^2 \right] \quad \text{Solution linéaire} \quad u = Lg$$

- Solution dans le domaine **spatial** (R_u matrice d'autocorrélation de u)

$$u_{est} = \left(H^* H + R_n R_u^{-1} \right)^{-1} H^* g$$

- Solution dans le domaine **fréquentiel**

$$\hat{u}_{est} = \frac{FTM^*}{|FTM|^2 + \frac{\sigma^2}{|\hat{r}_u|^2}} \hat{g}$$

Régularisation linéaire (Tikhonov)

- ◆ Chercher une solution dans un ensemble de solutions **admissibles** :

$$\left\{ \mathbf{u} / \frac{1}{N^2} \|\mathbf{g} - \mathbf{H}\mathbf{u}\|^2 \leq \sigma^2 \right\}$$

- ◆ Choisir une solution **régulière** pour stabiliser le processus d'inversion :

$$\min_{\mathbf{u}} \|\nabla \mathbf{u}\|^2 = \min_{\mathbf{u}} \sum_{i,j} |\nabla \mathbf{u}|_{i,j}^2$$

- ◆ Minimisation d'un critère (pénalisation) :

$$J(\mathbf{u}) = \|\mathbf{g} - \mathbf{H}\mathbf{u}\|^2 + \lambda \|\nabla \mathbf{u}\|^2$$

↙
Attache aux données

↘
Terme de régularisation

Régularisation linéaire(Tikhonov)

- ◆ Critère convexe, conditions du premier ordre :

$$H^*(Hu - g) - \lambda \Delta u = 0$$

- ◆ Descente de gradient

$$u^{n+1} = u^n + \delta_n (H^*(g - Hu) + \lambda \Delta u)$$



Résultat



Image originale @CNES



Image floue et bruitée
(@CNES, simulation SPOT5)



Une image restaurée

CONTENU global du cours

10 Introduction

- Formation des images en satellitaire, échantillonnage
- Déconvolution : ex de pb mal posé
- Régularisation linéaire

$$\|g - Hu\|_2^2 + \lambda \|\nabla u\|_2^2$$

◆ Minimisation de critère : approche variationnelle

- Espace fonctionnel pour la minimisation
- Existence, unicité de solution
- Introduction au calcul des variations,
- Introduction aux EDP en traitement d' image

$$\rightarrow \begin{cases} H^*(Hu - g) - \lambda \Delta u = 0 \\ \frac{\partial u}{\partial n} \Big|_{\partial\Omega} = 0 \end{cases}$$

$$\rightarrow \begin{cases} \frac{\partial u}{\partial t} = F(x, \nabla u, \nabla^2 u) \\ u(x, y, t = 0) = g(x, y) \end{cases}$$

Régularisation linéaire

- ◆ On va s'intéresser à la minimisation du critère

$$J(u) = \|g - Hu\|^2 + \|\nabla u\|^2$$

C'est-à-dire

$$J(u) = \int_{\Omega} [(g - Hu)(x)]^2 dx + \int_{\Omega} |\nabla u(x)|^2 dx$$

$\Omega \subset \mathbb{R}^2$ sous-ensemble ouvert borné

$u: \Omega \rightarrow \mathbb{R}$

$x \rightarrow u(x)$ niveau de gris au point $x = (x_1, x_2)$

Ce problème est-il bien posé du point vue mathématique?

Problème modèle $\inf_{u \in X} J(u)$

avec
$$J(u) = \int_{\Omega} [(g - Hu)(x)]^2 dx + \int_{\Omega} |\nabla u(x)|^2 dx$$

Questions :

- ◆ Existe-t-il un point minimum ?
- ◆ Est-il unique ?
- ◆ Quelle est la régularité du minimum ?
- ◆ Comment le calculer numériquement ?

Quelques définitions (rappel)

- ◆ Espaces L^p pour $1 \leq p < +\infty$

$$L^p(\Omega) = \left\{ u : \Omega \rightarrow \mathbb{R}, u \text{ mesurable t.q. } \left(\int_{\Omega} |u|^p dx \right)^{1/p} < +\infty \right\}$$

La norme est $\|u\|_{L^p(\Omega)} = \|u\|_p = \left(\int_{\Omega} |u(x)|^p dx \right)^{1/p}$

- ◆ L^∞ est défini par

$$L^\infty(\Omega) = \left\{ u : \Omega \rightarrow \mathbb{R}, u \text{ mesurable, } \exists C \text{ t.q. } |u(x)| \leq C \text{ pp sur } \Omega \right\}$$

et $\|u\|_{L^\infty(\Omega)} = \|u\|_\infty = \inf \{ C / |u(x)| \leq C \text{ pp } x \in \Omega \}$

- ◆ L^p est un espace de Banach (complet, normé) pour $1 \leq p \leq +\infty$

$L^p(\Omega)$ et dualité

- ◆ Si X est un espace de Banach, son **dual** X' est l'espace des formes linéaires continues sur X : $X' = \{ l : X \rightarrow \mathbb{R}, \text{ linéaire, continue } \}$

- ◆ La **norme duale** est $\|l\|_{X'} = \sup \{ l(x), x \in X, \|x\|_X \leq 1 \}$

On a $|l(u)| \leq \|l\|_{X'} \|u\|_X$

- ◆ Pour $1 < p < \infty$ $L^p(\Omega)' = L^q(\Omega)$ où $\frac{1}{p} + \frac{1}{q} = 1$

Pour $1 < p < \infty$, L^p est **réflexif** : $\left(L^p(\Omega)' \right)' = L^p(\Omega)$

- ◆ $L^1(\Omega)' = L^\infty(\Omega)$, mais $L^\infty(\Omega)' = ?$... il est strictement plus grand que $L^1(\Omega)$
- ◆ Cas particulier important : $p=2$ alors $L^2(\Omega)' = L^2(\Omega)$ et $L^2(\Omega)$ est un espace de Hilbert pour le produit scalaire

$$\langle f, g \rangle = \int_{\Omega} f(x)g(x)dx$$

Topologies

◆ Sur un espace de Banach X , on peut mettre **trois topologies**:

➤ La topologie **forte** sur X : $\|x_n - x\|_X \rightarrow 0$

➤ La topologie **faible** sur X : $l(x_n) \rightarrow l(x), \forall l \in X'$

➤ La topologie **faible*** sur X' : $l_n(x) \rightarrow l(x), \forall x \in X$

◆ Pourquoi toutes ces topologies?

➤ Si X est un espace réflexif et si une suite x_n est bornée, alors on peut extraire une sous suite convergente pour la topologie faible.

➤ Si une suite est bornée dans le dual X' , alors on peut en extraire une sous-suite convergente pour la topologie faible*.

Ces notions servent pour la démonstration de l'existence d'une solution au problème

$$\inf_{u \in X} F(u)$$

Convexité et semi-continuité inférieure

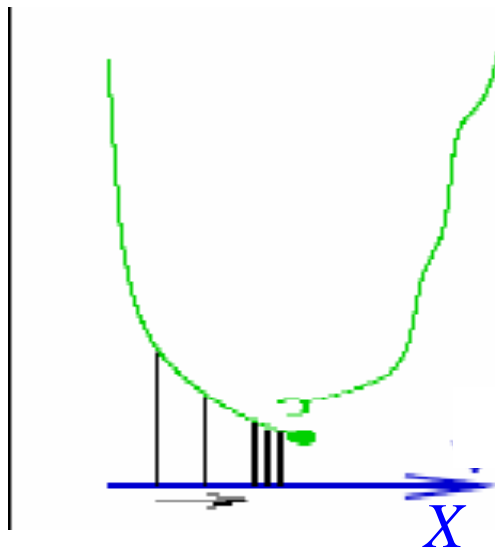
Soit X une espace de Banach et $F : X \rightarrow \mathbb{R}$.

◆ F est **convexe** si

$$F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y) \quad \forall x, y \in X \text{ et } \lambda \in [0,1]$$

◆ F est semi-continue inférieurement (**sci**) si

$$\liminf_{x_n \rightarrow x} F(x_n) \geq F(x)$$



Coercivité

◆ F est **coercive** si $\lim_{\|x\| \rightarrow +\infty} F(x) = +\infty$

◆ **Utilité de ces définitions ?**

Définir l'existence d'un minimum

➤ Si F est sci et coercive dans un espace réflexif alors F admet un minimum ...

Méthode directe du calcul des variations

On considère le problème de minimisation $\inf_{u \in X} F(u)$

- ◆ Considérer une suite minimisante $u_n \in X$, c'est-à-dire une suite satisfaisant

$$\lim_{n \rightarrow \infty} F(u_n) = \inf_{u \in X} F(u)$$

- ◆ Si F est coercive, on obtient une borne uniforme sur la suite u_n :

$$|u_n| \leq K$$

- ◆ Si X est réflexif alors on en déduit l'existence d'une sous suite u_{n_j} et $u_0 \in X$ tels que u_{n_j} converge faiblement vers u_0
- ◆ Si F est sci on en déduit

$$\inf_{u \in X} F(u) = \liminf_{n \rightarrow \infty} F(u_n) \geq F(u_0)$$

ce qui implique que

$$F(u_0) = \min_{u \in X} F(u)$$

Remarque

- ◆ Cas où F est une fonctionnelle intégrale

soit $f : \Omega \times \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}$ (avec $\Omega \subset \mathbb{R}^2$)

pour $u \in W^{1,p}(\Omega)$ on considère la fonctionnelle

$$F(u) = \int_{\Omega} f(x, u(x), Du(x)) dx$$

Si f est sci et coercive, F l'est aussi.

On en déduit l'existence d'un minimum dans $H^1 (=W^{1,2})$ pour le critère

$$J(u) = \int_{\Omega} [(g - Hu)(x)]^2 dx + \int_{\Omega} |\nabla u(x)|^2 dx$$

Gateaux différentiabilité

◆ **Définition :** dérivée de Gateaux

$$\lim_{\varepsilon \rightarrow 0^+} \frac{F(u + \varepsilon v) - F(u)}{\varepsilon}$$

Si cette limite existe pour tout $v \in X$ et s'il existe un élément de X' , que l'on notera $F'(u)$, tel que

$$\forall v \in X, \quad \lim_{\varepsilon \rightarrow 0^+} \frac{F(u + \varepsilon v) - F(u)}{\varepsilon} = \langle F'(u), v \rangle$$

alors F est dérivable au sens de Gateaux.

Equation d'Euler-Lagrange

- ◆ **Application** : si F est différentiable au sens de Gateaux et si le problème
- $$\inf_{u \in X} F(u)$$

a une solution u , alors nécessairement la solution vérifie :

$$F'(u) = 0$$

La réciproque est vraie si F est convexe.

Cette équation est l'équation d'**Euler-Lagrange**.

⇒ Calcul de l'équation d'Euler Lagrange pour $J(u)$.

Equation d'Euler-Lagrange

$$J(u) = \int_{\Omega} [(g - Hu)(x)]^2 dx + \int |\nabla u(x)|^2 dx$$

J est convexe, l'infimum existe et vérifie $J'(u) = 0$

Calcul de $J'(u)$

$$\lim_{\varepsilon \rightarrow 0^+} \frac{J(u + \varepsilon v) - J(u)}{\varepsilon} = \langle J'(u), v \rangle$$

Pour le deuxième terme de J , utiliser la formule de Green (intégration par partie) :

$$\int_{\Omega} \langle \nabla u, \nabla v \rangle dx = - \int_{\Omega} v \Delta u dx + \int_{\partial \Omega} v \frac{\partial u}{\partial n} ds$$

Equation d'Euler-Lagrange

$$J(u) = \int_{\Omega} [(g - Hu)(x)]^2 dx + \int_{\Omega} |\nabla u(x)|^2 dx$$

J est convexe, l'infimum existe et vérifie $J'(u) = 0$

soit

$$\begin{cases} H^*(Hu - g) - \lambda \Delta u = 0 \\ \left. \frac{\partial u}{\partial n} \right|_{\partial\Omega} = 0 \end{cases}$$

$$H : X \rightarrow X$$

$$u \rightarrow Hu = h^* u$$

$$H^* : X \rightarrow X$$

$$u \rightarrow H^* u = h^{**} u \quad \text{où } h^*(x) = \bar{h}(-x)$$

Unicité

- ◆ La fonctionnelle est convexe non strictement.
- ◆ Soient deux solutions u_1 et u_2 . Elles vérifient toutes les deux l'équation d'Euler Lagrange.
- ◆ En soustrayant les deux équations on obtient :

$$H^* H(u_1 - u_2) = \lambda \Delta(u_1 - u_2)$$

- ◆ En multipliant les deux membres par $(u_1 - u_2)$ et en intégrant par partie, on trouve

$$\int_{\Omega} \|H(u_1 - u_2)\|^2 = -\lambda \int_{\Omega} \|\nabla(u_1 - u_2)\|^2$$

$$\Rightarrow H(u_1 - u_2) = 0 \quad \text{et} \quad \nabla(u_1 - u_2) = 0$$

donc $u_1 = u_2 + cst$, et si H n'annule pas les constantes on en déduit $u_1 = u_2$

➔ si H n'annule pas les constantes, il y a unicité.

Equation dynamique (EDP)

- ◆ Equation d'Euler

$$\begin{cases} H^* (Hu - g) - \lambda \Delta u = 0 \\ \frac{\partial u}{\partial n} \Big|_{\partial\Omega} = 0 \end{cases}$$

- ◆ Schéma dynamique : $u(i, j, t)$

$$\longrightarrow \begin{cases} \frac{\partial u}{\partial t} = -\nabla^* J(g) + Hu + \lambda \Delta u \\ \frac{\partial u}{\partial n} \Big|_{\partial\Omega} = 0 \end{cases}$$

- ◆ Diffusion isotrope (Laplacien équation de la chaleur)

Résultat



Image originale @CNES



Image floue et bruitée
(@CNES, simulation SPOT5)



Une image restaurée

Comment introduire les EDP dans le traitement de l'image? 75

◆ Filtrage linéaire

Si u_0 une image est corrompue par du bruit, un moyen simple pour la lisser est de faire des moyennes locales :

$$u(x, y) = \frac{1}{5} (g(x-1, y) + g(x+1, y) + g(x, y-1) + g(x, y+1) + g(x, y))$$

ou plus généralement de faire une convolution discrète :

$$u(x, y) = \sum_{k,l} h(k, l) g(x-k, y-l)$$

Introduction des EDP

Et dans la formulation continue

$$u(x, y) = \int_{\mathbb{R}^2} h(x', y') g(x - x', y - y') dx' dy'$$

Pour que cette intégrale ait un sens il faut préalablement étendre l'image initiale à \mathbb{R}^2 tout entier :

- Par symétrie
- Par périodicité

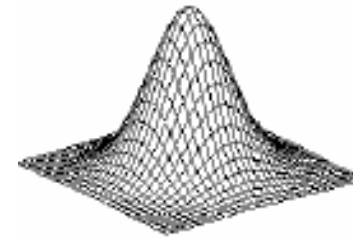
Exemple de noyau h : le noyau gaussien

$$u(x, y) = \int_{\mathbb{R}^2} G_t(x', y') g(x - x', y - y') dx' dy'$$

Introduction des EDP

où

$$G_t(x', y') = \frac{1}{4\pi t} e^{-\frac{(x'^2 + y'^2)}{4t}}$$



t = taille du filtre = paramètre d'échelle

Une remarque fondamentale (due à Konderink) : calculons
au point (x, y, t)

$$\frac{\partial u}{\partial t}, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \dots$$

Alors on trouve que ces dérivées vérifient l'EDP

$$\begin{cases} \frac{\partial u}{\partial t}(x, y, t) = \frac{\partial^2 u}{\partial x^2}(x, y, t) + \frac{\partial^2 u}{\partial y^2}(x, y, t) \\ u(x, y, t = 0) = g(x, y) \end{cases}$$

Introduction des EDP

C'est –à dire l'équation de la chaleur

$$\begin{cases} \frac{\partial u}{\partial t}(x, y, t) = \Delta u(x, y, t) \\ u(x, y, 0) = g(x, y) \end{cases}$$

Et donc filtrer une image par **convolution avec une gaussienne** est équivalent à **résoudre l'équation de la chaleur**. Cette remarque a été le point de départ de l'introduction des EDP en traitement d'images.

Mais l'équation de la chaleur est-elle une “bonne” EDP pour le traitement d'images ?

Propriétés de l'équation de la chaleur

- Si g est bornée alors il y a existence et unicité d'une solution vérifiant

$$\inf_{\mathbb{R}^2} g(x, y) \leq u(x, y, t) \leq \sup_{\mathbb{R}^2} g(x, y)$$

- Si on note T_t l'opérateur $u_0 \rightarrow u = T_t u_0$ alors :

$$(I1): T_t(0) = 0 \quad \text{et} \quad T_t(g + c) = T_t(g) + c$$

$$(I2): T_t(\tau_h g) = \tau_h(T_t(g))$$

$$\text{où } \tau_h(f)(x, y) = f(x + h_1, y + h_2), \quad h = (h_1, h_2)$$

$$(I3): T_t(Rg) = R(T_t(g))$$

$$\text{où } (Rf)(x, y) = f(R(x, y)), \quad \forall R \text{ une transformation}$$

orthogonale

Propriétés de l'équation de la chaleur

$$(I4): T_{t+s}(g) = T_t(T_s(g))$$

$$(I5): g \leq f \text{ alors } T_t(g) \leq T_t(f)$$

MAIS

$$\cdot u \in C^\infty(\mathbb{R}^2 \times (0, T)), \quad \forall T > 0$$

\Rightarrow lissage instantané

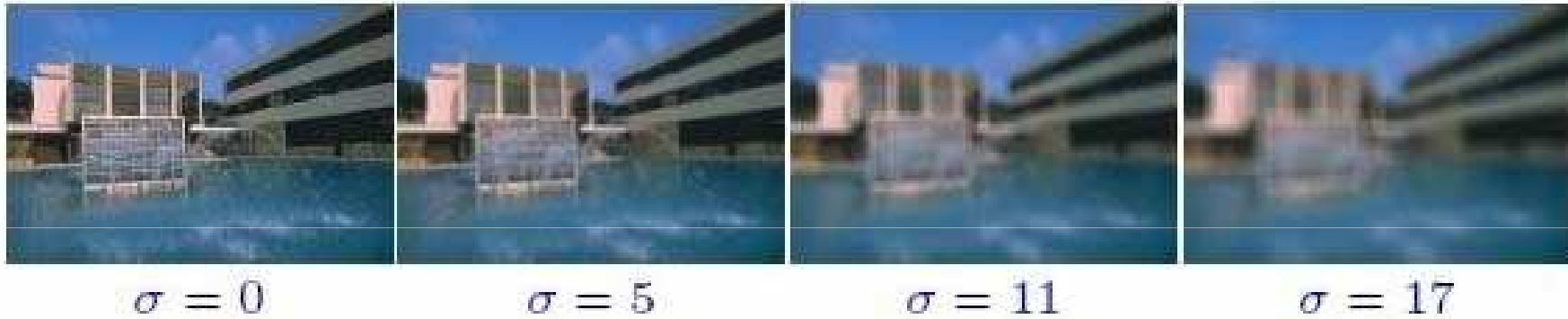
\Rightarrow perte des contours

- Si u est une solution, alors $h(u)$ avec h strictement croissante, n'est pas une solution.

\Rightarrow pas d'invariance morphologique

Exemple avec l'équation de la chaleur

81



Introduction de l'effet de flou dans le processus

⇒ vers un modèle non linéaire.

Filtrage non linéaire : le modèle de Malik et Perona

◆ Quelques notations

$$u : (x, y) \rightarrow u(x, y) \in \mathbb{R}, \quad u_x = \frac{\partial u}{\partial x}(x, y), \quad u_y = \frac{\partial u}{\partial y}(x, y),$$

$$u_{xx} = \dots$$

$$\nabla u(X) = \begin{pmatrix} u_x \\ u_y \end{pmatrix} \text{ vecteur gradient}$$

$$\nabla^2 u(X) = \begin{pmatrix} u_{xx} & u_{xy} \\ u_{xy} & u_{yy} \end{pmatrix} \text{ matrice hessienne}$$

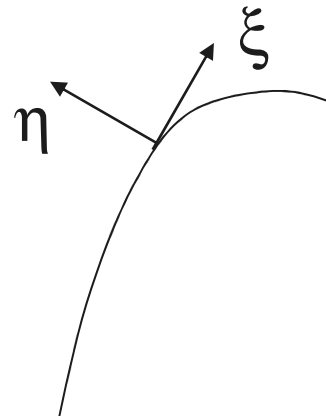
Notations (suite)

$C = \{ (x, y) \in \mathbb{R}^2 ; u(x, y) = c \}$ courbe de niveau de u

$\xi = \frac{1}{|\nabla u|} \begin{pmatrix} -u_y \\ u_x \end{pmatrix}$: vecteur tangent unitaire à C

$\eta = \frac{1}{|\nabla u|} \begin{pmatrix} u_x \\ u_y \end{pmatrix} = \frac{\nabla u}{|\nabla u|}$: vecteur normal unitaire à C

$$\xi \perp \eta$$



Notations (suite)

$$\begin{cases} u_{\eta\eta} = \eta^t \nabla^2 u \eta = \frac{1}{|\nabla u|^2} (u_x^2 u_{xx} + u_y^2 u_{yy} + 2u_x u_y u_{xy}) \\ u_{\xi\xi} = \xi^t \nabla^2 u \xi = \frac{1}{|\nabla u|^2} (u_x^2 u_{yy} + u_y^2 u_{xx} - 2u_x u_y u_{xy}) \end{cases}$$

Dérivée seconde dans la direction orthogonale η et tangentielle ξ .

Remarque : $u_{\eta\eta} + u_{\xi\xi} = \Delta u$

En fait on a pour toutes directions V et W orthogonales et unitaires $u_{VV} + u_{WW} = \Delta u$

Conséquences

L'équation de la chaleur est isotrope : aucune direction n'est privilégiée

⇒ Vers un modèle anisotrope

On a :
$$\frac{\partial u}{\partial t} = \Delta u = u_{xx} + u_{yy} = \operatorname{div}(\nabla u)$$

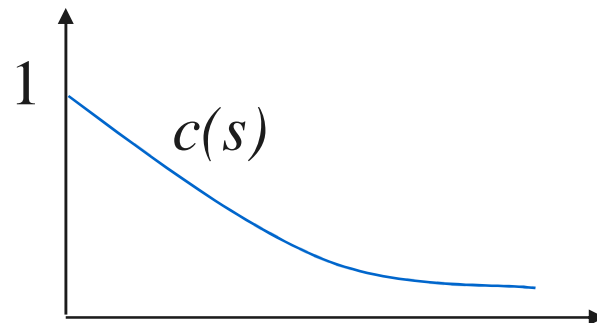
Un modèle plus général : on introduit un coefficient de diffusion $c(s)$

Diffusion non linéaire : le modèle de Malik et Perona

$$(PM) \begin{cases} \frac{\partial u}{\partial t} = \operatorname{div} \left(c(|\nabla u|^2) \nabla u \right) \\ u(x, y, t = 0) = g(x, y) \end{cases}$$

où $c(s)$ vérifie les hypothèses :

$$c(0) = 1, \quad c(+\infty) = 0, \quad c(s) > 0, \quad c(s) \text{ décroissante}$$



Propriétés formelles de (PM)

- ◆ Si $|\nabla u(x, y)| \cong 0$ i.e. à l'intérieur de zones homogènes :

$$c(|\nabla u(x, y)|) \cong 1, \quad \text{et} \quad \frac{\partial u}{\partial t} = \Delta u$$

⇒ lissage isotrope

- ◆ Si $|\nabla u(x, y)| \cong \infty$, i.e. au voisinage des contours :

$$c(|\nabla u(x, y)|) \cong 0, \quad \text{et} \quad \frac{\partial u}{\partial t} = 0$$

⇒ on ne fait rien : les contours sont préservés

Quel est le « meilleur » choix pour $c(s)$

- ◆ D'un point de vue analyse d'images : un développement formel

$$\operatorname{div}\left(c(|\nabla u|^2)\nabla u\right)=c(|\nabla u|^2)u_{\xi\xi}+\left(c(|\nabla u|^2)+2|\nabla u|^2c'(|\nabla u|^2)\right)u_{\eta\eta}$$

donc le modèle correspond à un lissage dans la direction ξ plus un lissage dans la direction η . Un bon choix serait :

- $c(0) = 1$ correspond à un lissage isotrope dans les zones homogènes
- Aucune diffusion dans la direction normale η :

$$c(s^2)+2s^2c'(s^2)\cong 0 \text{ lorsque } s \rightarrow +\infty$$

Choix de $c(s)$

C'est-à-dire $c(s) \approx \frac{a}{\sqrt{s}}$ ($s \rightarrow +\infty$)

et dans ce cas

$$(\text{PM}) \approx \frac{\partial u}{\partial t} = \operatorname{div} \left(a \frac{\nabla u}{|\nabla u|} \right) = a \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) = a \operatorname{curv}(u)(x, y)$$

$$\text{où } \operatorname{curv}(u)(x, y) = \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) = \frac{1}{|\nabla u|} u_{\xi\xi}$$

est la **courbure** de la ligne de niveau de u passant par (x, y) .

Choix de $c(s)$ (résumé)

- ◆ Du point de vue de l'analyse d'image, il serait souhaitable que

$$c(s) > 0, \quad c(0) = 1, \quad c(s) \approx \frac{1}{\sqrt{s}} \quad (s \rightarrow +\infty), \quad c \text{ décroissante}$$

- ◆ Et d'un point de vue mathématique ? Le problème principal concerne l'existence et l'unicité d'une solution.

L'équation de (PM) peut encore s'écrire sous la forme :

$$\frac{\partial u}{\partial t} = a_{11}(|\nabla u|^2)u_{xx} + 2a_{12}(|\nabla u|^2)u_{xy} + a_{22}(|\nabla u|^2)u_{yy}$$

Une bonne théorie pour aborder ce type d'équation est celle des EDP paraboliques, c'est-à-dire que les coefficients $a_{ij}(s)$ vérifient l'égalité

$$\sum_{i,j=1,2} a_{ij}(s)v_i v_j \geq 0, \quad \forall v \in \mathbb{R}^2, \quad \forall s > 0$$

Choix de $c(s)$ (point de vue mathématique)

ce qui est équivalent à

$$c(s) + 2c'(s) > 0, \quad \forall s > 0$$

- ◆ En résumé on veut :
$$\begin{cases} c : [0, +\infty[\rightarrow]0, +\infty[, & c \text{ décroissante} \\ c(0) = 1, & c(s) \approx \frac{1}{\sqrt{s}} \text{ quand } s \rightarrow +\infty \\ c(s) + 2sc'(s) > 0 \end{cases}$$

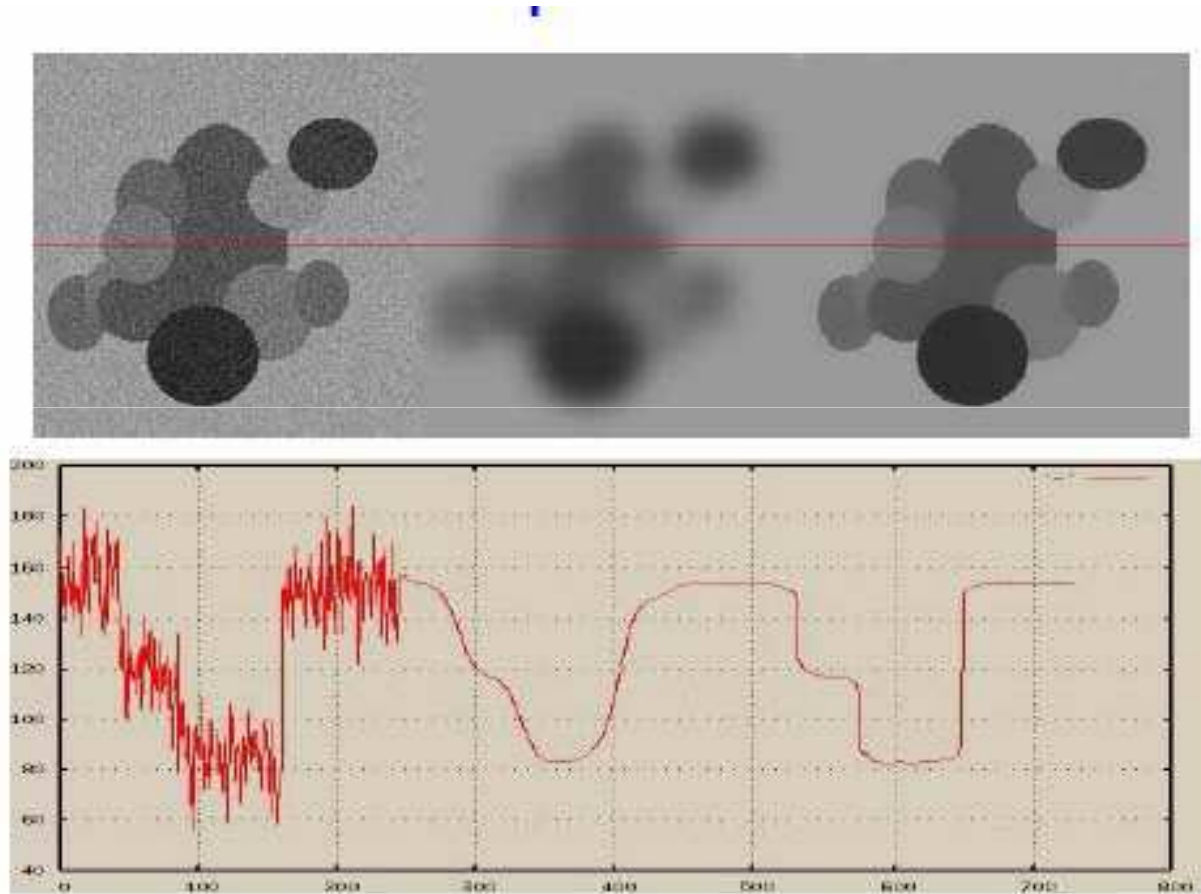
- ◆ Un choix qui convient est $c(s) = \frac{1}{\sqrt{1+s}}$

Modèle de Perona Malik (fin)

- ◆ On peut montrer en utilisant la théorie des opérateurs monotones l'existence et l'unicité d'une solution

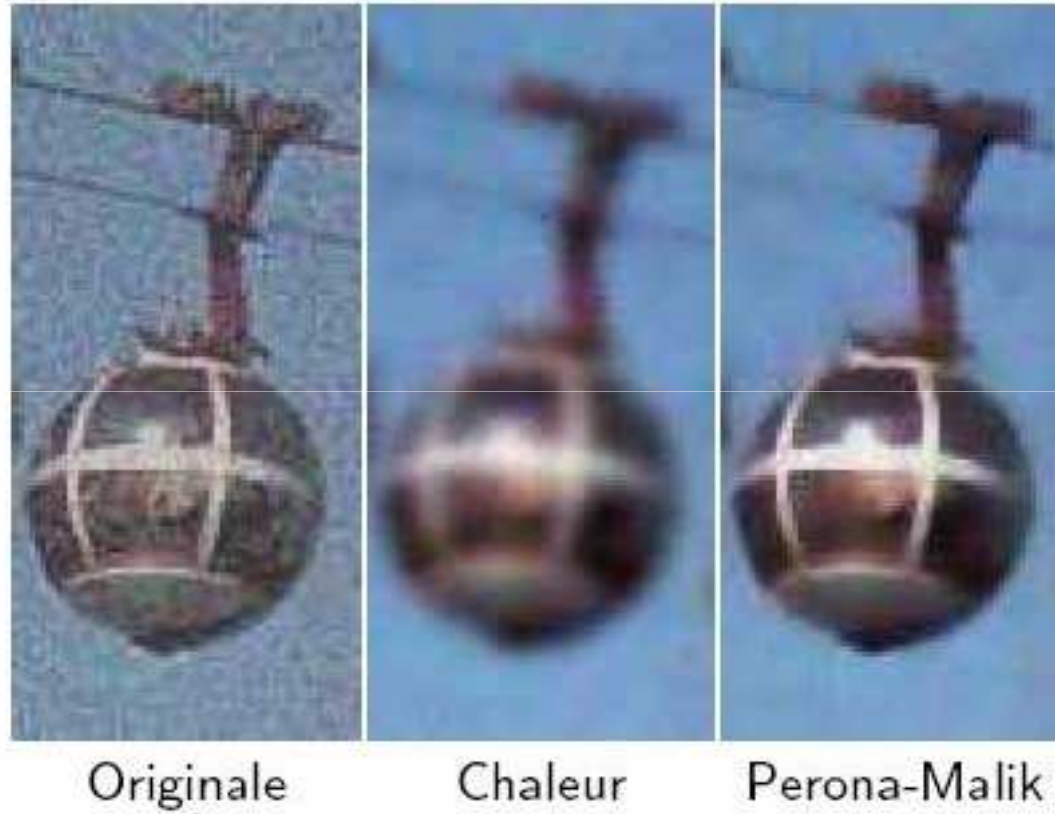
[H Brezis, “Opérateurs maximaux monotones et semi-groupes de contraction dans les espaces de Hilbert” North-Holland Publishing Compagny, 1973]

Quelques résultats expérimentaux



Originale bruitée / chaleur / Perona-Malik

Résultats



Minimisation de critère

$$\min_{u \in X \cap X_J} \|g - Hu\|_q^q + J(u)$$

Terme d'attache
aux données

Norme ou semi-norme
dans un espace
régularisant

- Quelles sont les normes, les bons espaces régularisant adaptés aux images ?
- Quelle norme choisir pour le terme de données ?
- existence, unicité d'une solution ?
- Algorithme de minimisation ?
- Evaluation des résultats ?

CONTENU (suite)

◆ Régularisation non linéaire

➤ Régularisation L^2/L^1

$$\|g - Hu\|_2^2 + \lambda \int \varphi(|\nabla u|) dx$$

➤ Espace BV, Variation Totale

$$\|g - Hu\|_2^2 + \lambda \int_{\Omega} |Du|$$

➤ Solutions dans l'espace BV, dualité, algorithmes

◆ Segmentation d'image, approche variationnelle

◆ Fonctionnelle de Mumford et Shah

◆ Notion de Γ -convergence : lien restauration/segmentation

◆ Et les représentations parcimonieuses ? $\|g - Hu\|_2^2 + \lambda \sum_{j,k} 2^{-\frac{j}{2}} |\langle u, \psi_{j,k} \rangle|$

➤ Régularisation dans le domaine de la transformée en ondelettes

➤ Solution dans un espace de Besov (et algorithmes)

➤ Modèles de régularisation / parcimonie

Norme l^1 et l^2

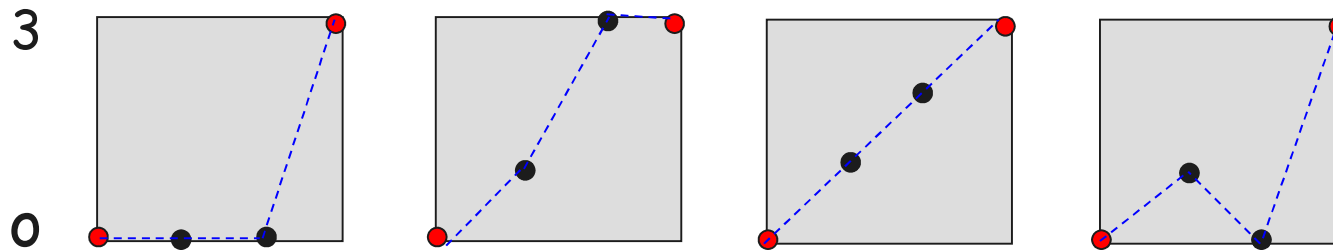
◆ Diminuer le poids des forts gradients dans le processus de minimisation. En variables discrètes :

- remplacer la norme l^2 par la norme l^1 .

$$E(u) = \|g - Hu\|_2^2 + \lambda \|\nabla u\|_1$$

$$\|\nabla u\|_1 = \sum_{i,j} |\nabla u|_{i,j} \quad \text{avec} \quad |\nabla u|_{i,j} = \sqrt{(u_{i,j} - u_{i-1,j})^2 + (u_{i,j} - u_{i,j-1})^2}$$

◆ Normes l^2 et l^1 : un petit exemple en 1D. $\|\nabla u\|_1 = \sum_i |u_i - u_{i-1}|$



$\ \cdot\ _2^2=9$	$\ \cdot\ _2^2=5$	$\ \cdot\ _2^2=3$	$\ \cdot\ _2^2=11$	$\ \nabla u\ _2^2 = \sum_i u_i - u_{i-1} ^2$ $\ \nabla u\ _1 = \sum_i u_i - u_{i-1} $
$\ \cdot\ _1=3$	$\ \cdot\ _1=3$	$\ \cdot\ _1=3$	$\ \cdot\ _1=5$	

Minimisation dans L^1

$$E(u) = \|g - Hu\|_2^2 + \lambda \|\nabla u\|_1$$

$$\|\nabla u\|_1 = \int_{\Omega} |\nabla u| dx$$

◆ Equation d'Euler $\frac{\partial u}{\partial t} = H^*(g - Hu) + \lambda \operatorname{div}\left(\frac{1}{|\nabla u|} \nabla u\right)$

◆ Rappel dans L^2 $E(u) = \|g - Hu\|_2^2 + \lambda \|\nabla u\|_2^2$

$$\frac{\partial u}{\partial t} = H^*(g - Hu) + \lambda \Delta u = H^*(g - Hu) + \lambda \operatorname{div}(\nabla u)$$

Conditions aux bords de Neuman $\frac{\partial u}{\partial \vec{n}} = 0$ sur $\partial\Omega$

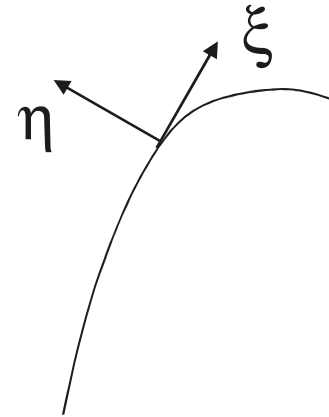
Régularisation L^2

◆ Terme de régularisation l^2 : $J_{reg}(u) = \int_{\Omega} |\nabla u|^2 dx$ $\nabla J_{reg}(u) = -div(\nabla u) = -\Delta u$

Lissage isotrope $\frac{\partial u}{\partial t} = -\nabla J_{reg}(u) = (u_{xx} + u_{yy}) = (u_{\xi\xi} + u_{\eta\eta})$

ξ et η vecteurs unitaires orthogonaux, ξ est dans la direction tangentielle à la ligne de niveau, η est orthogonal à la ligne de niveau, dans la direction du gradient

$$\eta = \frac{\nabla u}{|\nabla u|} \quad \xi \perp \eta$$



$$\begin{cases} u_{\eta\eta} = {}^t \eta \nabla^2 u \eta = \frac{1}{|\nabla u|^2} (u_{x_1}^2 u_{x_1 x_1} + u_{x_2}^2 u_{x_2 x_2} + 2u_{x_1} u_{x_2} u_{x_1 x_2}) \\ u_{\xi\xi} = {}^t \xi \nabla^2 u \xi = \frac{1}{|\nabla u|^2} (u_{x_1}^2 u_{x_2 x_2} + u_{x_2}^2 u_{x_1 x_1} - 2u_{x_1} u_{x_2} u_{x_1 x_2}) \end{cases}$$

Régularisation L^1

◆ terme de régularisation l^1 : $J_{reg}(u) = \int_{\Omega} |\nabla u| dx$ $\nabla J_{reg}(u) = -div\left(\frac{\nabla u}{|\nabla u|}\right)$

où $div(\vec{v}) = \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2}$ Remarquons que $div(\nabla u) = \Delta u$

$$div\left(\frac{\nabla u}{|\nabla u|}\right) = \frac{\partial}{\partial x_1} \left(\frac{u_{x_1}}{\sqrt{u_{x_1}^2 + u_{x_2}^2}} \right) + \frac{\partial}{\partial x_2} \left(\frac{u_{x_2}}{\sqrt{u_{x_1}^2 + u_{x_2}^2}} \right)$$

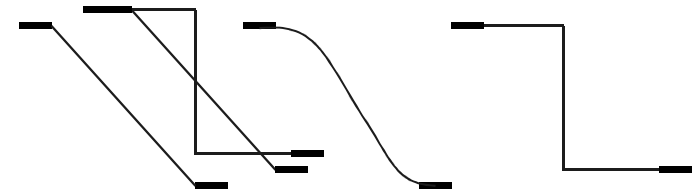
$$div\left(\frac{\nabla u}{|\nabla u|}\right) = \frac{\Delta u}{|\nabla u|} - \frac{1}{|\nabla u|} \cdot u_{\eta\eta} = \frac{1}{|\nabla u|} \cdot u_{\xi\xi}$$

Le lissage est anisotrope $\frac{\partial u}{\partial t} = -\nabla J_{reg}(u) = \frac{1}{|\nabla u|} u_{\xi\xi} = curv(u)$

- ◆ Le lissage est uniquement dans la direction tangentielle à la courbe de niveau, il n'y a donc pas de diffusion dans la direction du gradient.

Régulariser **et** préserver les contours

$$\text{Min}_u \sum_{i,j} |(\nabla u)_{i,j}|^\alpha \quad \alpha \approx 1 \rightarrow$$



$$(0 < \alpha)$$

⑩ On cherche u qui minimise $J(u)$:

ϕ fonction de régularisation

$$J(u) = \|g - Hu\|^2 + \lambda \sum_{i,j} \phi(|\nabla u|_{i,j})$$

■ Zones homogènes \longrightarrow régularisation

$$(\phi \approx t^\alpha \quad 1 < \alpha)$$

■ Contours \longrightarrow préservation

$$(\phi \approx t^\alpha \quad 1 \geq \alpha)$$

Classe des fonctions de régularisation φ

- ◆ Equation d'Euler

$$\frac{\partial u}{\partial t} = H^*(g - Hu) + \lambda \operatorname{div} \left(\frac{\varphi'(|\nabla u|)}{|\nabla u|} \nabla u \right)$$

Coefficient de pondération

- Lissage des zones homogènes :

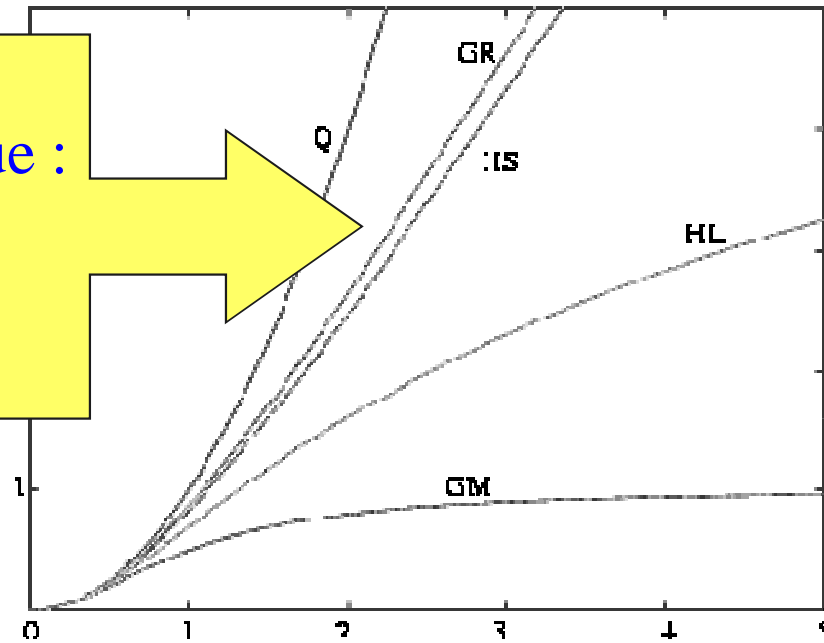
$$|\nabla u| \approx 0 \longrightarrow \frac{\varphi'(|\nabla u|)}{|\nabla u|} \approx 1$$

- Préservation des contours :

$$|\nabla u| \approx \infty \longrightarrow \frac{\varphi'(|\nabla u|)}{|\nabla u|} \approx 0$$

Fonctions $\varphi : l^2/l^1$

φ non quadratique :
*préservation
des contours*



Q	Quadratique	u^2
GM	Geman & McClure	$u^2 / (1 + u^2)$
HS	Hyper Surfaces	$2\sqrt{1 + u^2} - 2$
HL	Hebert & Leahy	$\log(1 + u^2)$
GR	Green	$2 \log(\cosh(u))$

Modèle explicite de contours

- ◆ Développements semi-quadratiques, sous certaines conditions en φ (principalement $\varphi(\sqrt{t})$ concave)

$$\varphi(t) = \inf_b \Psi(t, b)$$

■ Ψ quadratique en t

■ Ψ convexe en b ,

expression analytique du minimum

- ◆ Développement de $\varphi(|\nabla u|_{i,j})$ en tout i,j
- ◆ variable auxiliaire $b_{i,j}$: contours
- ◆ Minimisation de $J^*(u,b)$

$$J^*(u, b) = \|g - Hu\|^2 + \lambda^2 \sum_{i,j} \Psi(|\nabla u|_{i,j}, b_{i,j})$$

- [D. Geman & G.Reynolds « Constrained restoration and the recovery of discontinuities » *IEEE Trans. on PAMI*, 14(3), p. 367-383, 1992].
- [D. Geman & C.Yang « Nonlinear image recovery with half-quadratic regularization » *IEEE Trans. on Image Processing*, 4(7), p.932-946, 1995].

Algorithme semi-quadratique

◆ Minimisations **alternées** en u et b



■ u fixé \longrightarrow expression de b

■ b fixé \longrightarrow minimum en u quadratique (gradient conjugué).

◆ Initialisation $u^0=0$ ou $u^0=g$

◆ convergence (en fonction de la convexité de φ)

• [P.Charbonnier, L. Blanc-Féraud, G.Aubert, M. Barlaud, « Deterministic Edge-Preserving Regularization in Computed Imaging » *IEEE Trans. On Image Processing*, 6(2), 1997].

• [G.Aubert, L.Vese « A variational method in image recovery » *SIAM Journal of Numerical analysis*, 35(4), 1997].

• [A. Delaney, Y. Bresler « Globally convergent edge-preserving regularized reconstruction :An application to limited-angle tomography » *IEEE Trans. On Image Processing*, 7(2), 1998].

Algorithme semi-quadratique

◆ Développement multiplicatif $\varphi(t) = \inf_{b \in [0,1]} \{bt^2 + \psi(b)\}$ avec $b_{\text{inf}} = \frac{\varphi'(t)}{t}$

soit $\Psi(t, b) = bt^2 + \psi(b)$

◆ exemples

■ fonction hypersurface

$$\varphi(t) = \sqrt{1+t^2} - 1 \quad \psi(b) = b + \frac{1}{2b} - \frac{5}{2} \quad \frac{\varphi'(t)}{t} = \frac{2}{\sqrt{1+t^2}}$$

■ fonction Geman & McClure

$$\varphi(t) = \frac{t^2}{1+t^2} \quad \psi(b) = (\sqrt{b} - 1)^2 \quad \frac{\varphi'(t)}{t} = \frac{2}{(1+t^2)^2}$$

Algorithme semi-quadratique

◆ Avec
$$\varphi(t) = \inf_{b \in [0,1]} \{bt^2 + \psi(b)\} \quad , \quad b_{\text{inf}} = \frac{\varphi'(t)}{t}$$

soit
$$\Psi(t, b) = bt^2 + \psi(b)$$

on minimise le critère augmenté :

$$J^*(u, b) = \|g - Hu\|^2 + \lambda^2 \sum_{i,j} b_{i,j} |\nabla u|_{i,j}^2 + \psi(b_{i,j})$$

◆ Minimisations **alternées** en u et b

■ u fixé $\rightarrow b_{i,j} = \frac{\varphi'(|\nabla u|_{i,j})}{|\nabla u|_{i,j}}$

■ b fixé \rightarrow minimum en u quadratique (gradient conjugué).

◆ Schéma semi-implicite de résolution de l'EDP

$$u^{n+1} = u^n + \delta t \left\{ H^* (g - Hu^{n+1}) + \lambda^2 \operatorname{div} \left(\frac{\varphi'(|\nabla u^n|)}{|\nabla u^n|} \nabla u^{n+1} \right) \right\}$$

Modèle de contours

◆ Minimisation

- $b_{i,j} = 1 \rightarrow$ zone homogène : régularisation en $|\nabla u|_{i,j}^2$
- $b_{i,j} = 0 \rightarrow$ contour : pas de régularisation

$$J^*(u, b) = \|g - Hu\|^2 + \lambda^2 \sum_{i,j} b_{i,j} |\nabla u|_{i,j}^2 + \psi(b_{i,j})$$

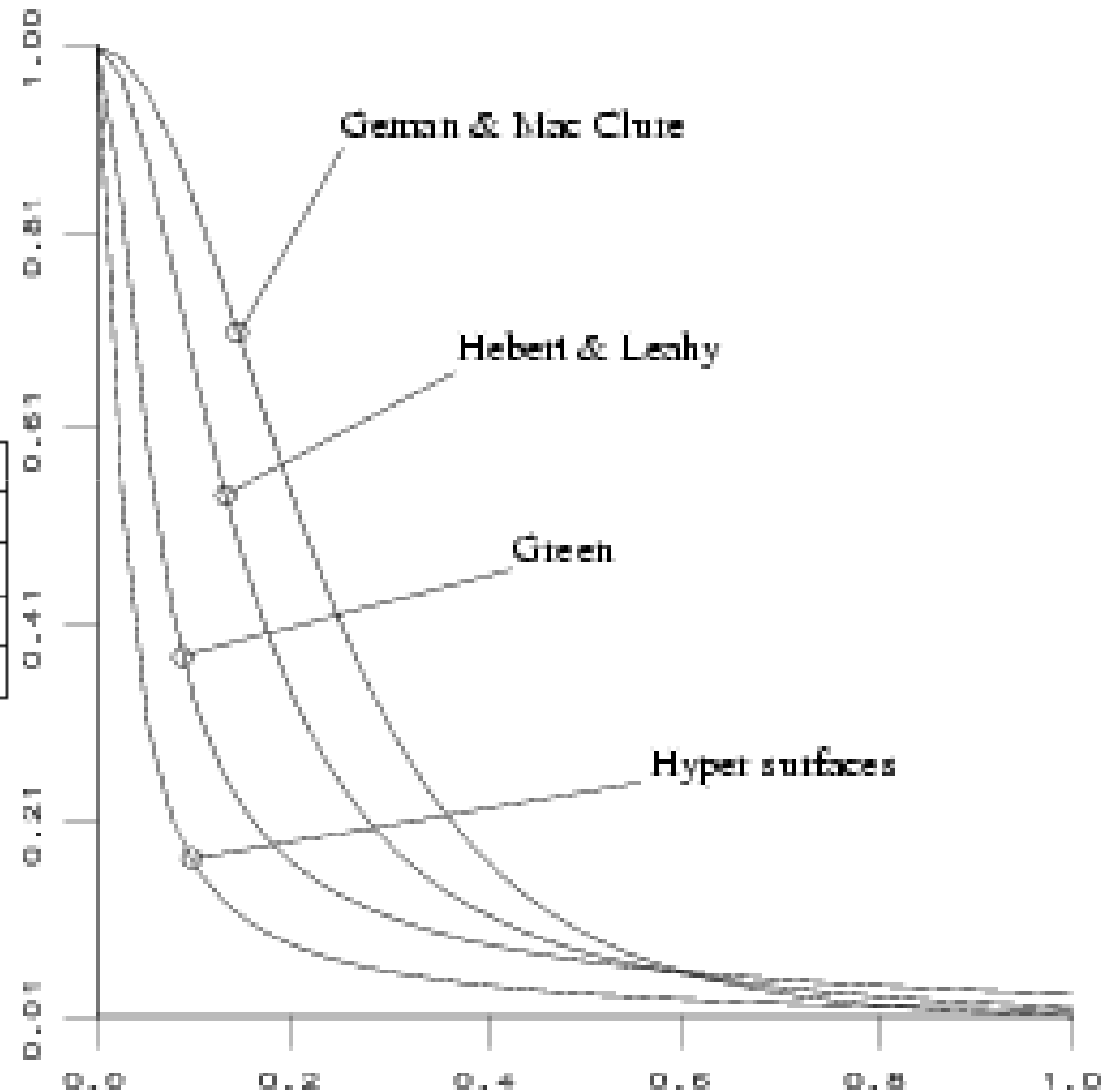
◆ Diffusion

- $b_{i,j} = 1 \rightarrow$ zone homogène : diffusion isotrope en Δu
- $b_{i,j} = 0 \rightarrow$ contour : pas de diffusion

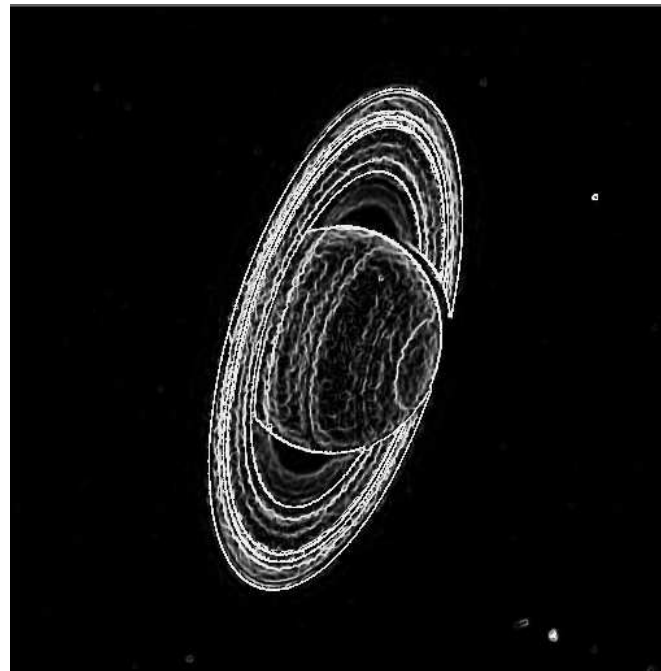
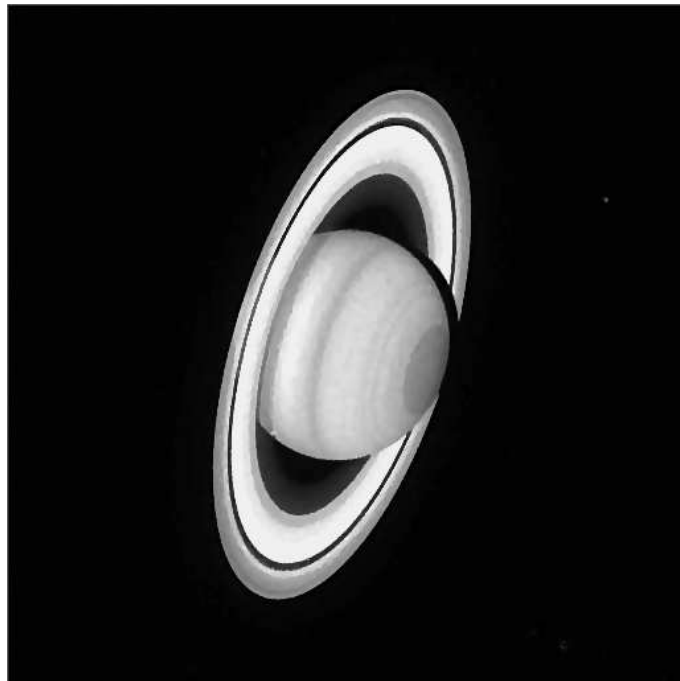
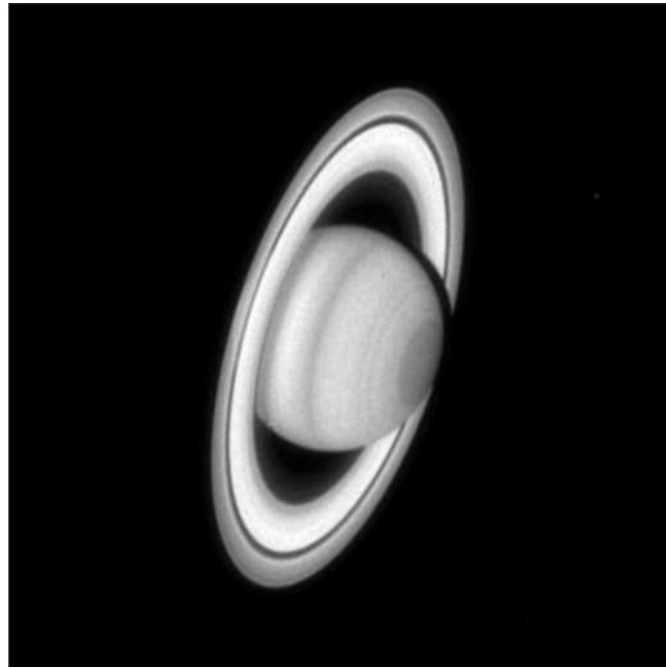
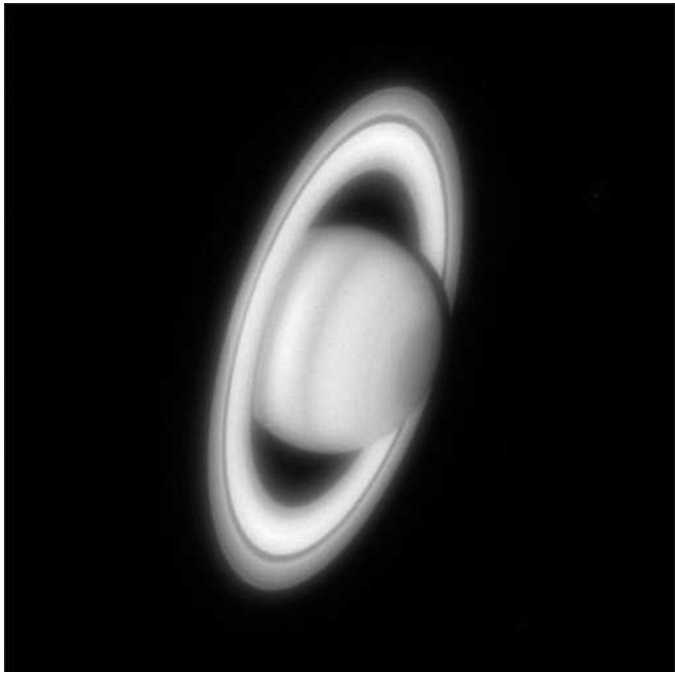
$$\frac{\partial u}{\partial t} = H^*(g - Hu) + \lambda^2 \operatorname{div}(b \nabla u)$$

Fonctions de pondérations b

Q	Quadratique	u^2
GM	Geman & McClure	$u^2 / (1 + u^2)$
HS	Hyper Surfaces	$2\sqrt{1 + u^2} - 2$
HL	Hebert & Leahy	$\log(1 + u^2)$
GR	Green	$2 \log(\cosh(u))$



Saturn (Hubble telescope)



observed	Wiener
edge preserving	b



Nîmes, image originale 512 x 512 (@CNES)



Nîmes, image dégradée: flou + bruit gaussien $\sigma \sim 1.4$



Nîmes, image restaurée, $\lambda=0.5$, $\delta=10$



Image originale



Régularisation L^2



Image observée g



Régularisation semi-quadratique
 $\lambda=0.5, \delta=10$

Interprétation stochastique

Cadre discret : $u(x) \longleftrightarrow u_{i,j}$

$g, u =$ champ de $N \times N$ variables aléatoires (pixels)

- Étude des probabilités
- Estimateurs statistiques
- Contraintes : introduites sur la loi $P(u)$

Interprétation stochastique : attache aux données

- ◆ **Estimateur du maximum de vraisemblance :**

$$\underset{\mathbf{u}}{\text{Max}} P(\mathbf{g} | \mathbf{u})$$

- ◆ **Loi du bruit**

cas additif blanc gaussien indépendant des observations

$$\begin{array}{l} \mathbf{g} = \mathbf{H}\mathbf{u} + \mathbf{n} \\ \mathbf{n} \sim \mathbf{N}(0, \sigma^2 \mathbf{I}) \end{array} \longrightarrow \begin{array}{l} P(\mathbf{g} | \mathbf{u}) = P(\mathbf{g} - \mathbf{H}\mathbf{u} = \mathbf{n}) \\ P(\mathbf{g} | \mathbf{u}) \propto \exp(-\|\mathbf{g} - \mathbf{H}\mathbf{u}\|^2 / 2\sigma^2) \end{array}$$

$$\underset{\mathbf{u}}{\text{Max}} P(\mathbf{g} | \mathbf{u}) \longleftrightarrow \underset{\mathbf{u}}{\text{Min}} \|\mathbf{g} - \mathbf{H}\mathbf{u}\|^2 / 2\sigma^2$$

H et σ supposés connus

Interprétation stochastique : terme de régularisation

◆ **Choix de la probabilité $P(\mathbf{u})$**

Contraintes locales \longrightarrow **champ de Markov**

$$P_T(\mathbf{u}) = \frac{1}{Z_T} e^{-\frac{\Phi(\mathbf{u})}{T}} \quad \text{où} \quad \Phi(\mathbf{u}) = \lambda^2 \sum_c \varphi(D_c \mathbf{u} / \delta)$$

Opérateur de **différence**

Cliques c : \longleftarrow sur les cliques c

groupes de pixels voisins deux à deux

$$\Phi(\mathbf{u}) = \lambda^2 \sum_{i,j} \varphi\left(\frac{|D_x \mathbf{u}|_{i,j}}{\delta}\right) + \varphi\left(\frac{|D_y \mathbf{u}|_{i,j}}{\delta}\right)$$

Exemple : voisinage ordre 1 / cliques associées



Interprétation stochastique

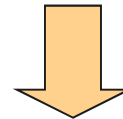
Maximum a **posteriori** : $\text{Max}_u P(u/g)$

$$P(u/g) \propto P(g/u) P(u) \quad (\text{Bayes})$$



Vraisemblance : loi du bruit $\mathbf{n} = \mathbf{g} - \mathbf{H}\mathbf{u}$
cas gaussien $\propto \exp(-\|\mathbf{g} - \mathbf{H}\mathbf{u}\|^2 / 2\sigma^2)$

Loi **a priori** : modèle de Markov 2D
 $\propto \exp(-\Phi(U))$

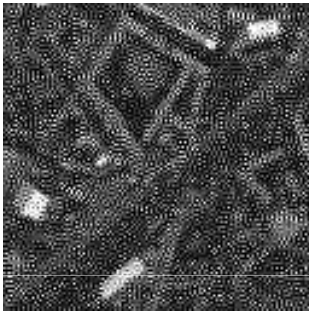


$$J(\mathbf{u}) = \sum_{i,j} (\mathbf{g} - \mathbf{H}\mathbf{u})_{i,j}^2 + \lambda^2 \sum_{i,j} \varphi\left(\frac{|D_x \mathbf{u}|_{i,j}}{\delta}\right) + \varphi\left(\frac{|D_y \mathbf{u}|_{i,j}}{\delta}\right)$$

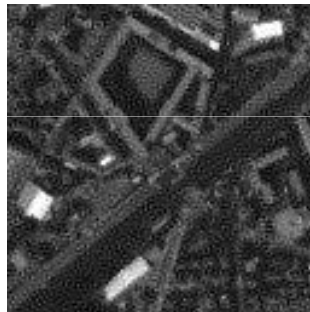
[Hebert & Leahy 94, Green 90, Geman & Reynolds 92...]

Choix des paramètres ?

$\lambda=0.1, \delta=5$



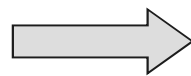
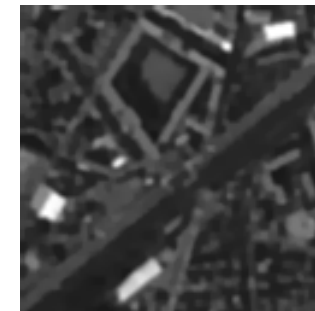
$\lambda=0.3, \delta=50$



$\lambda=0.3, \delta=0.5$



$\lambda=2, \delta=5$



Estimation stochastique

Espace de minimisation ?

- ◆ Critère en variables continues : $u : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$
 $x \rightarrow u(x)$

$$J(u) = \int_{\Omega} |g - Hu|^2 dx + \lambda \int_{\Omega} |\nabla u| dx$$

- ◆ A priori il faut minimiser dans l'espace $\{Hu \in L^2(\Omega), \nabla u \in L^1(\Omega)^2\}$
- ◆ Un tel espace, comme l'espace de Sobolev $W^{1,1}(\Omega) = \{u \in L^1(\Omega), \nabla u \in L^1(\Omega)^2\}$ est
 - **Non reflexif** : difficulté pour montrer existence d'une solution
 - **trop régulier**, il ne contient pas des fonctions qui ont des sauts.

[Evans & Gariepy "Measure Theory and Fine properties of functions" Studies in advanced mathematics, CRC Press, Berlin 1992].

- ◆ Un espace de minimisation contenant des fonctions qui peuvent avoir des discontinuités à travers des courbes \rightarrow **BV**
 Espace des fonctions à **variations bornées**.

Espace BV

- ◆ Le gradient est considéré comme une mesure et on calcule la **variation totale** de cette mesure.

$$J_{TV}(u) = \int_{\Omega} |Du| = \sup \left\{ \int_{\Omega} u \operatorname{div}(p) \mid p \in C_c^1(\Omega, \mathbb{R}^2), \|p\|_{\infty} \leq 1 \right\}$$

La norme du sup est définie par : $\|p\|_{\infty} = \sup_{x \in \Omega} |p(x)|$

C_c^1 est l'ensemble des fonctions C^1 à support compact dans Ω .

L'opérateur de dérivation est reporté sur une fonction test régulière.

[Luigi Ambrosio, Nicola Fusco, Diego Pallara “Functions of bounded variation and free discontinuity problems“ Oxford ; New York : Clarendon Press, cop. 2000].

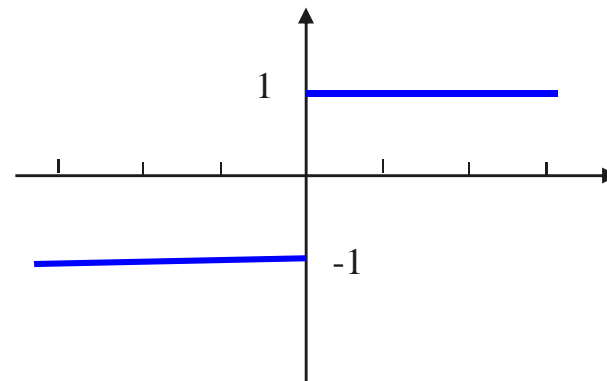
Exemples

◆ $u \in C^1(\Omega)$

$$\text{alors } \int_{\Omega} u \operatorname{div}(p) \, dx = - \int_{\Omega} \nabla u \cdot p \, dx \quad \text{ainsi } \int_{\Omega} |Du| = \int_{\Omega} |\nabla u| \, dx$$

◆ u définie sur $[-1,1]$ par $u(x) = \begin{cases} -1 & \text{si } -1 \leq x < 0 \\ 1 & \text{si } 0 < x \leq 1 \end{cases}$

$$\text{alors } Du = 2\delta_0 \quad \text{et} \quad \int_{-1}^1 |Du| = 2$$

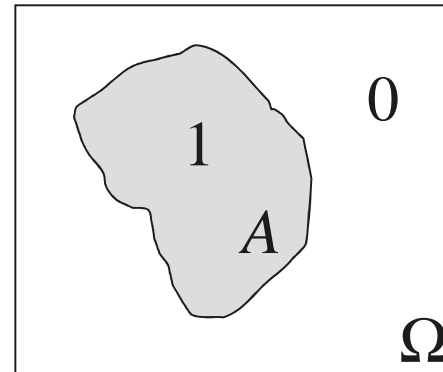


Espace BV

- ◆ Exemple en 2D : χ_A une fonction caractéristique d'ensemble :

$$\chi_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$$

$$\int_{\Omega} |D\chi_A| = \text{Per}_{\Omega}(A)$$



Espace BV

◆ Le critère à minimiser est défini par $J(u) = \int_{\Omega} |g - Hu|^2 dx + \lambda \int_{\Omega} |Du|$

qui est alors minimisé dans l'espace $\left\{ u : \Omega \rightarrow \mathbb{R} / Hu \in L^2(\Omega), \int |Du| < \infty \right\}$

◆ si H est un opérateur linéaire continu n'annulant pas les constantes alors il y a existence et unicité d'une solution au problème de minimisation dans BV.

[A. Chambolle, P-L. Lions, Numer. Math. 76(2),p.167-188,1997]

Espace BV

◆ En discret, ou si u régulière $J_{TV}(u) = \int_{\Omega} |Du| = \sum_{i,j} |\nabla u|_{i,j} = \|\nabla u\|_1$

◆ Alors...intérêt de BV ???

- Bon espace de fonctions pour les images géométriques (cartoon)
- Résultats d'existence et unicité de solution
- Nouveaux algorithmes de minimisation

[A. Chambolle "An Algorithm for Total Variation Minimization and Applications."
Journal of Mathematical Imaging and Vision 20 (1-2): 89-97, January - March, 2004.]

Formule de la co-aire

- ◆ Formule de la co-aire.

si $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ est Lipschitz et $g \in L^1(\mathbb{R}^2)$ alors

$$\int_{\mathbb{R}^2} g(x) |\nabla u(x)| dx = \int_{\mathbb{R}} \left[\int_{u=\rho} g(x) ds \right] d\rho$$

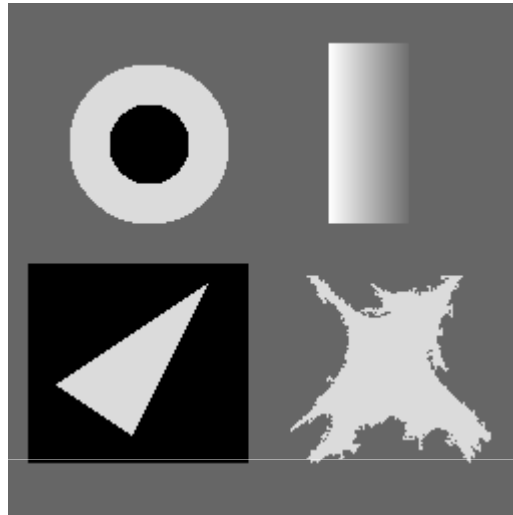
- ◆ Extension pour u dans BV et sur un domaine borné Ω

$$\int_{\Omega} |Du| = \int_{\mathbb{R}} \left[\int_{\{x/u(x)=\rho\}} ds \right] d\rho$$

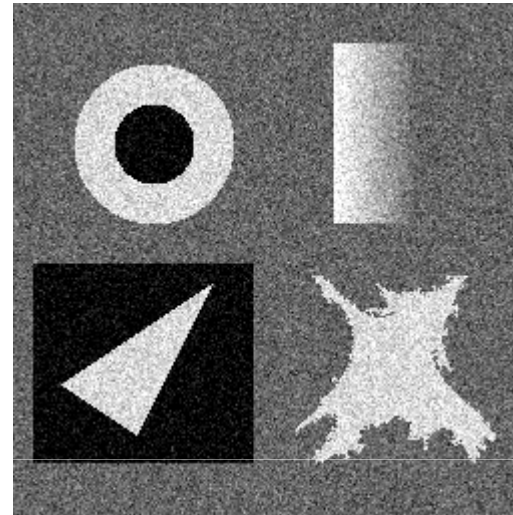
la VT de u est la somme des longueurs de ses lignes de niveaux.

Exemple

originale



bruitée



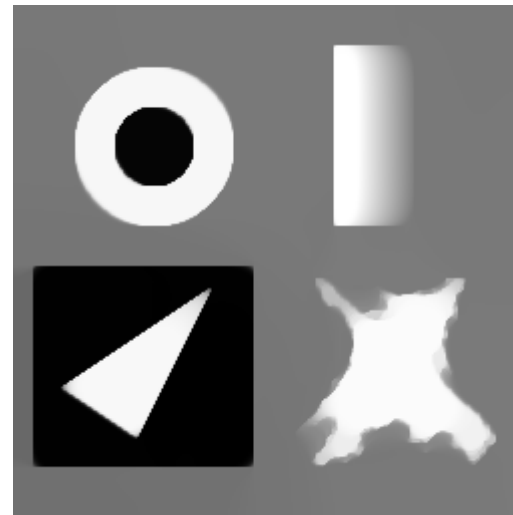
régularisation

L^2



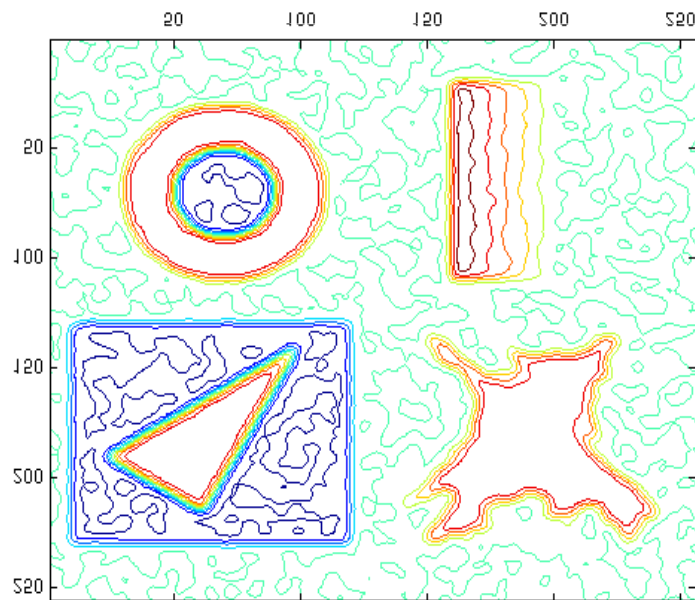
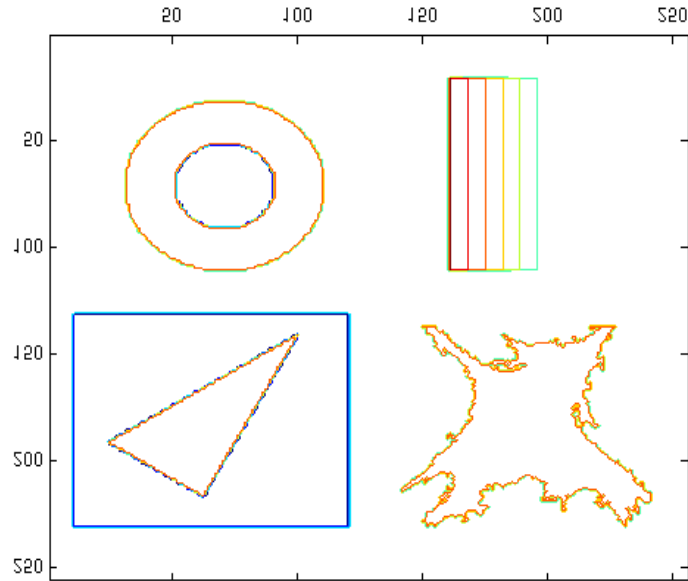
régularisation

L^1

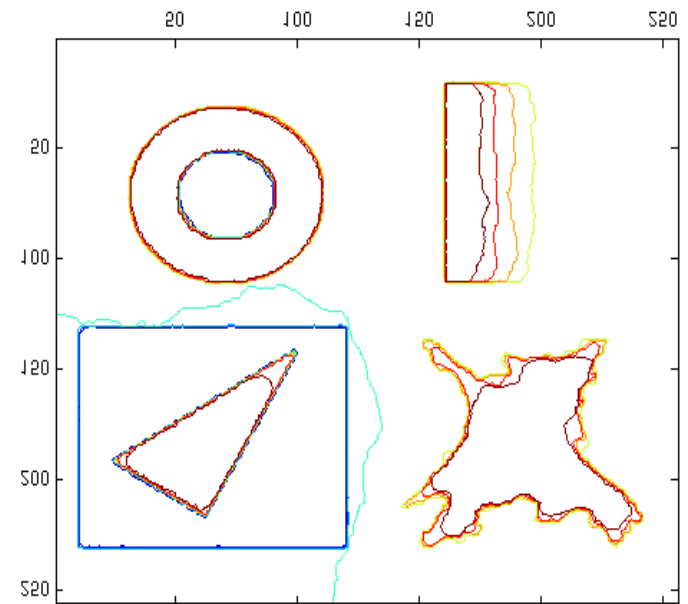


Lignes de niveaux

originale



régularisation L^2



régularisation L^1

Définition des sauts

Points de Lebesgues

- ◆ Pour toute fonction $u \in L^1$ on a pour $x \in \Omega$ p.p.:

$$u(x) = \lim_{r \rightarrow 0} \frac{1}{|B(x, r)|} \int_{B(x, r)} u(y) dy \quad (*)$$

où $B(x, r)$ est la boule unité de centre x et rayon r . Un point x où (*) est vrai est appelé un **point de Lebesgue** de u .

Sauts

- ◆ Le complément des points de Lebesgue de u , noté S_u , est appelé l'ensemble des **sauts** de u .

Limite approchée

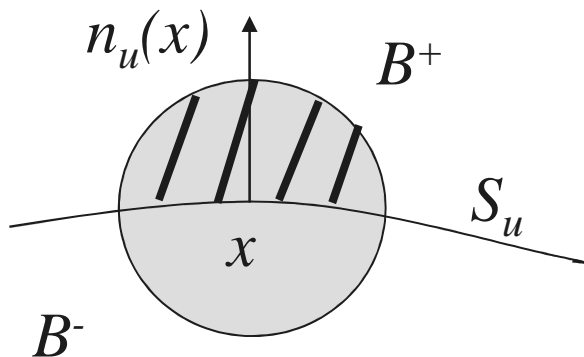
- ◆ Si $u \in BV$ on peut dire plus que (*). Si $x \in S_u$, on peut définir p.p. une normale $n_u(x)$ et des limites

- ◆ Limite supérieure

$$u^+(x) = \lim_{r \rightarrow 0} \frac{1}{|B^+(x, r)|} \int_{B^+(x, r)} u(y) dy$$

- ◆ Limite inférieure

$$u^-(x) = \lim_{r \rightarrow 0} \frac{1}{|B^-(x, r)|} \int_{B^-(x, r)} u(y) dy$$



On a toujours $u^+(x) \geq u^-(x)$.

L'ensemble des sauts de u est S_u ,

l'ensemble des x où $u^+(x) \neq u^-(x)$.

Décomposition dans BV

[Giusti 84, Evans & Gariepy 92, Ambrosio 98]

$$Du(\Omega) = \underbrace{\nabla u \cdot dx}_{\text{Partie absolument continue}} + \underbrace{(u^+ - u^-) n_u d\mathcal{H}^1|_{S_u}}_{\text{Partie singulière (sauts)}} + \underbrace{C(u)}_{\text{Partie Cantor}}$$

Partie absolument continue

Partie singulière (sauts)

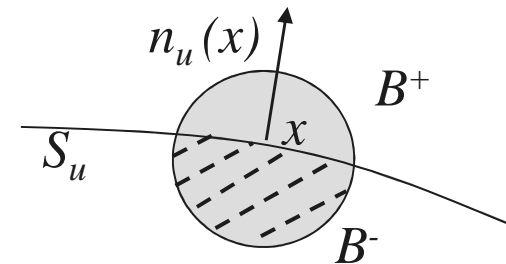
Partie Cantor

- S_u = ensemble des sauts de u = ensemble des points x où $u^+(x) \neq u^-(x)$

(\approx ensemble des points de discontinuité de u).

- \mathcal{H}^1 : mesure de Hausdorff (longueur pour des courbes non régulières).

- $n_u(x) : (x \in S_u)$ normale orientée vers la plus grande valeur de u .



Décomposition dans BV ou SBV

SBV

$$\int_{\Omega} |Du| = \int_{\Omega} |\nabla u| dx + \int_{S_f} (u^+ - u^-) d\mathcal{H}^1 + \int_{\Omega \setminus S_f} |C(u)|$$

➤ Dans SBV, avec φ convexe

$$\int_{\Omega} \varphi(|Du|) = \int_{\Omega} \varphi(|\nabla u|) dx + \varphi^{\infty}(1) \int_{S_f} (u^+ - u^-) d\mathcal{H}^1$$

$$\varphi^{\infty}(1) = \lim_{s \rightarrow +\infty} \frac{\varphi(s)}{s}$$

*Partie
régularisation sur
le gradient*

*Longueur pondérée
des courbes de
discontinuité*

Régularisation dans BV

- ◆ BV est le “bon” espace de fonctions pour les images, car c’est un espace de fonctions qui peuvent être discontinues sur des courbes, donc qui peuvent représenter des images avec des contours.

- ◆ Difficultés :
 - Pas d’équation d’Euler simple dans BV (on travaille avec des mesures)
 - Comment numériquement appréhender les solutions dans BV ?

- ◆ Solutions :
 - Algorithme de projection proposé par A.Chambolle [Workshop MIA 02]
 - Plus généralement algorithme en dualité
 - Variable auxiliaire et Γ -convergence,
 - Modèle explicite des contours par des courbes représentées par des ensembles de niveaux.
 - ...

Variables discrètes

- ◆ En variables discrètes, la variation totale (TV) est l'ensemble des accroissements de f

$$\int_{\Omega} |Du| = \int_{\Omega} |\nabla u| dx \approx \sum_{i,j} |\nabla u|_{i,j}$$

- ◆ On choisit

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

Variables discrètes

- ◆ La norme discrète L^1 recommandée est

$$|\nabla u|_{i,j} = \sqrt{(u_{i+1,j} - u_{i,j})^2 + (u_{i,j+1} - u_{i,j})^2}$$

- ◆ Il est préférable de discrétiser en bout de chaîne. Si on considère le gradient dans L^1 , l'équation d'Euler est

$$H^*(Hu - g) - \lambda^2 \operatorname{div} \left(\frac{\varphi'(|\nabla u|)}{|\nabla u|} \nabla u \right) = 0$$

$$\frac{\partial u}{\partial n} = 0$$

Variables discrètes

- ◆ On introduit la divergence discrète, par analogie à la divergence continue, par

$$\mathbf{div} = -\nabla^*$$

$$\forall \mathbf{p} \in \mathbf{R}^{N^2} \times \mathbf{R}^{N^2} \text{ et } \forall \mathbf{u} \in \mathbf{R}^{N^2} \quad \langle -\mathbf{div}(\mathbf{p}), \mathbf{u} \rangle_{\mathbf{R}^{N^2}} = \langle \mathbf{p}, \nabla \mathbf{u} \rangle_{\mathbf{R}^{N^2} \times \mathbf{R}^{N^2}}$$

- ◆ On vérifie facilement que

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

Minimisation dans L^1

- ◆ Problème de linéarisation du terme de régularisation.

$$\frac{\partial u}{\partial t} = H^*(g - Hu) + \lambda^2 \operatorname{div} \left(\frac{\varphi'(|\nabla u|)}{|\nabla u|} \nabla u \right)$$

- ◆ Algorithme itératif

$$u^{n+1} = u^n + \delta t \left\{ H^*(g - Hu) + \lambda^2 \operatorname{div} \left(\frac{\varphi'(|\nabla u|)}{|\nabla u|} \nabla u \right) \right\}$$

- ◆ Pour $\varphi(t) = |t|$, Rudin Osher Fatemi (ROF 92) ont proposé un algorithme résolvant l'équation d'Euler associée, numériquement stable. Estimation du paramètre λ en utilisant la contrainte

$$\frac{1}{|\Omega|} \int_{\Omega} (g - Hu)^2 dx = \sigma^2$$

si $u_{ij}^n = u(ih, jh, n\Delta t)$ une itération est

$$u_{ij}^{n+1} = u_{ij}^n + \frac{\Delta t}{h} \left\{ \Delta_-^x u_{ij}^n \left(\frac{\Delta_+^x u_{ij}^n}{\left((\Delta_+^x u_{ij}^n)^2 + (m(\Delta_+^y u_{ij}^n, \Delta_-^y u_{ij}^n))^2 \right)^{1/2}} \right) + \right. \\ \left. \Delta_-^y u_{ij}^n \left(\frac{\Delta_+^y u_{ij}^n}{\left((\Delta_+^y u_{ij}^n)^2 + (m(\Delta_+^x u_{ij}^n, \Delta_-^x u_{ij}^n))^2 \right)^{1/2}} \right) \right\} - \Delta t \lambda^n (u_{ij}^n - g_{ij})$$

avec $\Delta_{\pm}^x u_{ij} = \pm(u_{i\pm 1j} - u_{ij})$ et $m(a, b) = \min \text{mod}(a, b) = \left(\frac{\text{sgn } a + \text{sgn } b}{2} \right) \min(|a|, |b|)$

et $\frac{\Delta t}{h^2} \leq c$ pour la stabilité

La fonction minmod est un **limiteur** qui permet d'obtenir un **schéma numérique non oscillant**. Par exemple $\min \text{mod}(\Delta_+, \Delta_-) = 0$ si $\text{sgn}(\Delta_+) \neq \text{sgn}(\Delta_-)$

Algorithme de A. Chambolle

- ◆ Vraie minimisation dans BV (et non gradient dans L^1)

$$\int_{\Omega} \varphi(|Du|) = \int_{\Omega} \underbrace{\varphi(|\nabla u|)}_{\text{Partie régularisation sur le gradient}} dx + \underbrace{\varphi^{\infty}(1) \int_{S_f} (u^+ - u^-) d\mathcal{H}^1}_{\text{Longueur pondérée des courbes de discontinuité}}$$

Partie
régularisation sur le
gradient

Longueur pondérée
des courbes de
discontinuité

- ◆ On s'intéresse à la minimisation de la fonctionnelle (bruit uniquement)

$$\text{Min}_u \frac{1}{2\lambda} \|u - g\|^2 + \int_{\Omega} |Du|$$

$$TV(u) = \int |Du| = \sup \left\{ \int u \operatorname{div}(p) / \|p\|_{\infty} \leq 1 \right\}$$

- ◆ Algorithme de projection, obtenu par dualité.

Quelques notions de dualité

[Rockafellar, R.T., *Convex Analysis*, Princeton University Press, 1970]

Soit X un espace de Hilbert, et soit $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ telle que $f \neq +\infty$

- ◆ Définition : **Conjuguée de Legendre-Fenchel** (ou fonction **polaire**) :

$$f^* : X \rightarrow \mathbb{R} \cup \{+\infty\}$$

$$f^*(s) = \sup_{x \in X} \{\langle s, x \rangle - f(x)\}$$

- ◆ Propriété : La fonction conjuguée f^* est une fonction convexe sci (NB : pas d'hypothèse sur f , on peut avoir $f^* \equiv +\infty$).

- ◆ Théorème : Si f est une fonction convexe et sci, alors

- $f^* \neq +\infty$

- $f^{**} = f$

Quelques notions de dualité

◆ **Exemple** : fonction duale de $f(u) = \frac{1}{2\lambda} \|u - g\|^2$ est

$$f^*(q) = \frac{1}{2\lambda} \|\lambda q + g\|^2 - \frac{1}{2\lambda} \|g\|^2$$

En effet $f^*(q) = \sup_u \left\{ \langle u, q \rangle - \frac{1}{2\lambda} \|u - g\|^2 \right\}$

Soit $f^*(q) = \sup_u \left\{ \langle u, q + g \rangle - \frac{1}{2\lambda} \|u\|^2 - \frac{1}{2\lambda} \|g\|^2 \right\}$

Le maximum est atteint en $u = \lambda(q + g)$

Fonctions d'appui et indicatrice

Si K est un ensemble convexe fermé non vide de X .

La **fonction indicatrice** est $\mathfrak{N}_K(v) = \begin{cases} 0 & \text{si } v \in K \\ +\infty & \text{sinon} \end{cases}$

La **fonction d'appui** est : $\delta_K(s) = \sup_{x \in K} \langle s, x \rangle$

Les fonctions indicatrice et d'appui sont **convexes** et **sci**, et **conjuguées** l'une de l'autre.

$$\mathfrak{N}_K^* = \delta_K \quad \text{et} \quad \mathfrak{N}_K^{**} = \delta_K^* = \mathfrak{N}_K$$

Fonctions d'appui et indicatrice

Théorème :

Toute fonction d'appui $\delta_K(s) = \sup_{x \in K} \langle s, x \rangle$ associée à un convexe fermé

non vide K est convexe sci et homogène de degré un

$$\forall t > 0, \forall x \in X, \quad f(tx) = t f(x)$$

Inversement toute fonction $f \neq +\infty$ convexe, homogène de degré un et sci est la fonction d'appui d'un ensemble convexe fermé K_f défini par

$$K_f = \left\{ s \in X, \forall x \in X, \langle s, x \rangle \leq f(x) \right\}$$

Quelques notions de dualité

- ◆ Propriété: Inégalité de Fenchel

soit $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ $\forall (x, s) \in X \times X$ $f(x) + f^*(s) \geq \langle s, x \rangle$

- ◆ Théorème : Dualité de Fenchel

Soient f et g deux fonctions $X \rightarrow \mathbb{R} \cup \{+\infty\}$ convexes.

On suppose que (contraintes de qualification) :

$\exists x_0 \in X$ tel que $f(x_0) < +\infty$, $g(x_0) < +\infty$ et f continue en x_0

Alors on a

$$\inf_{x \in X} \{f(x) + g(x)\} = - \min_{s \in X} \{f^*(s) + g^*(-s)\}$$

Pb primal

Pb dual

NB : Le problème dual a toujours une solution (si hypothèse thm OK), même si le problème primal n'en a pas, et même lorsque l'inf = $-\infty$

Quelques notions de dualité

◆ Relations d'extrémalité

Soient f et g deux fonctions $X \rightarrow \mathbb{R} \cup \{+\infty\}$ vérifiant les hypothèses du théorème de dualité de Fenchel.

Alors le problème dual a toujours une solution. Soit s une telle solution.

Alors les solutions x du problème primal, s'il en a, sont caractérisées par les relations d'extrémalité :

$$\begin{cases} g(x) + g^*(s) = \langle x, s \rangle \\ f(x) + f^*(-s) = -\langle x, s \rangle \end{cases}$$

Algorithme de A. Chambolle

- ◆ TV est la fonction d'appui d'un ensemble convexe K_{TV}

$$TV(u) = \int |Du| = \sup \left\{ \int u \operatorname{div}(p) / \|p\|_{\infty} \leq 1 \right\}$$

$$TV(u) = \delta_{K_{TV}}(u) = \sup_{q \in K_{TV}} \langle u, q \rangle$$

$$K_{TV} = \left\{ \operatorname{div}(p) / p \in X \times X \quad \|p\|_{\infty} \leq 1 \right\}$$

Le problème dual

TV est une fonction d'appui : $TV(u) = \sup_{q \in K_{TV}} \langle u, q \rangle$

Sa polaire est $TV^*(v) = \mathfrak{N}_{K_{TV}}(v) = \begin{cases} 0 & \text{si } v \in K_{TV} \\ +\infty & \text{sinon} \end{cases}$

C'est la fonction caractéristique de l'ensemble K_{TV} .

Et comme la fonction duale de $f(u) = \frac{1}{2\lambda} \|u - g\|^2$ est

$$f^*(q) = \frac{1}{2\lambda} \|\lambda q + g\|^2 - \frac{1}{2\lambda} \|g\|^2$$

Alg. Chambolle

Problème primal

$$\underset{u}{\text{Min}} \frac{1}{2\lambda} \|g - u\|^2 + TV(u)$$

Problème dual

$$\underset{q}{\text{Min}} \frac{1}{2\lambda} \|q - g\|^2 + TV^*\left(\frac{q}{\lambda}\right)$$

Comme TV^* est la fonction caractéristique de K_{TV} , le pb dual est une projection orthogonale

$$\underset{\frac{q}{\lambda} \in K_{TV}}{\text{Min}} \frac{1}{2\lambda} \|q - g\|^2 \longrightarrow q = \Pi_{\lambda K_{TV}}(g)$$

$$u = g - \Pi_{\lambda K_{TV}}(g)$$

Calcul de la projection Π_K

On considère

$$TV(u) = \|Qu\|_1, \quad Q: X \rightarrow \prod_{1 \leq \theta \leq r} \mathbb{R}^{n_\theta}$$

Soit $B = \left\{ p \in \prod_{1 \leq \theta \leq r} \mathbb{R}^{n_\theta}, \|p\|_\infty \leq 1 \right\}$. On a $\forall q, \|q\|_1 = \sup_{p \in B} \langle p, q \rangle$

$$TV(u) = \|Qu\|_1 = \sup_{p \in B} \langle Qu, p \rangle = \sup_{p \in B} \langle u, Q^* p \rangle$$

Or $TV(u) = \delta_{K_{TV}}(u) = \sup_{s \in K_{TV}} \langle u, s \rangle$ donc $K_{TV} = Q^* B$

$$\text{Min}_{q \in \lambda K_{TV}} \|q - g\|^2 \Leftrightarrow \text{Min}_{p \in B} \|Q^* p - g/\lambda\|^2$$

Calcul de la projection Π_K

- ◆ Pour calculer numériquement projection $\Pi_{\eta K_J}(f)$ on copie l'algorithme de itératif de A. Chambolle : on doit calculer

$$\text{Min}_{|p|_{i,j} \leq 1} \left\| Q^* p - f / \eta \right\|^2$$

qui est une minimisation contrainte. En utilisant les multiplicateurs de Lagrange on obtient

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau(Q(Q^* p^n - f / \eta))_{i,j}}{1 + \tau|(Q(Q^* p^n - f / \eta))_{i,j}|}$$

Si $\tau \leq 1 / \|Q^*\|^2$ alors

$$\eta Q^* p^n \rightarrow \eta Q^* \hat{p} = \Pi_{\eta K}(f)$$

Pas différence numérique sur les résultats de restauration...

Restoration/decomposition models via duality

Class of **BV-IP** problems

$$\inf_{u \in X} TV(u) + \lambda \|g - u\|_p^p \quad 1 < p < +\infty$$

- ◆ when $p=2$, it is the Rudin Osher Fatemi (ROF) problem.
 - ◆ case $p=1$ is treated separately.
 - ◆ when $1 < p < \infty$, a solution exists and is unique.
 - ◆ the minimization is not easy due to the non-differentiability of TV.
- we use **duality** to rewrite the problem in the form

$$\inf_{u \in K} J(u), \quad K = \left\{ u / \|g - u\|_\infty \leq \alpha \right\}$$

BV- ℓ^p model and duality

- ◆ The dual problem is defined by

$$\inf_{\{q \in X \times X, \|q\|_\infty \leq 1\}} \langle -\operatorname{div}(q), g \rangle_X + \beta \|\operatorname{div}(q)\|_{p'}^{p'}$$

where $\beta = (\lambda p)^{-1/(p-1)} - \lambda (\lambda p)^{-p'}$ and $\frac{1}{p} + \frac{1}{p'} = 1$

- ◆ Using the extremality relations, the solution u^* of the primal problem is deduced from the solution q^* of the dual problem :

$$u^* = g - \beta p' \|\operatorname{div}(q^*)\|^{p'-2} \operatorname{div}(q^*)$$

- ◆ The dual problem is a **differentiable problem** under a convex constraint, which can be minimized by using standard or accelerated algorithm.

Généralisation à un problème inverse linéaire

- ◆ Résoudre le problème de minimisation

$$\underset{u}{\text{Min}} \frac{1}{2\lambda} \|g - Hu\|_{X_1}^2 + J(u)$$

où J est une norme ou une semi norme dans un espace régularisant.

Typiquement, en discret $J(u) = \|Qu\|_1$



Généralisation de l'algorithme de A. Chambolle [[Chambolle-MIA02](#)]

où $A=I$ et $J(u) = TV$

Algorithme général

L'algorithme peut être réécrit ;

- $v^n = u^n + \mu H^*(g - Hu^n)$ ← Descente de gradient à pas fixe sur $\|g - Hu\|^2$
- $u^{n+1} = \left(I - \Pi_{\lambda\mu K_J}\right) (v^n)$ ← Débruitage sur v^n

[Demol-Defrise-Daubechies04, Figueiredo-Nowak04, Bect-Blanc-Féraud-Aubert Chambolle04]

On peut réécrire encore

$$u^{n+1} = \left(I - \Pi_{\lambda\mu K_J}\right) \left(u^n + \mu H^*(g - Hu^n)\right)$$

■ Algorithme de « type proximal », Forward-Backward

[Wajs-Combettes05, Chaux-Pesquet-Combettes-Wajs07...]

mais sur critère à l'analyse [Combettes 2009]

Algorithmes pour régularisation L^1

$$\mathop{\text{Inf}}_u \frac{1}{2\lambda} \|g - Hu\|_2^2 + \|Qu\|_1 \qquad \mathop{\text{Inf}}_u f(u) + g(u)$$

◆ Des dizaines d'articles sur le sujet

- Dualité
- Second ordre
 - Point intérieurs [Goldfarb, Yin 2005]
 - Second Order Cone Beam Programming [Boyd, Vandenberghe 2004]
- Ensemble de niveaux et algo. graph-cut [Darbon, Sigelle 2005]
- Sous-gradients projeté [Polyak 1987]
- Forward-Backward spitting [Lions 1979, Combettes 2005]
- Méthodes rapides par algorithme de Nesterov [P.Weiss 2008]
- Douglas-Rachford [Lions 1979, Combettes 2007]
-

CONTENU (suite)

◆ Régularisation non linéaire

➤ Régularisation L^2/L^1

$$\|g - Hu\|_2^2 + \lambda \int_{\Omega} \varphi(|\nabla u|) dx$$

➤ Espace BV, Variation Totale

$$\|g - Hu\|_2^2 + \lambda \int_{\Omega} |Du|$$

➤ Solutions dans l'espace BV, dualité, algorithmes

◆ Segmentation d'image, approche variationnelle

➤ Fonctionnelle de Mumford et Shah

➤ Notion de Γ -convergence : lien restauration/segmentation

◆ Et les représentations parcimonieuses ? $\|g - Hu\|_2^2 + \lambda \sum_{j,k} 2^{-\frac{j}{2}} |\langle u, \psi_{j,k} \rangle|$

➤ Régularisation dans le domaine de la transformée en ondelettes

➤ Solution dans un espace de Besov (et algorithmes)

➤ Modèles de régularisation / parcimonie

Segmentation d'image

- ◆ **Segmentation** : trouver une description de l'image en termes de **régions homogènes** et de **contours**.
- ◆ **Littérature très importante**. Synthèse des méthodes variationnelles dans [Morel & Solimini 94].
- ◆ **Formalisation mathématique** : problème de **minimisation globale** qui combine des mesures appropriées pour les régions et les contours.

$$J_{MS}(u, \Gamma) = \int_{\Omega} (g - u)^2 dx + \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx + \alpha \mathcal{H}^1(\Gamma)$$

- $u \in H^1(\Omega \setminus \Gamma) = \left\{ u \in L^2(\Omega \setminus \Gamma), \nabla u \in L^2(\Omega \setminus \Gamma) \right\}$
- $\Gamma \subset \Omega$, fermé dans Ω = ensemble des bords des régions homogènes.
- \mathcal{H}^1 mesure de Hausdorff de dimension 1, (longueur pour des courbes non régulières).

Analogie segmentation/régularisation

$$J_{MS}(u, \Gamma) = \int_{\Omega} (g - u)^2 dx + \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx + \alpha \mathcal{H}^1(\Gamma)$$

⑩ Formulation faible : minimisation pour $u \in SBV$ de

$$\tilde{J}_{MS}(u) = \int_{\Omega} (g - u)^2 dx + \int_{\Omega} |\nabla u|^2 dx + \alpha \mathcal{H}^1(S_u)$$

[De Giorgi, Carriero & Leaci 89]

(u, S_u) où $u \in SBV$, est minimum de $\tilde{J}_{MS}(u)$



(u, Γ) avec $\Gamma = S_u \cup N$ ($N \mathcal{H}^1$ négligeable) est minimum de $J_{MS}(u, \Gamma)$

Analogie segmentation/régularisation

⑩ Formulation faible : minimisation de

$$\tilde{J}_{MS}(u) = \int_{\Omega} (g - u)^2 dx + \underbrace{\int_{\Omega} |\nabla u|^2 dx}_{\text{régularisation}} + \underbrace{\alpha \mathcal{H}^1(S_u)}_{\text{segmentation}} \quad u \in SBV$$

⑩ Régularisation dans BV :

$$\int_{\Omega} \varphi(|Du|) = \int_{\Omega} \underbrace{\varphi(|\nabla u|)}_{\text{régularisation}} dx + \underbrace{\varphi^{\infty}(1) \int_{S_f} (u^+ - u^-) d\mathcal{H}^1}_{\text{segmentation}}$$

⑩ Régularisation par Variation Totale :

$$\int_{\Omega} |Du| = \int_{\Omega} \underbrace{|\nabla u|}_{\text{régularisation}} dx + \int_{S_f} \underbrace{(u^+ - u^-)}_{\text{segmentation}} d\mathcal{H}^1$$

Minimisation

- ◆ Difficultés:
 - Non convexité de la fonctionnelle en (u, Γ)
 - Non unicité de la solution
 - Difficulté pour modéliser des courbes dans \mathbb{R}^2 et minimiser par rapport à des surfaces et des courbes.
 - Il existe des résultats de régularité de la solution (Γ est une union de courbes régulières) dans \mathbb{R}^2 mais pas dans \mathbb{R}^3 .
- ◆ Approximation par Γ -convergence :
 - la fonctionnelle est approchée par une suite de fonctionnelles elliptiques. Des mesures sur des hypersurfaces de dimension $n-1$ (comme $\mathcal{H}^{n-1}(\Gamma)$) peuvent être approchée par des fonctionnelles elliptiques définies dans \mathbb{R}^n .
 - L'idée de base est d'introduire une nouvelle variable b définie de $\Omega \subset \mathbb{R}^2$ dans \mathbb{R} qui contrôle l'ensemble inconnu S_u .
 - Exemple d'approximation: [Ambrosio Tortorelli 90]

$$J_{AT}^\varepsilon(u, b) = \int_{\Omega} (g - u)^2 dx + \lambda^2 \int_{\Omega} b^2 |\nabla u|^2 dx + \frac{1}{\varepsilon} \int_{\Omega} (b - 1)^2 dx + \varepsilon \int_{\Omega} |\nabla b|^2 dx$$

- Variable auxiliaire modélisant les contours, régularisation sur les contours (équivalent minimisation de la longueur des courbes de discontinuités)

Γ -convergence de fonctionnelles

- ◆ Voir dans [Ambrosio Tortorelli 90] La suite de fonctionnelles J_ε tend vers J_{MS} au sens de la Γ -convergence

$$J_{AT}^\varepsilon(u, b) = \int_{\Omega} (g - u)^2 dx + \lambda^2 \int_{\Omega} b^2 |\nabla u|^2 dx + \frac{1}{\varepsilon} \int_{\Omega} (b - 1)^2 + \varepsilon \int_{\Omega} |\nabla b|^2 dx$$

- ◆ Définition : $J_\varepsilon(u)$ Γ -converge vers $J_0(u)$ si $\forall u_0 \in L^1(\Omega)$:

$$\triangleright \forall u_\varepsilon \rightarrow u_0 \quad \liminf_{\varepsilon \rightarrow 0} J_\varepsilon(u_\varepsilon) \geq J_0(u_0)$$

$$\triangleright \exists u_\varepsilon \rightarrow u_0 \quad \lim_{\varepsilon \rightarrow 0} J_\varepsilon(u_\varepsilon) = J_0(u_0)$$

Γ -convergence de fonctionnelles

- ◆ **Propriété « utile »** : Toute suite de minimiseurs u_ε de J_ε admet une sous-suite convergente dont la limite u_0 est un minimiseur de la fonctionnelle limite.
- ◆ Pour trouver un minimiseur de J_{MS} , on minimise successivement les fonctionnelles J_ε pour ε décroissant. Ainsi on ne minimise que des approximations régulières de la fonctionnelle limite.

$$J_{AT}^\varepsilon(u, b) = \int_{\Omega} (g - u)^2 dx + \lambda^2 \int_{\Omega} b^2 |\nabla u|^2 dx + \frac{1}{\varepsilon} \int_{\Omega} (b - 1)^2 dx + \varepsilon \int_{\Omega} |\nabla b|^2 dx$$

Γ -converge vers

$$J_{MS}(u, \Gamma) = \int_{\Omega} (g - u)^2 dx + \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx + \alpha \mathcal{H}^1(\Gamma)$$

Lien avec la restauration

- ◆ Pour la segmentation, une fonctionnelle d'approximation est

$$J_{AT}^{\varepsilon}(u, b) = \int_{\Omega} (g - u)^2 dx + \lambda^2 \int_{\Omega} b^2 |\nabla u|^2 dx + \frac{1}{\varepsilon} \int_{\Omega} (b - 1)^2 dx + \varepsilon \int_{\Omega} |\nabla b|^2 dx$$

- ◆ Pour la restauration avec la variable auxiliaire modélisant les contours, la

fonctionnelle est pour $\varphi_{GM}(t) = \frac{t^2}{1+t^2}$ $\psi(b) = (\sqrt{b} - 1)^2$

et après changement de variable b en b^2 :

$$J(u, b) = \int_{\Omega} (g - u)^2 dx + \lambda^2 \int_{\Omega} b^2 |\nabla u|^2 dx + \int_{\Omega} (b - 1)^2 dx$$

- ➔ lissage de la variable b

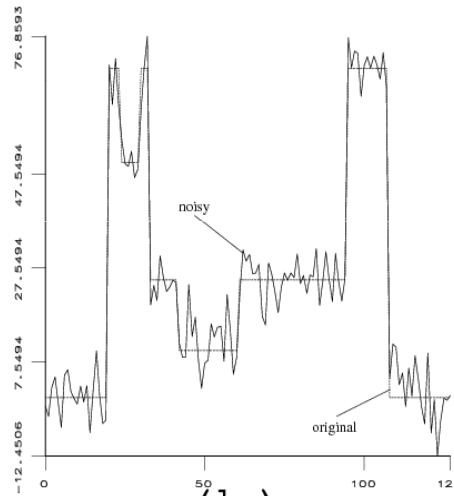
Un résultat

◆ débruitage



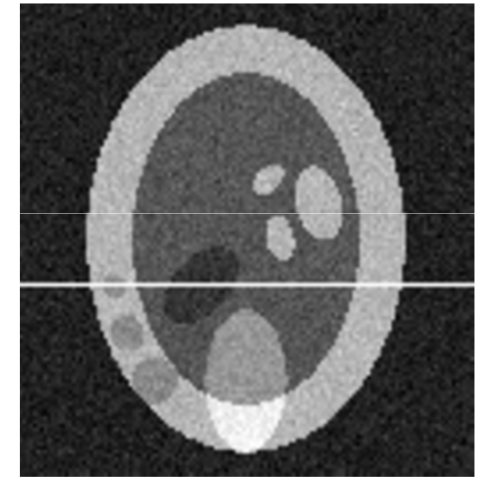
(a)

Image originale
Synthétique constante par
morceaux



(b)

coupe



(c)

Image bruitée
Bruit gaussien
SNR=15 db

Un résultat

- ◆ Fonctionnelle de débruitage : $J(u) = \int_{\Omega} (g - u)^2 dx + \lambda^2 \int_{\Omega} \varphi_{GM} \left(\frac{|\nabla u|}{\delta} \right) dx$

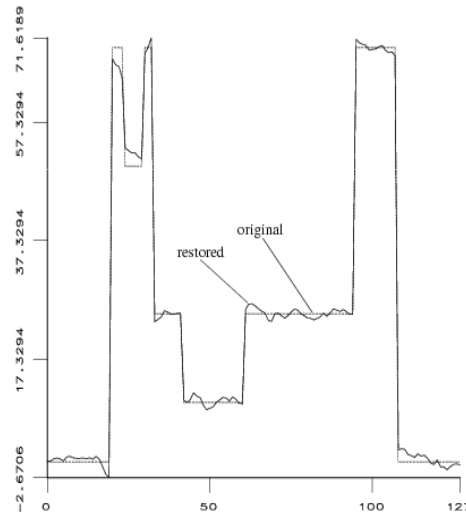
$$J^*(u, b) = \int_{\Omega} (g - u)^2 dx + \lambda^2 \int_{\Omega} b \left(\frac{|\nabla u|}{\delta} \right)^2 dx + \int_{\Omega} (\sqrt{b} - 1)^2 dx$$



(a)

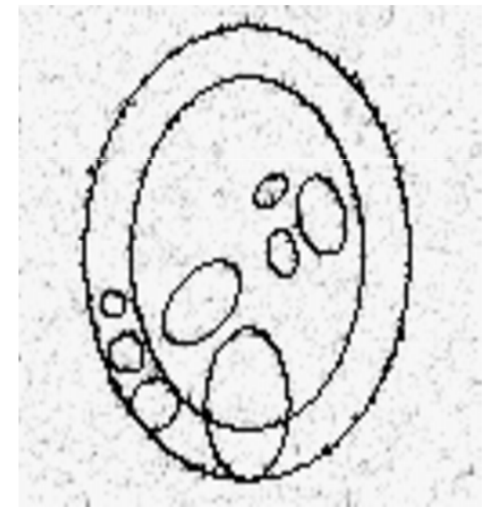
Image restaurée
par modèle semi-
quadratique avec

$$\varphi_{GM}(t) = \frac{t^2}{1+t^2} \quad \psi(b) = (\sqrt{b} - 1)^2$$



(b)

coupe



(c)

Variable auxiliaire b à
convergence
(sur l'image restaurée)

Un résultat avec d'autres paramètres

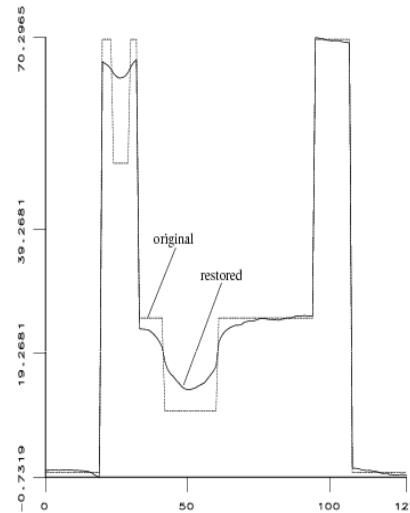
- ◆ Autre jeu de paramètres (λ, δ) pour plus de lissage



(d)

Image restaurée
par modèle semi-
quadratique avec

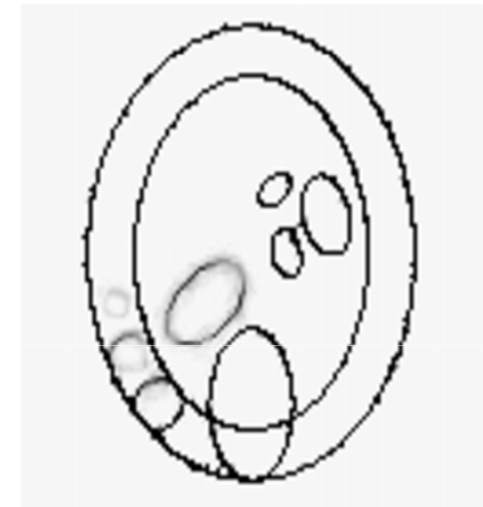
$$\varphi_{GM}(t) = \frac{t^2}{1+t^2}$$



(e)

coupe

$$\psi(b) = (\sqrt{b} - 1)^2$$



(f)

Variante auxiliaire b à
convergence
(sur l'image restaurée)

Lissage de la variable b

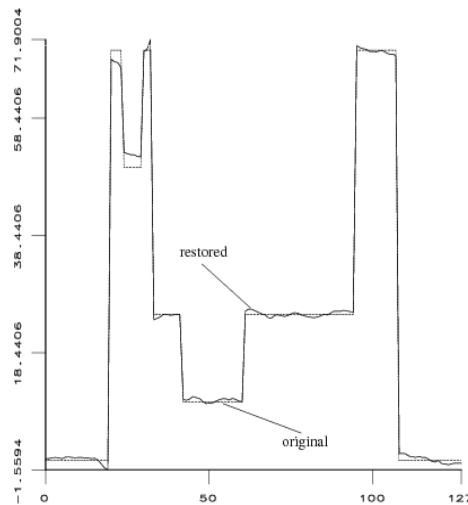
- ◆ Lissage de b avec φ_{GM}

$$J^*(u, b) = \int_{\Omega} (g - u)^2 dx + \lambda^2 \int_{\Omega} b^2 |\nabla u|^2 dx + \int_{\Omega} (b - 1)^2 dx + \int_{\Omega} \varphi_{GM}(b) dx$$



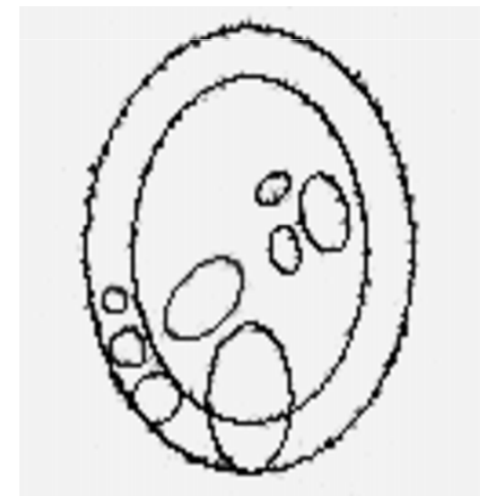
(a)

Image restaurée



(b)

coupe



(c)

Variable auxiliaire b

Lissage de la variable b avec ε

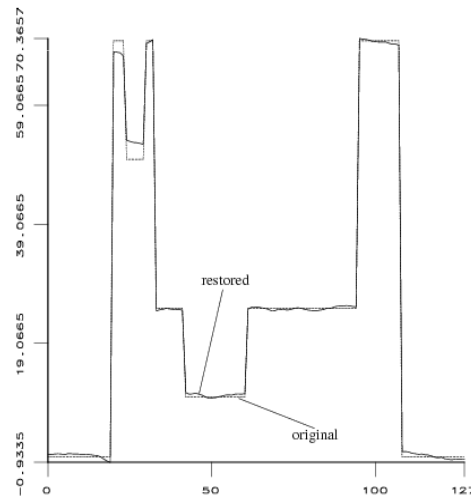
- ◆ Lissage de b avec φ_{GM}

$$J^*(u, b) = \int_{\Omega} (g - u)^2 dx + \lambda^2 \int_{\Omega} b^2 |\nabla u|^2 dx + \frac{1}{\varepsilon} \int_{\Omega} (b - 1)^2 dx + \varepsilon \int_{\Omega} \varphi_{GM}(b) dx$$



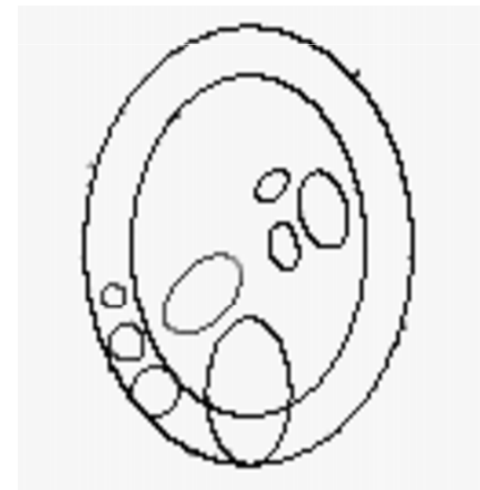
(d)

Image restaurée



(e)

coupe



(f)

Variable auxiliaire b

Du discret au continu

- ◆ Difficulté en discret (en 2D) pour modéliser des courbes
- ◆ Fonctionnelle 1D discrétisée avec un pas h

$$h \sum_i (u_i - g_i)^2 + \sum_i h b_i^2 \left(\frac{u_{i+1} - u_i}{h} \right)^2 + \frac{h}{\varepsilon} \sum_i (1 - b_i)^2 + \alpha \varepsilon h \sum_i \varphi \left(\frac{b_{i+1} - b_i}{h} \right)$$

- ◆ Pour que les résultats théoriques et pratiques soient corrects, il faut que le pas de discrétisation tende vers 0 avec ε . Sinon le terme impose $b_i \rightarrow 1 \quad \forall i$ quand $\varepsilon \rightarrow 0$

Résultat théorique de Γ -convergence sous les hypothèse adéquates vers la fonctionnelle de Mumford et Shah

[Aubert Blanc-Féraud March 2006]

CONTENU (suite)

⑩ Régularisation non linéaire

➤ Régularisation L^2/L^1

$$\|g - Hu\|_2^2 + \lambda \int_{\Omega} \varphi(|\nabla u|) dx$$

➤ Espace BV, Variation Totale

$$\|g - Hu\|_2^2 + \lambda \int_{\Omega} |Du|$$

➤ Solutions dans l'espace BV, dualité, algorithmes

⑩ Segmentation d'image, approche variationnelle

➤ Fonctionnelle de Mumford et Shah

➤ Notion de Γ -convergence : lien restauration/segmentation

⑩ Et les représentations parcimonieuses ? $\|g - Hu\|_2^2 + \lambda \sum_{j,k} 2^{-\frac{j}{2}} |\langle u, \psi_{j,k} \rangle|$

➤ Régularisation dans le domaine de la transformée en ondelettes

➤ Solution dans un espace de Besov (et algorithmes)

➤ Modèles de régularisation / parcimonie

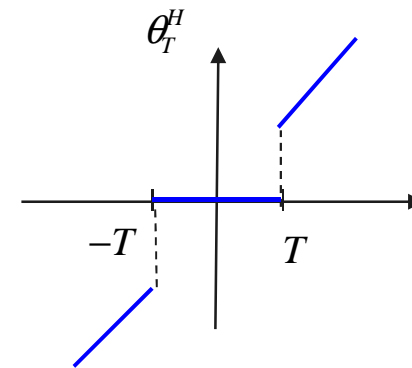
Régularisation par ondelettes

$$\|g - Hu\|_2^2 + \lambda \sum_{i,j=1}^M \left| \langle u, \psi_{i,j} \rangle \right| = \|g - Hu\|_2^2 + \lambda \|\Psi u\|_1$$

- ◆ Domaine des ondelettes : Représentation **parcimonieuse** du signal → bonne **séparation signal/bruit**
- ◆ Débruitage : $g = u + n \rightarrow c^g = c^u + c^n$

⑩ Par **seuillage** : $\hat{c}^u = \theta(c^g)$

■ seuillage **dur** : $\theta_T^H(t) = \begin{cases} t & \text{si } |t| > T \\ 0 & \text{si } |t| \leq T \end{cases}$

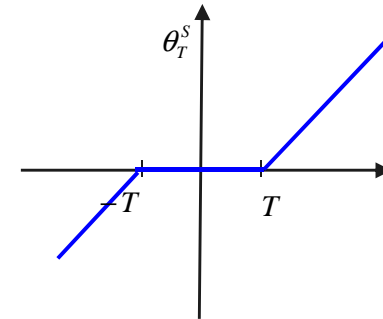


Le seuillage dur préserve les forts coefficients (supérieurs à T) = contours
→ pas de lissage des contours.

Les petits coefficients (inférieurs à T) sont mis à 0 : dans les zones où les transitions de l'image sont faibles, on fait une moyenne locale des coefficients bruités.

Débruitage par ondelettes

■ seuillage mou : $\theta_T^s(t) = \begin{cases} t - T & \text{si } t > T \\ t + T & \text{si } t < -T \\ 0 & \text{si } |t| \leq T \end{cases}$



Les forts coefficients sont atténués de T .

La fonction de seuillage est plus régulière (continue).

L'estimateur de **seuillage est optimal** (au sens risque minimax) pour le seuil $T = \sigma\sqrt{2\log M^2}$, parmi les opérateurs diagonaux dans une base d'ondelettes.

[D. Donoho I. Johnstone, 1994 - ...]

Interprétation variationnelle du seuillage des coefficients en ondelettes

- Domaine des ondelettes : Représentation parcimonieuse du signal
→ Trouver u qui minimise

$$\|g - u\|^2 + \lambda \#_{\{c^u \neq 0\}}$$

$$\|c^u - c^g\|^2 + \lambda \#_{\{c^u \neq 0\}}$$

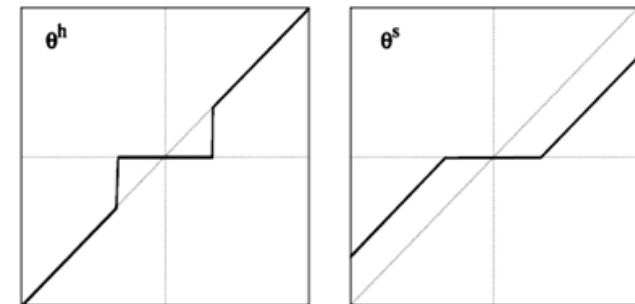
Solution : $c^u_{opt} = \theta_\lambda(c^g)$ seuillage **dur**.

- Trouver u qui minimise

$$\|u - g\|_2^2 + \lambda \|c^u\|_1$$

$$\Leftrightarrow \|c^u - c^g\|_2^2 + \lambda \|c^u\|_1$$

Solution : $c^u_{opt} = \theta_\lambda(c^g)$ seuillage **doux**



Régularisation par TO

- Extension à la déconvolution, **modèle de régularisation par ondelettes (Analyse)**

$$(1) \quad \min_u \left\| g - Hu \right\|_2^2 + \lambda \left\| \Psi u \right\|_1$$

- Si Ψ est une base, il y a équivalence avec le **modèle de parcimonie (Synthèse)**

$$(2) \quad \min_{c_u} \left\| g - H\Psi^* c_u \right\|_2^2 + \lambda \left\| c_u \right\|_1$$

- Algorithmes en deux temps de « type proximal »

[Demol-Defrise-Daubechies04, Figueiredo-Nowak04, Bect-Blanc-Féraud-Aubert-Chambolle04, Combettes 2009] pour le critère de régularisation (1)

[Wajs-Combettes05, Chaux-Pesquet-Combettes-Wajs07...] pour le critère de parcimonie (2)

Comparaison ?

- ◆ Modèle de régularisation

$$\min_u \|g - Hu\|_2^2 + \lambda \|\Psi u\|_1$$

- ◆ Modèle de parcimonie

$$\min_{c_u} \|g - H\Psi^* c_u\|_2^2 + \lambda \|c_u\|_1$$

- Equivalence si transformée sur une base orthogonale
- Utilisation de transformée redondante pour l'invariance par translation (arbre dual, complexe, curvelet...)
- Quel modèle choisir et quelle transformée?

Comparaison ?

Une question ouverte

◆ Un résultat [Elad, Milanfar, Rubinstein 07]

➤ Régularisation $U = \arg \min_u \|g - Hu\|_2^2 + \lambda \|\Psi u\|_1$

➤ Parcimonie $V = \arg \min_{c_u} \|g - HRC_u\|_2^2 + \lambda \|c_u\|_1$

Pour tout Ψ de rang plein, il existe R tel que $\forall g, U = RV$

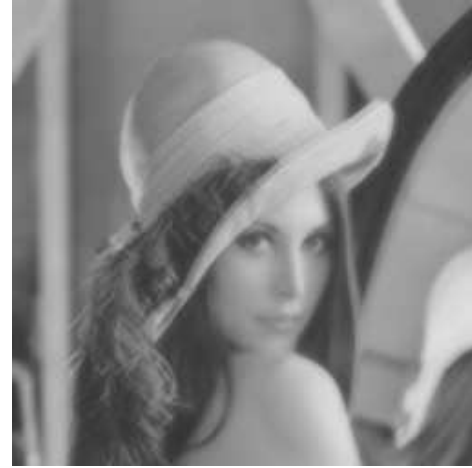
L'inverse n'est pas vrai.

Les a priori de parcimonie sont plus généraux mais pas nécessairement mieux adaptés

Deconvolution results



Lena image



blurred image



blurred+noisy



Wavelet regul.

Modèle de régularisation par ondelettes

$$\min_u \left\| g - Hu \right\|_2^2 + \lambda \left\| \Psi u \right\|_1$$

Ψ est redondante pour l'invariance par translation, c'est une trame et non une base.



image floue et bruitée



Image restaurée avec régularisation par transformée en ondelette complexe sur arbre dual (N. Kingsbury)

Modèle de parcimonie

$$\min_{c_u} \left\| g - H\Psi^* c_u \right\|_2^2 + \lambda \|c_u\|_1$$

Ψ est un dictionnaire de formes, un ensemble de transformée en ondelettes par exemple. La minimisation se fait directement sur les coefficients en ondelettes



image floue et bruitée



Image restaurée avec régularisation par transformée en ondelette complexe sur arbre dual (N. Kingsbury)

Comparaison TV, Ondelettes et Parcimonie

Les ondelettes sont les ondelettes complexes sur arbre duales de N. Kingsbury, c'est une trame et on a une base, il y a de la redondance. Cette transformée est pratiquement invariante par translation et rotation

Image



Restaurée par modèle de parcimonie



Restaurée par modèle de régularisation ondelettes



Restaurée par modèle de régularisation TV

TV et ondelettes

Besov et BV

On peut montrer que

$$B_{1,1}^1([0,1]) \subset BV([0,1]) \subset B_{1,+\infty}^1([0,1])$$

C'est-à-dire que la norme BV s'encadre par deux normes de

Besov

$$A \|u\|_{B_{1,+\infty}^1} \leq \|u\|_{BV} \leq B \|u\|_{B_{1,1}^1}$$

$$\|u\|_{B_{1,+\infty}^1} = \sup_{j \leq J+1} \left(\sum_{m=0}^{2^j-1} 2^{\frac{-j}{2}} \left| \langle u, \psi_{j,m} \rangle \right| \right)$$

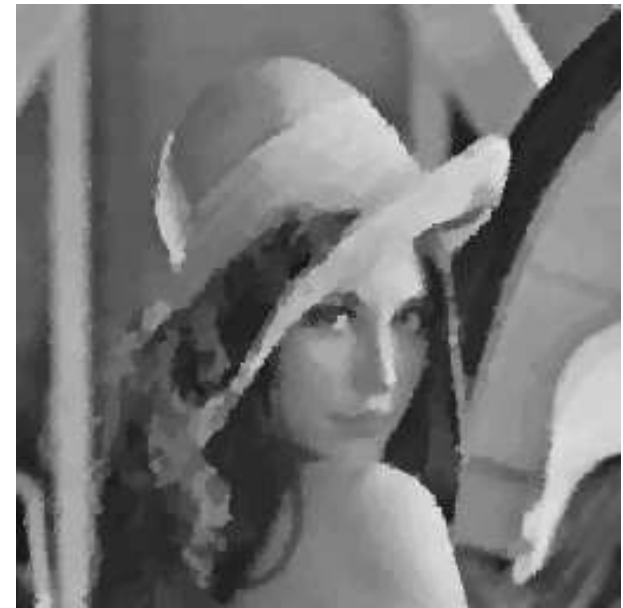
$$\|u\|_{B_{1,1}^1} = \sum_{j=0}^{J+1} \sum_{m=0}^{2^j-1} 2^{\frac{-j}{2}} \left| \langle u, \psi_{j,m} \rangle \right|$$

$$\|u\|_{BV} = \|u\|_1 + \int |Du|$$

Modèle BV- L^1

$$\underset{u}{\text{Min}} \frac{1}{2\lambda} \|g - u\|_1 + \|\nabla u\|_1$$

Résultat pour du bruit poivre et sel 40%



Attache aux données 1¹

- ◆ Restauration (flou et bruit) et passage d'un échantillonnage irrégulier à une grille régulière. Restauration par minimisation de fonctionnelle avec régularisation par Variation totale et décomposition en ondelettes complexes sur un arbre dual de Nick Kingsbury

[J W Selesnick, R G Baraniuk, and N G Kingsbury: "The Dual-Tree Complex Wavelet Transform," IEEE Signal Processing Magazine, vol 22, no 6, pp 123-151, Nov. 2005].

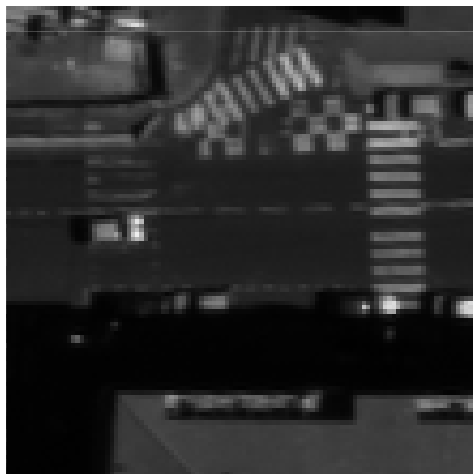
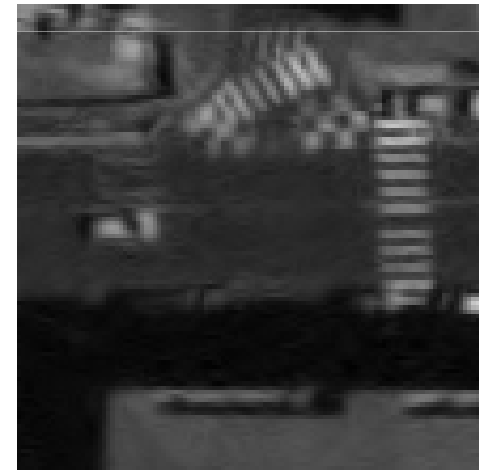


Image observée
/ image originale



régularisation par VT
BV-1¹



régularisation par
ondelettes complexes
1¹ -1¹

En résumé

$$\|g - Au\|_Z + J_Y(u)$$

- ◆ Quel est l'espace Z (et la norme) à utiliser à la place de L^2 , pour modéliser les textures et/ou le bruit dans un algorithme de restauration ?
- ◆ Quel est l'espace Y (et la norme) à utiliser pour la régularisation ?
- ◆ Interprétation stochastique : estimation des paramètres ?
- ◆ Analyse théorique du problème de minimisation ?
- ◆ Algorithmes de minimisation ?
- ◆ Évaluation, analyse des erreurs ?
- ◆ Et le multispectral / hyperspectral ?

Remerciements

- ◆ Gilles Aubert, J-A Dieudonné Laboratory, University of Nice-Sophia Antipolis
- ◆ Antonin Chambolle, CMAPX, Ecole Polytechnique Paris,
- ◆ Jean-François Aujol, CMLA, ENS Cachan, Paris
- ◆ Julien Bect, ex master student in Ariana
- ◆ Pierre Charbonnier, LCPC, Strasbourg
- ◆ Pierre Weiss, LMT, Toulouse
- ◆ Mikael Carlavan, Ariana, Sophia Antipolis
- ◆ et les autres collaborateurs

Peyresq juillet 2009

