Méthodes variationnelles et variation totale en restauration, MATIM (IMA203)

Saïd Ladjal

Introduction

Dans ce court texte nous rappelons le principe des méthodes variationnelles dans le cadre de la restauration des images. Ensuite, nous étudions dans le détail une méthode variationnelle de débruitage utilisant la variation totale comme terme de régularisation.

1 Régularisation

Les images numériques sont acquises par des appareils physiques. L'appareil photographique mesure l'énergie lumineuse qui atteint le capteur à partir de la scène photographiée. D'abord, une image est stockée sous la forme d'un nombre fini de valeurs. Cela implique que la représentation que l'on a de l'image est discrète, en particulier nous ne pouvons pas espérer obtenir une valeur d'énergie lumineuse qui varie trop finement en fonction de la position. À l'opposé, on conçoit que la luminosité réelle de la scène peut varier très rapidement en fonction de la position (penser à un appareil qui photographie un paysage arboré, on imagine bien que l'on ne pourra pas distinguer dans l'image les détails trop fins tels que les pores des feuilles des arbres. À moins d'augmenter de beaucoup la résolution de l'appareil...). Ainsi l'acquisition des images introduit une dégradation inévitable. Le but de la restauration est d'essayer d'atténuer l'effet de la dégradation.

1.1 Modèle d'acquisition

Une image parfaite, f_0 est observée par un appareil photographique qui acquiert l'image discrète g. La relation qui lie g à f_0 est :

$$g = \Pi(f_0 * h + b).$$

Où h représente tous les flous subis par l'image (flou optique et flou d'intégration de l'énergie de chaque pixel du capteur) et b est le bruit qui s'ajoute à l'image. Enfin, Π est l'opération d'échantillonnage sur la grille du capteur.

f_0 est une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ et g est une fonction de $\mathbb{Z}^2$ dans $\mathbb{R}$. Le but de restauration étant de retrouver f_0 à partir de g, on pourrait penser que la tâche est impossible car l'espace des fonctions définie sur $\mathbb{R}^2$ est beaucoup plus vaste que celui des fonctions définies sur $\mathbb{Z}^2$. Nous éclaircissons ce problème au paragraphe suivant.

1.1.1 Cas de la bande limitée

Ici, nous voulons simplifier le modèle d'acquisition afin de ne plus faire apparaître d'images définies sur des domaines différents. Pour cela on va supposer que l'image parfaite, f_0 est à bande limitée (**i.e.** sa transformée de Fourier est à support borné). On sait, par le théorème de Shannon, que cela implique qu'il y a équivalence entre la donnée de f_0 ou la donnée de ses échantillons sur une grille régulières dont les points sont espacés de $1/L$ où L est la largeur de la bande spectrale qui supporte toute la transformée de Fourier de f_0. Si nous appelons f_d la fonction de $\mathbb{Z}^2$ dans $\mathbb{R}$ qui représente l'échantillonnage de f_0, il est clair que f_d dépend linéairement de f. On ne rentre pas dans le détail de cette dépendance (voir les cours de base de traitement du signal).

S'il y a une équivalence entre la donnée de f_0 et la donnée de ses échantillons, alors g (l'image effectivement acquise) dépend linéairement de f_d. Par ailleurs si l'on translate f_d d'une distance entière, alors il est évident que cela correspond à une translation de f_0 de la même distance entière et, comme g dépend de f_0 par une convolution, g est encore translatée de la même distance. Il y a alors une relation linéaire et invariante par translation qui donne g à partir de f_d. g est donc le résultat de la convolution de f_d par un certain filtre que nous noterons h_d (attention, h_d n'est pas forcément le résultat de l'échantillonnage du noyau h). Tout ceci se résume dans le modèle suivant :

$$g = h_d * f_d + b.$$

Dans la pratique on ne manipule que des images de taille finie. Si l'on adapte le dernier modèle au cas de la dimension finie et que l'on représente les images par des vecteurs de $\mathbb{R}^N$ (où N est le nombre de pixels de l'image) alors on a

$$G = A.F_0 + B$$

où G, F et B sont des vecteurs de $\mathbb{R}^N$ et A une matrice de taille $N \times N$. De plus, A représente une convolution discrète. *a priori* la donnée de A demande de connaître N^2 coefficients (ce qui fait 10^{12} coefficients si on manipule des images de un million de pixels), mais comme A est la matrice d'une convolution elle a pour vecteurs propre les ondes pures (vecteurs de Fourier). A est donc diagonalisable dans la base de Fourier et sa description ne demande que N coefficients qui sont ses valeurs propres dans la base de Fourier. Ces coefficients sont la transformée de Fourier du noyau h_d.

1.2 Inversion directe

La solution la plus simple au problème de restauration est d'inverser la matrice A et d'appliquer le résultat à G cela donne :

$$F = A^{-1}G = F_0 + A^{-1}B$$

où F est l'image estimée. La matrice A^{-1} est encore diagonalisable dans la base de Fourier et ses valeurs propres sont les inverses des valeurs propres de A. Or, les noyaux de convolution ont souvent une décomposition dans la base de Fourier (les valeurs propres de A) qui décroît vite en fonction de la fréquence. Par contre le bruit B (que l'on peut supposer blanc c'est-à-dire décorrélé spatialement) distribue son énergie de manière uniforme en fonction de la fréquence. Le terme $A^{-1}.B$ a toutes les chances d'être aberrant du fait du rehaussement de l'énergie du bruit dans les hautes fréquences (voir les transparents et TP pour les expériences).

Pour trouver une estimation F qui soit plus acceptable il faut ajouter une hypothèse de régularité sur l'image F afin de refuser les termes aberrants tels que $A^{-1}.B$. Cela justifie l'introduction de la régularisation et son implémentation pratique par le biais des méthodes variationnelles.

1.3 Régularisation

L'inversion correspond à la minimisation de $\|g - h * f\|^2$, c'est-à-dire que l'on cherche une image, f, qui une fois qu'elle a subi les dégradations de l'appareil photographique donne l'image observée g. Pour introduire une connaissance sur les images naturelles, l'approche variationnelle consiste à demander à ce que l'image restaurée soit régulière. La régularité que l'on demande à l'image peut être déduite de modèles probabilistes (voir cours sur Markov) ou tout autre régularité (voir ce qui suit sur la variation totale) qui se constate dans les images. La manière d'introduire cette contrainte de régularité est de minimiser une fonctionnelle qui mélange la fidélité aux observations (attache aux données, $\|g - h * f\|^2$) et une énergie de régularité, par exemple l'intégrale du gradient au carré. Soit le problème

$$f \text{ minimise } E(f) = \|Af - g\|^2 + \lambda \iint \|\nabla f\|^2. \text{ (A représente la convolution par h)}$$

1.4 Méthodes de de minimisation

Nous donnons ici une méthode de minimisation *ad hoc* adaptée au cas où l'énergie à minimiser est quadratique et le modèle d'acquisition ne fait intervenir que la convolution. Cette méthode a l'avantage d'être très rapide car elle passe par la transformation de Fourier. Par ailleurs elle exprime le résultat dans le domaine de Fourier ce qui permet de faire le lien avec le filtrage de Wiener qui se déduit directement d'hypothèses statistiques sur les signaux traités.

Soit donc à minimiser la fonctionnelle :

$$E(f) = \|Af - g\|^2 + \lambda \iint \|\nabla f\|^2$$

où A est une matrice représentant une convolution, f l'image à trouver et g l'image observée. f et g sont des vecteurs de $\mathbb{R}^N$ où N est le nombre de pixels de l'image. A est une matrice carrée de taille $N \times N$. On peut réécrire cette énergie sous la forme :

$$E(f) = \|Af - g\|^2 + \lambda(\|D_x f\|^2 + \|D_y f\|^2)$$

où D_x est la matrice qui représente la dérivation en x. Elle transforme une image $I(x, y)$ en l'image $I(x + 1, y) - I(x, y)$. Aux bords près, on constate que D_x est la matrice de la convolution par le noyau $h_x = [1 \ -1]$. De même D_y représente le dérivateur en y et elle effectue la convolution par le noyaux $h_y = [1 \ -1]^T$.

Or, l'égalité de Parseval nous dit que la norme au carré [1] d'un vecteur de $\mathbb{R}^N$ est égale à la norme au carré [2] de sa transformée de Fourier. On a donc

$$E(f) = \sum_w \left[|\hat{h}(w)\hat{f}(w) - \hat{g}(w)|^2 + \lambda |\hat{f}(w)|^2 \left(|\widehat{h_x}(w)|^2 + |\widehat{h_y}(w)|^2 \right) \right]$$

où le $\hat{k}$ signifie "transformée de Fourier de k" et w parcourt l'ensemble des fréquences possibles (qui est un ensemble bi-dimensionnel pour les images). On constate que, pour chaque w, la valeur $\hat{f}(w)$ n'apparaît que dans un seul terme de la somme définissant E. Pour minimiser E nous arrivons donc à la conclusion qu'il faut minimiser chacun des termes de la dernière somme par rapport à une variable $\hat{f}(w)$ qui n'apparaît dans une aucun autre terme. Si z est une variable complexe et que l'on veut minimiser l'expression

$$|z\alpha - \beta|^2 + |z|^2 \gamma$$

où α, β sont des complexes et γ un réel positif (trouver leurs expressions pour faire coïncider cette expression avec un terme de la somme définissant E). Alors z doit valoir

$$z = \frac{\overline{\alpha}\beta}{|\alpha|^2 + \gamma}$$

On obtient donc comme minimiseur de E une fonction f dont la transformée de Fourier vérifie

$$\hat{f}(w) = \frac{\overline{\hat{h}(w)}}{|\hat{h}(w)|^2 + \lambda \left(\widehat{h_x}(w)|^2 + |\widehat{h_y}(w)|^2 \right)} \hat{g}(w).$$

Ainsi pour minimise E il suffit de calculer la transformée de Fourier de g et de la multiplier point par point par le coefficient décrit ci-dessus. Enfin, on calcule la transformée de Fourier inverse pour trouver le minimiseur.

1.5 Débruitage

On peut appliquer les même raisonnements que ci-dessus pour trouver une solution au problème du débruitage. Dans un problème de débruitage la dégradation est réduite à un ajout de bruit. Ceci revient à prendre pour noyau h le Dirac, dont la transformée de Fourier est constante égale à 1. En reprenant l'équation ci-dessus dans ce cas-là on a

$$\hat{f}(w) = \frac{1}{1 + \lambda \left(\widehat{h_x}(w)|^2 + |\widehat{h_y}(w)|^2 \right)} \hat{g}(w)$$

Dans la suite nous introduisons un autre type d'énergie de régularité. Comme vous le verrez en TP, cette nouvelle régularité donne de meilleurs résultats de débruitage, mais la méthode de minimisation est bien plus compliquée.

1. D'où l'intérêt d'avoir choisi des normes quadratiques.
2. à une constante près, mais comme nous voulons minimiser E, cela ne change rien au problème de minimiser E ou une constante fois E.

2 Débruitage par variation totale

Au début des années 1990 les auteurs Rudin, Osher et Fatemi ont proposé une méthode de débruitage basée sur la variation totale. Ayant observé une image g suivant le modèle d'observation

$$g = f_0 + b$$

où b est le bruit (supposé blanc de puissance totale σ^2) et f_0 l'image parfaite. Les auteurs proposent comme version débruitée de g l'image f qui vérifie

$$\begin{cases} f \text{ minimise } \iint \|\nabla f\| \\ \text{sous la contrainte } \|g-f\|^2 = \sigma^2 \end{cases}$$

Ceci signifie que l'on suppose l'espace des images mieux représenté par la régularité que l'on appelle "variation totale". On note TV la variation total

$$TV(f) = \iint |\nabla f|.$$

Pour résoudre ce problème sous contrainte il suffit de choisir un paramètre de régularisation λ et de minimiser sans contrainte la fonctionnelle

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

Nous allons introduire dans la section suivantes des notions qui généralisent la notion de gradient et qui seront des outils pour une méthode de minimisation de fonctionnelles faisant intervenir la variation totale.

2.1 Sous-différentielles, transformée de Fenchel et fonctionnelles d'ordre 1

Dans cette section on se donne une fonctionnelle convexe que l'on note J définie de $\mathbb{R}^N$ à valeurs dans $\mathbb{R} \cup \{+\infty\}$. Le produit scalaire usuel est noté $< x|y >$ où x et y sont des vecteurs de $\mathbb{R}^N$.

Dire que J est convexe signifie que

$$\forall x, y \in \mathbb{R}^N, \theta \in [0,1], J(\theta x + (1-\theta)y) \leqslant \theta J(x) + (1-\theta)J(y)$$

avec la convention que la somme de $+\infty$ avec n'importe quel réel (ou $+\infty$) vaut $+\infty$. L'ensemble de définition de J est l'intérieur de l'ensemble des points où elle prend des valeurs finies, il est convexe. En dimension finie, ce qui est notre cas, J est continue sur son ensemble de définition.

Pour illustrer les définitions on prendra la fonctionnelle exemple : $J_0(x) = |x|$ définie sur $\mathbb{R}$ ($N = 1$).

Définition 1. Sous différentielle

Aux points x intérieurs à l'ensemble de définition de J la sous-différentielle est un ensemble de vecteurs de $\mathbb{R}^N$ que l'on not $\partial J(x)$ qui vérifient :

$$v \in \partial J(x) \Leftrightarrow \forall y \in \mathbb{R}^N J(y) \geqslant J(x) + < v|y-x >$$

Cet ensemble est non vide et il est réduit à un seul élément lorsque J est différentiable en x, de plus il est convexe et fermé. On peut le regarder comme l'ensemble des vecteurs qui définissent (par produit scalaire) un hyperplan qui passe sous le graphe de J au point x. On comprends alors que lorsqu'il est réduit à un point, il n'y a qu'un seul hyperplan qui passe sous le graphe ce qui caractérise les points où une fonction convexe est différentiable.

Exemple :

$$\partial J_0(x) = \begin{cases} \{1\} & \text{si} \quad x > 0 \\ [-1, 1] & \text{si} \quad x = 0 \\ \{-1\} & \text{si} \quad x < 0 \end{cases}$$

Proposition 1. critère de minimisation

Si x_0 est un point de l'ensemble de définition de J alors : J atteint sont minimum en x_0 si et seulement si $0 \in \partial J(x_0)$.

démonstration :
J atteint son minimum en x_0 si et seulement si
$\forall y \in \mathbb{R}^N, J(y) \geqslant J(x_0)$ si et seulement si
$\forall y \in \mathbb{R}^N, J(y) \geqslant J(x_0) + < 0|y - x >$ si et seulement si
$0 \in \partial J(x_0)$.$\square$

Définition 2. Transformée de Fenchel

On appelle transformée de Fenchel de J que l'on note J^* la fonctionnelle définie sur $\mathbb{R}^N$ par

$$J^*(v) = \sup_{x \in \mathbb{R}^N} < v|x > -J(x).$$

Il faut la voir comme une fonctionnelle définie sur l'ensemble des hyperplans (dualité introduite par le produit scalaire). Pour un vecteur v définissant un hyperplan, $J^*(v)$ vaut le maximum de la différence entre le graphe de J et l'hyperplan défini par v.

Si le sup est atteint en un point x alors $v \in \partial J(x)$. (le démontrer)

Proposition 2.

J^* est une fonctionnelle convexe et

$$J^{**}(x) = J(x)$$

(aux points x de l'ensemble de définition de J). Autrement dit appliquer deux fois la transformation de Fenchel à une fonctionnelle revient à ne pas la modifier.

Démonstration :
Convexité : Soit u, v deux vecteurs et θ un réel de l'intervalle $[0, 1]$.

$$\forall x, < \theta u + (1-\theta)v|x > -J(x) = \theta(< u|x > -J(x)) + (1-\theta)(< v|x > -J(x)) \leqslant \theta J^*(u) + (1-\theta)J^*(v).$$

La dernière inégalité vient de la définition de J^* comme un sup. Donc

$$J^*(\theta u + (1-\theta)v) \leqslant \theta J^*(u) + (1-\theta)J^*(v)$$

Égalité entre J et J^{**} :

$$J^{**}(x) = \sup_{v\in\mathbb{R}^N} < x|v > -J^*(v).$$

Or, pour tout vecteur v

$$J^*(v) \geqslant < v|x > -J(x) \Rightarrow < x|v > -J^*(v) \leqslant < x|v > -(< v|x > -J(x)) = J(x)$$

D'où

$$J^{**}(x) \leqslant J(x)$$

Soit $v_0 \in \partial J(x)$ alors $J^*(v_0) =< v_0|x > -J(x)$ (par définition de la sous-différentielle (faire un dessin...))

On a

$$J^{**}(x) \geqslant < x|v_0 > -J^*(v_0) = J(x)$$

Ce qui termine la preuve de $J^{**} = J\square$.

Définition 3. Fonctionnelle d'ordre 1

J est dite d'ordre 1 si

$$\forall \lambda \in \mathbb{R}, x \in \mathbb{R}^N, J(\lambda x) = |\lambda| J(x)$$

Proposition 3.

Si J est d'ordre 1 alors il existe un fermé convexe $K \subset \mathbb{R}^N$ tel que

$$J^*(v) = \begin{cases} 0 & si \quad v \in K \\ +\infty & si \quad v \notin K \end{cases}$$

De plus $K = \partial J(0)$.

Démonstration :

D'abord comme $J(0) = J(0x) = 0J(x) = 0$ on a toujours $J^*(v) \geqslant 0$ (prendre $x = 0$ dans le sup définissant J^*). Si $J^*(v) > 0$ alors il existe x tel que

$$< v|x > -J(x) = A > 0$$

Mais alors la suite nx vérifie

$$< v|nx > -J(nx) = n(< v|x > -J(x)) = nA$$

et nA tend vers $+\infty$ avec n. Donc $J^*(v) = +\infty$.

On a donc montré que J^* ne peut prendre que les deux valeurs 0 ou $+\infty$. Si on appelle $L = \partial J(0)$, on a

$v \in L$ si et seulement si

$\forall y, J(y) \geqslant < v|y > -J(0)$ si et seulement si

$\forall y, 0 \geqslant < v|y > -J(y)$ si et seulement si

$J^*(v) \leqslant 0$ (on a déjà que $J^*(v) \geqslant 0$). Il y a donc équivalence entre $J^*(v) = 0$ (i.e. $v \in K$) et $v \in \partial J(0)$ qui est un ensemble convexe fermé[3].$\square$

Exemple :

Montrer directement que

$$J_0^*(v) = \begin{cases} 0 & \text{si} \quad v \in [-1, 1] \\ +\infty & \text{si} \quad |v| > 1 \end{cases}$$

3. montrer en exercice que tout sous-différentielle est un ensemble fermé

Proposition 4. *Pour tous vecteurs x et v on a*

$$v \in \partial J(x) \Leftrightarrow x \in \partial J^*(v)$$

On a
$v \in \partial J(x)$ si et seulement si (1)
$\forall y J(y) \geqslant < v|y - x > + J(x)$ ssi
$\forall y, < v|y > - J(y) \leqslant < v|x > - J(x)$ ssi
$J^*(v) = < v|x > - J(x)$ ssi
$J^*(v) + J(x) = < v|x >$ ssi (2)
$J^{**}(x) + J^*(v) = < v|x >$ (on a remplacé J par J^{**} qui lui est égal) ssi
$x \in \partial J^*(v)$ (on retourne le raisonnement de la ligne (2) vers la ligne (1) en changeant J par J^*) $\square$.

2.2 Retour au problème du débruitage

Soit donc la fonctionnelle

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

La sous-différentielle de E est la somme [4] du gradient de $\frac{1}{2}\|f - g\|^2$ et de la sous-différentielle de TV [5] (démontrer que la sous-différentielle d'une somme de fonctionnelle est la somme des sous-différentielles). Le gradient de $\frac{1}{2}\|f - g\|^2$ est $f - g$ (le démontrer). E est définie sur tout l'espace des images. D'après le critère de minimisation on a

$$\begin{gathered}
f \text{ minimise } E \Leftrightarrow \\
0 \in \partial E(f) = f - g + \lambda \partial TV(f) \Leftrightarrow \\
\frac{g-f}{\lambda} \in \partial TV(f) \Leftrightarrow \\
f \in \partial TV^*(\frac{g-f}{\lambda}) \text{ (proposition 4)} \qquad (1)
\end{gathered}$$

Si maintenant on cherche w qui minimise

$$F(w) = TV^*(\frac{w}{\lambda}) + \frac{1}{2}\|w - g\|^2$$

alors on a

$$\begin{gathered}
w \text{ minimise } F \Leftrightarrow \\
0 \in \partial TV^*(\frac{w}{\lambda}) + w - g \Leftrightarrow \\
g - w \in \partial TV^*(\frac{w}{\lambda}) \qquad (2)
\end{gathered}$$

4. la somme de deux ensembles est définie comme la somme des points qui s'écrivent comme somme d'un élément de chaque ensemble

5. le facteur $\frac{1}{2}$ a été ajouté pour la clarté des calculs, il revient à prendre un λ différent pour se rament au cas sans facteur

En identifiant les équations (1) et (2) on a

$$f \text{ minimise } E \Leftrightarrow$$

$$w = g - f \text{minimise } F$$

Or TV est une fonctionnelle d'ordre 1 et alors TV^* ne prend que les valeur 0 et $+\infty$. Donc si w minimise F il faut au moins que $\frac{w}{\lambda} \in K = \partial TV(0) \Leftrightarrow w \in \lambda K$ (proposition 3). Réciproquement parmi tous les $w \in \lambda K$ il faut choisir le minimiseur de $\|g - w\|^2$ pour minimiser F. Il faut donc choisir w comme le projeté orthogonal de g sur l'ensemble λK. Et enfin prendre $f = g - w$ pour obtenir un minimiseur de E.

On a donc ramené la minimisation de notre fonctionnelle à la projection sur un convexe. Dans la section suivante on donne une définition explicite de K ainsi que l'algorithme de projection sur K qui a été développé par Antonin Chambolle.

2.3 Algorithme de minimisation

On commence par définir dans le cas discret ce que sont le gradient et la divergence. Nous considérons une image de taille $M \times N$ définie sur l'ensemble $[\![0, M-1]\!] \times [\![0, N-1]\!]$.

Définition 4. gradient discret

Le gradient d'une image discrète est une fonction définie sur le même ensemble $[\![0, M-1]\!] \times [\![0, N-1]\!]$ dont la valeur en chaque point est un vecteur à deux dimensions (le nuémro de la dimension est noté en exposant) :

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

avec

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

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

Définition 5. divergence discrète

Si p est un champ de vecteurs défini sur $[\![0, M-1]\!] \times [\![0, N-1]\!]$, c'est-à-dire qu'il fait correspondre pour chaque point (i,j) un vecteur $(p^1_{i,j}, p^2_{i,j})$. La divergence de p notée div p est une fonction à valeurs réelles définie sur $[\![0, M-1]\!] \times [\![0, N-1]\!]$ par

$$(div\ p)_{i,j} = \begin{cases} p^1_{i,j} - p^1_{i-1,j} & si \quad 0 < i < M-1 \\ p^1_{i,j} & si \quad i = 0 \\ -p^1 i,j & si \quad i = M-1 \end{cases} + \begin{cases} p^2_{i,j} - p^2_{i,j-1} & si \quad 0 < j < N-1 \\ p^2_{i,j} & si \quad j = 0 \\ -p^2 i,j & si \quad j = N-1 \end{cases}$$

Hors des bords la divergence d'un champ de vecteurs est la somme de la dérivée en x de la composante x avec la dérivée en y de la composante y. Le gradient transforme une image en un champ de vecteur et la divergence transforme un champ de vecteur en une image. La divergence du gradient d'une image est le laplacien de l'image.

L'ensemble $K = \partial J(0)$ que nous avons vu plus haut est caractérisé ainsi

$$K = \{\text{div } p \ : \ \forall i,j \|p_{i,j}\| \leqslant 1\}$$

Autrement dit K est l'ensemble des images qui peuvent s'écrire comme la divergence d'un champ de vecteurs dont la norme de chaque vecteur ne dépasse pas 1. Nous ne faisons pas la démonstration ici.

Il nous faut projeter g sur λK. Si on sait projeter g sur K alors pour projeter g sur λK il suffit de projeter g/λ sur K puis multiplier le résultat par λ.

Nous décrivons l'algorithme de projection sur K. τ est une constante qui doit être prise $< \frac{1}{8}$. L'algorithme est itératif.

Projection Pour projeter une image g sur K, faire :

Initialiser un champ de vecteurs p^0 à $p^0 = 0$.

Calculer le champ de vecteur p^{n+1} à partir de p^n par la formule

$$p_{i,j}^{n+1} = \frac{p^n + \tau \left(\nabla \left(\text{div } (p^n) - g\right)\right)_{i,j}}{1 + \tau \| \left(\nabla \left(\text{div } (p^n) - g\right)\right)_{i,j} \|}$$

Dans la pratique 20 ou 30 itérations suffisent.

À la fin, le projeté de g sur K est div p^∞ (la divergence du champ de vecteurs limite)

2.4 Exemples

Voir transparents et TP.