



, width: 40%)

Projet de recherche

Master 1 Mathématiques

Le modèle de Potts pour les problèmes inverses en imagerie

Auteurs : Almahaboub Fouad et Daniel Ashraful

Sous l'encadrement de Georges Koepfler

*Basé sur l'article : Joint image reconstruction and segmentation using the Potts model
de Martin Storath, Andreas Weinmann, Jürgen Friel et Michael Unser*

Table des matières

1	Introduction	3
2	Fondements théoriques	3
2.1	Cadre théorique des problèmes inverses en imagerie	3
2.1.1	Difficultés d'inversion d'un problème mal posé	3
2.1.1.1	Reconstruction à partir des données observées	4
2.1.2	Approche bayésienne	4
2.2	Formulation du problème d'optimisation du modèle de Potts	5
2.2.1	Un problème de type non-convexe et non-lisse	6
3	Discrétisation du modèle de Potts	6
3.1	Lagrangien	7
3.2	Lagrangien augmenté et méthode des direction alternées	7
3.3	Application de l'ADMM au modèle de Potts	9
3.3.1	Formulation du problème dans le cas anisotrope	9
3.3.2	Lagrangien augmenté appliqué modèle de Potts	9
3.3.3	Minimisation ADMM	10
3.4	Discrétisation isotropique	11
3.5	Discrétisation générale	14
4	Solutions des sous-problèmes	16
4.1	Fonctionnelle de Potts univariée	16
4.2	Cas d'application de régularisation de Tikhonov	17
5	Convergence	17
6	Conclusion	21
7	Références	22

1 Introduction

Le modèle de Potts, également connu sous le nom de modèle de Mumford-Shah constant par morceaux, est un outil puissant pour la reconstruction et la segmentation d'images dans le cadre des problèmes inverses en imagerie. Ce projet est divisé en deux parties principales. La première partie se concentre sur l'établissement d'un cadre théorique pour les problèmes inverses en imagerie. Cette section explore les caractéristiques des problèmes mal posés, introduit des concepts de régularisation, et formalise le problème d'optimisation que l'on veut étudier. La deuxième partie du projet est dédiée à l'étude algorithmique de la résolution du problème d'optimisation tout en définissant un cadre théorique.

2 Fondements théoriques

Fondamentalement, les problèmes inverses consistent à déterminer les paramètres d'un modèle à partir d'observations ou de mesures du phénomène étudié. Contrairement aux problèmes directs, où l'on part des paramètres pour prédire les observations, les problèmes inverses nécessitent de remonter des observations aux paramètres du modèle.

2.1 Cadre théorique des problèmes inverses en imagerie

En imagerie, on est souvent amenés à devoir construire une image à partir d'une autre : on peut par exemple imaginer qu'une image a été altérée et l'on souhaite la reconstruire, ou on peut vouloir en faire une de toute pièce à partir d'une autre. On identifie ce procédé à un problème inverse : on dispose d'une image f et on veut obtenir une image u . On considère que f est le résultat d'une opération effectuée sur u . On modélise cela sous la forme d'une équation générale $f = Au + \eta$, où :

- A est un opérateur linéaire.
- η est la réalisation d'un bruit, une perturbation.

2.1.1 Difficultés d'inversion d'un problème mal posé

Pour comprendre ce qu'est un problème mal posé, définissons ce qu'est un problème bien posé au sens de Hadamard (Hadamard, 1923). Considérons une équation d'observation $f = A(u)$, où A est un opérateur linéaire entre deux espaces vectoriels normés. Les trois conditions de Hadamard pour qu'un problème soit bien posé sont :

- **Existence d'une solution :** Pour chaque observation f dans l'espace des données, il existe une solution u dans l'espace des images telle que $f = A(u)$.
- **Unicité de la solution :** Pour chaque observation f , il existe au plus une solution u telle que $f = A(u)$.
- **Stabilité de la solution par rapport aux données :** Pour toute suite de données f_n convergeant vers f , les solutions correspondantes u_n doivent converger vers la solution u .

$$\lim_{n \rightarrow \infty} A(u_n) = A(u) \implies \lim_{n \rightarrow \infty} u_n = u \quad (1)$$

Le problème est mal posé si l'une de ces propriétés n'est pas satisfaite. En pratique, les problèmes inverses ne vérifient souvent pas l'une ou l'autre de ces conditions. L'existence peut être assurée en élargissant l'ensemble des solutions (par exemple, en cherchant la solution d'une équation différentielle au sens des distributions), ou en redéfinissant le problème avec moins d'équations (comme dans les moindres carrés). C'est un procédé classique consistant à rétablir l'existence en relaxant la notion de solution. Dans le cas d'un problème à plusieurs solutions, l'unicité peut être obtenue en ajoutant des informations (a priori) sur le modèle pour réduire

l'ensemble des solutions possibles et choisir l'une d'elles. La condition de stabilité n'est (en général) jamais vérifiée pour les problèmes inverses rencontrés en traitement d'images. Ceci rend le calcul pratique de la solution impossible car un peu de bruit sur les données peut rendre la solution calculée très éloignée de la vraie solution. Cela veut dire qu'il ne sera pas possible (indépendamment de méthodes numériques) d'approcher de façon satisfaisante la solution du problème inverse.

2.1.1.1 Reconstruction à partir des données observées

Reprenons la relation $f = Au + \eta$, où nous supposons connu f , c'est-à-dire les données dont on dispose, ainsi que la matrice A et la loi de η . Nous souhaitons retrouver u , ou en tout cas en trouver une bonne approximation. Une première idée consisterait à résoudre le problème d'optimisation $\arg \min_u \|f - Au\|^2$. On considère une fonction de coût F définie par :

$$F(u) = \|f - Au\|^2. \quad (2)$$

Pour minimiser cette fonction, on calcule son gradient et on résout l'équation $\nabla F(u) = 0$:

$$\nabla F(u) = 2A^T(Au - f). \quad (3)$$

En posant $\nabla F(u) = 0$, on obtient :

$$A^T Au = A^T f. \quad (4)$$

Si $A^T A$ est inversible, alors :

$$u = (A^T A)^{-1} A^T f. \quad (5)$$

Malheureusement, bien souvent A ne sera pas inversible, ou alors A sera mal conditionnée, ce qui rend cette approche problématique.

2.1.2 Approche bayésienne

Pour pallier ces problèmes mal posés, on peut adopter une approche bayésienne. En effet, dans l'approche probabiliste, une image est considérée comme étant à variables discrètes (de dimension $M \times N$) et est modélisée par une variable aléatoire multidimensionnelle U définie sur l'ensemble des pixels. Une image u est donc une réalisation de cette variable aléatoire et la calculer consiste à choisir un estimateur de la variable aléatoire U afin d'obtenir sa valeur. On cherche la loi conditionnelle de U sachant F . En utilisant le théorème de Bayes, on calcule la probabilité de vraisemblance suivante :

$$\mathbb{P}_{F=f}(U = u) = \frac{\mathbb{P}_{U=u}(F = f)\mathbb{P}(U = u)}{\mathbb{P}(F = f)}. \quad (6)$$

Dans les problèmes auxquels nous sommes confrontés, nous avons des connaissances a priori sur U : c'est pourquoi nous décidons ici de modéliser $\mathbb{P}(U = u)$ par une fonction $x \mapsto e^{-\lambda R(u)}$ avec R une fonction connue en fonction du problème, et λ un réel strictement positif. Quand on s'intéresse à la quantité de $\mathbb{P}_{U=u}(F = f)$, U est fixé, de sorte que le seul aléa qui reste concerne η . En des termes plus probabilistes, la loi $\mathbb{P}_{U=u}(F = f)$ ne dépend que du rapport entre u et f , c'est-à-dire de la loi du bruit η . Si le bruit η est indépendant de U , alors $\mathbb{P}_{U=u}(F = f) = \mathbb{P}(F = Au + \eta) = \mathbb{P}(\eta = f - Au)$. La distribution de probabilité de la vraisemblance est donnée par la distribution de probabilité du bruit. On considère alors que $\mathbb{P}_{U=u}(F = f)$ est proportionnel à $\exp\left(-\frac{\|f - Au\|^2}{2\sigma^2}\right)$, où σ^2 est la variance de la loi de η : c'est ce que l'on appelle la vraisemblance. On peut alors simplifier l'écriture de $\mathbb{P}_{F=f}(U = u)$ en disant qu'elle est proportionnelle à

$\exp\left(-\frac{\|f-Au\|^2}{2\sigma^2} - \lambda R(u)\right)$. On souhaite alors trouver le u qui maximise cette probabilité telle que f ait la plus grande chance d'être le résultat d'une mesure, selon le modèle de formation $f = Au + \eta$. Autrement dit, on veut déterminer $\arg \max_u \mathbb{P}_{F=f}(U = u)$. Cette estimation peut alors être atteinte par une stratégie du maximum a posteriori. Par croissance du logarithme, on veut donc déterminer $\arg \max_u \ln(\mathbb{P}_{F=f}(U = u))$, c'est-à-dire

$$\arg \max_u -\frac{\|f - Au\|^2}{2\sigma^2} - \lambda R(u), \quad (7)$$

et donc $\arg \min_u \frac{1}{2\sigma^2}\|f - Au\|^2 + \lambda R(u)$.

La plupart des problèmes de reconstruction d'images consisteront donc à résoudre un problème d'optimisation $\arg \min_u \frac{1}{2\sigma^2}\|f - Au\|^2 + \lambda R(u)$. On dira que $\frac{1}{2\sigma^2}\|f - Au\|^2$ est le terme d'attache aux données, et on dira que $\lambda R(u)$ est le terme de régularisation. Il s'agit là de favoriser les images régulières. Pour illustrer ces propos, nous pouvons donner l'exemple d'un bruit blanc gaussien où la vraisemblance est donnée par :

$$\mathbb{P}_{U=u}(F = f) = \frac{1}{(2\pi)^{MN/2} \sigma_\eta^{MN}} \exp\left(-\frac{(f - Au)^T (f - Au)}{2\sigma_\eta^2}\right). \quad (8)$$

Et :

$$\max_u \{\mathbb{P}_{U=u}(F = f)\} \iff \min_u \|f - Au\|^2. \quad (9)$$

L'estimateur du maximum de vraisemblance sous l'hypothèse gaussienne du bruit est l'estimateur des moindres carrés.

2.2 Formulation du problème d'optimisation du modèle de Potts

Dans la section précédente, nous avons examiné les bases théoriques des problèmes inverses en imagerie, en mettant en évidence les difficultés rencontrées par les problèmes mal posés, tels que l'instabilité des solutions et la non-unicité. Il est donc essentiel d'utiliser des méthodes de régularisation afin d'obtenir des solutions stables et fiables. Nous introduisons ainsi le modèle de Potts, également connu sous le nom de modèle de Mumford-Shah constant par morceaux. Ce modèle est particulièrement adapté pour les problèmes de reconstruction et de segmentation d'images, car il pénalise la longueur des discontinuités dans l'image, favorisant ainsi des solutions constantes par morceaux. D'un point de vue mathématique, on cherche à minimiser une fonctionnelle de coût composée de deux termes : un terme de fidélité aux données et un terme de régularisation. Cette fonctionnelle peut être exprimée comme suit :

$$\min_u (\gamma \|\nabla u\|_0 + \|Au - f\|_2^2), \quad (10)$$

où :

- $\|Au - f\|_2^2$ est le terme de fidélité aux données, mesurant la différence entre les données observées f et l'opération de reconstruction Au , où A est l'opérateur linéaire appliqué à l'image u . Il assure que l'image reconstruite u est cohérente avec les données mesurées. Plus la différence est petite, plus u est proche des données observées.
- $\|\nabla u\|_0$ est le terme de saut ou de régularisation, qui pénalise le nombre de discontinuités (ou sauts) dans l'image u . Cette norme compte le nombre de pixels où le gradient de u n'est pas nul, c'est-à-dire où il y a des changements abrupts de valeurs de pixels. Il favorise les solutions constantes par morceaux, où les régions de l'image sont homogènes et séparées par

des frontières nettes. Cela réduit le bruit et les artefacts dans l'image reconstruite, rendant la segmentation plus claire et les régions homogènes mieux définies.

- $\gamma > 0$ est le paramètre de régularisation qui contrôle l'équilibre entre la fidélité aux données et la pénalité de saut. Il est ajusté pour trouver un compromis entre une reconstruction trop lisse (perte de détails importants) et une reconstruction trop bruitée (présence d'artefacts).

Dans ce contexte, le modèle de Potts agit bien comme une méthode de régularisation dans le sens des problèmes inverses. Il introduit une pénalité sur les discontinuités de l'image, ce qui aide à stabiliser la solution en favorisant les reconstructions avec des régions homogènes séparées par des frontières nettes (images constantes par morceaux). Cela réduit l'impact des perturbations dans les données sur la solution finale, ce qui est crucial pour les problèmes inverses mal posés.

2.2.1 Un problème de type non-convexe et non-lisse

Le terme de régularisation $\|\nabla u\|_0$ utilise la norme L^0 , comptant le nombre de discontinuités dans l'image u . Mathématiquement, ce terme peut être exprimé comme :

$$\|\nabla u\|_0 = \sum_{i,j} \mathbb{1}_{\nabla u(i,j) \neq 0} \quad (11)$$

où $\mathbb{1}$ est la fonction indicatrice qui vaut 1 si l'argument est vrai, sinon 0. La nature non-convexe de $\|\nabla u\|_0$ signifie que la fonction de coût associée peut posséder plusieurs minima locaux. Le terme $\|\nabla u\|_0$ est également non-lisse, car il implique des discontinuités dans l'image u . Étant donné que la norme L^0 n'est pas différentiable, cela signifie que $\nabla(\|\nabla u\|_0)$ n'existe pas. En raison de cette nature discrète des discontinuités pénalisées, les méthodes d'optimisation basées sur le gradient, telles que la descente de gradient, ne peuvent pas être appliquées directement sans adaptations spécifiques.

3 Discrétisation du modèle de Potts

La norme $\|\nabla u\|_0$ est complexe à manipuler. Elle implique de compter les discontinuités dans un espace continu, ce qui nécessite des outils mathématiques avancés comme la mesure de Hausdorff. Cela rend la résolution directe de ce problème extrêmement difficile. Nous avons donc besoin d'une discrétisation permettant de transformer le problème continu en un problème discret, où des techniques de régularisation peuvent être appliquées de manière plus efficace et faciliter les calculs numériques. En discrétisant l'image u en une grille de pixels, on peut reformuler le problème de minimisation de manière discrète :

$$u^* = \arg \min_{u \in \mathbb{R}^{m \times n}} \left(\gamma \sum_{s=1}^S \omega_s \|\nabla_{p_s} u\|_0 + \|Au - f\|_2^2 \right) \quad (12)$$

où :

- u^* est la solution optimale recherchée.
- γ est un paramètre de régularisation qui équilibre la fidélité aux données et la régularité de la solution.
- ω_s sont des poids associés à chaque terme de la somme.
- $\nabla_{p_s} u$ représente les différences finies de u dans la direction du vecteur de déplacement p_s .
- $\|\cdot\|_0$ est la norme ℓ^0 , utilisée pour mesurer le nombre de discontinuités dans u .
- A est un opérateur linéaire qui modélise le système d'imagerie, par exemple, la transformation de Radon dans le cas de la tomographie.
- f est l'élément de l'espace des données observées, tel qu'un sinogramme en tomographie.

- $\|Au - f\|_2^2$ est le terme de fidélité aux données, mesurant l'écart au carré entre les données reconstruites Au et les données observées f .

Avant de s'attaquer au problème, introduisons d'abord plusieurs définitions essentielles pour la résolution.

3.1 Lagrangien

Définition

En optimisation, le lagrangien (ou fonction lagrangienne) est une fonction permettant d'étudier les problèmes (d'optimisation) avec contraintes. On l'utilise pour établir des conditions d'optimalité, pour construire des problèmes duaux ou pour analyser la perturbation de problèmes. Le lagrangien est construit à partir des multiplicateurs de Lagrange : si on considère le problème suivant :

$$\forall x \in E \subset \mathbb{R} \quad \min x \in G \text{ avec } G = \{x \in E \mid \varphi_i(x) = 0, i = 1, \dots, m, \psi_j(x) \leq 0, j = 1, \dots, p\} \quad (13)$$

Le lagrangien du problème s'écrit comme :

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i \varphi_i(x) + \sum_{j=1}^p \mu_j \psi_j(x) \quad (14)$$

3.2 Lagrangien augmenté et méthode des direction alternées

Lagrangien augmenté :

Les algorithmes du lagrangien augmenté sont une certaine classe d'algorithmes pour résoudre des problèmes d'optimisation sous contraintes. Elles présentent des similitudes avec les méthodes de pénalité dans le sens où elles remplacent un problème d'optimisation sous contraintes par une série de problèmes sans contrainte et ajoutent un terme de pénalité à l'objectif ; la différence est qu'une méthode du lagrangien augmenté ajoute encore un autre terme, conçu pour agir comme un multiplicateur de Lagrange. Le Lagrangien augmenté est apparenté, mais non identique à la méthode des multiplicateurs de Lagrange.

Vu différemment, l'objectif sans contrainte est le lagrangien du problème contraint, avec un terme de pénalité supplémentaire (l'augmentation).

On cherche à résoudre le problème sous contraintes suivant : $\min f(x)$ sous conditions $c_i = 0 \forall i \in \varepsilon$ où ε désigne les indices des contraintes d'égalité. Ce problème peut être résolu comme une série de problèmes de minimisation sans contrainte. Pour référence, on liste d'abord la k -e étape de l'approche de la méthode de pénalité :

$$\min \varphi_k(x) = f(x) + \mu_k \sum_{i \in \varepsilon} c_i(x)^2 \quad (15)$$

La méthode de pénalité résout ce problème, puis à l'itération suivante, elle résout à nouveau le problème en utilisant une plus grande valeur de μ_k (et en utilisant l'ancienne solution comme estimation initiale ou « démarrage à chaud »).

La méthode du lagrangien augmenté utilise l'objectif non contraint suivant :

$$\min \varphi_k(x) = f(x) + \frac{\mu_k}{2} \sum_{i \in \varepsilon} c_i(x)^2 + \sum_{i \in \varepsilon} \lambda_i c_i(x) \quad (16)$$

et après chaque itération, en plus de la mise à jour des μ_k , la variable λ est également mise à jour selon la règle :

$$\lambda_i \leftarrow \lambda_i + \mu_k c_i(x_k) \quad (17)$$

où x_k est la solution du problème sans contrainte à la k -e étape, c'est-à-dire :

$$x_k = \arg \min \varphi_k(x) \quad (18)$$

La variable λ est une estimation du multiplicateur de Lagrange, et la précision de cette estimation s'améliore à chaque étape.

Méthode des directions alternées :

La méthode des directions alternées est une variante du schéma du lagrangien augmenté qui utilise des mises à jour partielles pour les variables duales. Cette méthode est souvent appliquée pour résoudre des problèmes tels que :

$$\min_x f(x) + g(x) \quad (19)$$

Ceci est équivalent au problème contraint : $\min_{x,y} f(x) + g(y)$ avec $x = y$. Bien que ce changement puisse sembler anodin, le problème peut maintenant être envisagé en utilisant des méthodes d'optimisation sous contraintes (en particulier, la méthode du lagrangien augmenté), et la fonction objectif est séparable en x et y . La double mise à jour nécessite de résoudre une fonction de proximité en x et y en même temps ; la technique ADA (ADMM en anglais pour Alternating Direction Method of Multipliers) permet de résoudre approximativement ce problème en résolvant d'abord pour x avec y fixe, puis en résolvant pour y avec x fixe.

Pour se fixer les idées, nous allons donner un cas d'application simple de l'algorithme des directions alternées.

L'algorithme des directions alternées est un algorithme de résolution de problèmes d'optimisation décomposables, qui cherche à adapter l'algorithme du lagrangien augmenté à ce contexte, alors que cet algorithme détruit cette décomposabilité. Il est typiquement utilisé pour minimiser la somme de deux fonctions dépendant de variables différentes couplées par une contrainte affine :

$$\begin{cases} \inf_{x,y} f(x) + g(y) \\ Ax + By = c \end{cases} \quad (20)$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ et $g : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{+\infty\}$, $A \in \mathbb{R}^{p \times n}$, $B \in \mathbb{R}^{p \times m}$ et $c \in \mathbb{R}^p$

L'algorithme suppose implicitement que minimiser f ou g seul est facile. C'est un algorithme trouvant rapidement une solution avec peu de précision, mais qui demande beaucoup d'itérations pour déterminer une solution avec précision.

Le lagrangien augmenté associé est la fonction :

$$\ell_r : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R} \cup \{+\infty\}, \quad (21)$$

dont la valeur en $(x, y, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ s'écrit :

$$\ell_r(x, y, \lambda) = f(x) + g(y) + \lambda^T (Ax + By - c) + \frac{r}{2} \|Ax + By - C\|_2^2 \quad (22)$$

où $r \geq 0$ est le paramètre d'augmentation. Si $r = 0$ on retrouve le lagrangien du problème.

Une itération passe d'un quadruplet $(x_k, y_k, \lambda_k, r_k) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}$ au suivant $(x_{k+1}, y_{k+1}, \lambda_{k+1}, r_{k+1})$ comme suit :

1. Minimisation en x : $x_{k+1} \in \arg \min_{x \in \mathbb{R}^n} \ell_{r_k}(x, y_k, \lambda_k)$
2. Minimisation en y : $y_{k+1} \in \arg \min_{y \in \mathbb{R}^m} \ell_{r_k}(x_{k+1}, y, \lambda_k)$
3. Nouvelle variable duale : $\lambda_{k+1} := \lambda_k + r_k(Ax_{k+1} + By_{k+1} - c)$
4. Test d'arrêt : $Ax_{k+1} + By_{k+1} \approx c$
5. Nouveau paramètre d'augmentation : r_{k+1}

C'est la minimisation séparée en x et y qui permet de faire de la décomposition lorsque f et g sont séparables et si l'on minimisait d'abord en y puis en x avant de mettre à jour λ , on obtiendrait des itérés différents.

3.3 Application de l'ADMM au modèle de Potts

L'ADMM va décomposer le modèle de Potts en sous-problèmes plus simples qui peuvent être résolus de manière itérative et alternée. Chaque sous-problème est généralement plus facile à résoudre grâce à la séparation des variables et des contraintes.

On introduit une variante anisotrope de la régularisation, où le terme de régularisation $\|\nabla u\|_0$ est discrétisé différemment selon les axes $p_1 = (1, 0)$ et $p_2 = (0, 1)$ et, avec des poids associés $\omega_1, \omega_2 = 1$. Cela conduit à une somme séparée des différences finies horizontales et verticales pour u , ce qui peut être écrit comme :

$$\|\nabla u\|_0 = \|\nabla_{p_1} u\|_0 + \|\nabla_{p_2} u\|_0 \quad (23)$$

Ensuite, le terme de régularisation s'écrit

$$\|\nabla u\|_0 = \|\nabla_{p_1} u\|_0 + \|\nabla_{p_2} u\|_0 = \sum_{(i,j)} |u_{i+1,j} - u_{i,j}| + \sum_{(i,j)} |u_{i,j+1} - u_{i,j}|. \quad (24)$$

3.3.1 Formulation du problème dans le cas anisotrope

L'objectif initial est de minimiser la fonctionnelle :

$$u \mapsto \gamma \|\nabla u\|_0 + \|Au - f\|^2 \quad (25)$$

la discrétisation dans le cas anisotrope nous permet de reformuler le problème de Potts comme le problème d'optimisation sous contrainte suivant :

$$\min_{u_1, u_2, v} \gamma \left(\|\nabla_{p_1} u_1\|_0 + \|\nabla_{p_2} u_2\|_0 \right) + \|Av - f\|_2^2, \quad (26)$$

soumis à $v - u_1 = 0$, $v - u_2 = 0$, $u_1 - u_2 = 0$.

3.3.2 Lagrangien augmenté appliqué modèle de Potts

Le Lagrangien augmenté \mathcal{L} pour ce problème est exprimé comme suit :

$$\begin{aligned} \mathcal{L}(u_1, u_2, v, \lambda_1, \lambda_2, \rho) &= \gamma \left(\|\nabla_{p_1} u_1\|_0 + \|\nabla_{p_2} u_2\|_0 \right) + \|Av - f\|_2^2 \\ &\quad + \langle \lambda_1, v - u_1 \rangle + \frac{\mu}{2} \|v - u_1\|_2^2 + \langle \lambda_2, v - u_2 \rangle \\ &\quad + \frac{\mu}{2} \|v - u_2\|_2^2 + \langle \rho, u_1 - u_2 \rangle + \frac{\nu}{2} \|u_1 - u_2\|_2^2. \end{aligned} \quad (27)$$

Les contraintes font maintenant partie intégrante de la fonctionnelle \mathcal{L} , les variables $\lambda_1, \lambda_2, \rho$ sont les multiplicateurs de Lagrange de dimension $(m \times n)$. Le paramètre $\nu > 0$ contrôle la liaison entre u_1 et u_2 et $\mu > 0$ leur couplage à v .

Nous reformulons \mathcal{L} de manière pratique :

$$\begin{aligned}\mathcal{L}(u_1, u_2, v, \lambda_1, \lambda_2, q) = & \gamma \left(\left\| \nabla_{p_1} u_1 \right\|_0 + \left\| \nabla_{p_2} u_2 \right\|_0 \right) + \|Av - f\|_2^2 \\ & + \frac{\mu}{2} \left\| v - u_1 + \frac{\lambda_1}{\mu} \right\|_2^2 + \frac{\mu}{2} \left\| v - u_2 + \frac{\lambda_2}{\mu} \right\|_2^2 \\ & - \frac{\mu}{2} \left\| \frac{\lambda_1}{\mu} \right\|_2^2 - \frac{\mu}{2} \left\| \frac{\lambda_2}{\mu} \right\|_2^2 + \frac{\nu}{2} \left\| u_1 - u_2 + \frac{\rho}{\nu} \right\|_2^2 - \frac{\nu}{2} \left\| \frac{\rho}{\nu} \right\|_2^2.\end{aligned}\tag{28}$$

3.3.3 Minimisation ADMM

Nous utilisons la méthode des directions alternées des multiplicateurs (ADMM en anglais). L'objectif est de minimiser \mathcal{L} en la considérant comme étant une fonction séparable de variables u_1, u_2 , et v , tout en ajustant itérativement les multiplicateurs de Lagrange λ_1, λ_2 , et ρ .

Au préalable nous introduisons un lemme nous permettant de simplifier les calculs :

Lemme : Pour $a, b_1, \dots, b_N \in \mathbb{R}$ et $x_1, \dots, x_N > 0$, on a

$$\sum_i x_i (a - b_i)^2 = \left(\sum_i x_i \right) \left(a - \frac{\sum_i b_i x_i}{\sum_i x_i} \right)^2 + C,\tag{29}$$

où $C \in \mathbb{R}$ est une constante qui ne dépend pas de a .

Preuve :

$$\begin{aligned}\sum_i x_i (a - b_i)^2 &= a^2 \left(\sum_i x_i \right) - 2a \left(\sum_i b_i x_i \right) + \sum_i b_i^2 x_i \\ &= \left(\sum_i x_i \right) \left(a^2 - 2a \frac{\sum_i b_i x_i}{\sum_i x_i} + \frac{\sum_i b_i^2 x_i}{\sum_i x_i} \right) \\ &= \left(\sum_i x_i \right) \left(a^2 - 2a \frac{\sum_i b_i x_i}{\sum_i x_i} + \left(\frac{\sum_i b_i x_i}{\sum_i x_i} \right)^2 - \left(\frac{\sum_i b_i x_i}{\sum_i x_i} \right)^2 + \frac{\sum_i b_i^2 x_i}{\sum_i x_i} \right)^{(30)} \\ &= \left(\sum_i x_i \right) \left(a - \frac{\sum_i b_i x_i}{\sum_i x_i} \right)^2 - \left(\frac{\sum_i b_i x_i}{\sum_i x_i} \right)^2 + \frac{\sum_i b_i^2 x_i}{\sum_i x_i}.\end{aligned}$$

Les deux derniers termes ne dépendent pas de a , ce qui démontre l'affirmation. \square

Ainsi le lemme permet de réarranger les termes quadratiques en une forme plus gérable qui sépare le terme linéaire du reste.

Pour illustrer ce réarrangement, considérons g la fonctionnelle séparée associées aux variables u_1 de la fonctionnelle \mathcal{L} c'est à dire :

$$u_1 \mapsto \gamma \left\| \nabla_{p_1} u_1 \right\|_0 + \frac{\mu}{2} \left\| v - u_1 + \frac{\lambda_1}{\mu} \right\|_2^2 + \left\| \nabla_{p_1} u_1 \right\|_0 + \frac{\nu}{2} \left\| u_1 - u_2 + \frac{\rho}{\nu} \right\|_2^2\tag{31}$$

on a que :

$$\begin{aligned}
\min_{u_1} g &\Leftrightarrow \min_{u_1} 2\gamma \|\nabla_{p_1} u_1\|_0 + \mu \left\| u_1 - \left(v + \frac{\lambda_1}{\mu} \right) \right\|_2^2 + \|\nabla_{p_1} u_1\|_0 + \nu \left\| u_1 - \left(u_2 - \frac{\rho}{\nu} \right) \right\|_2^2 \\
&\Leftrightarrow \min_{u_1} 2\gamma \|\nabla_{p_1} u_1\|_0 + (\mu + \nu) \left\| u_1 - \frac{\mu \left(v + \frac{\lambda_1}{\mu} \right) + \nu \left(u_2 - \frac{\rho}{\nu} \right)}{\mu + \nu} \right\|_2^2 \\
&\Leftrightarrow \min_{u_1} \frac{2}{\mu + \nu} \gamma \|\nabla_{p_1} u_1\|_0 + \left\| u_1 - \frac{1}{\mu + \nu} (\mu \nu + \nu u_2 + \lambda_1 + \rho) \right\|_2^2
\end{aligned} \tag{32}$$

De manière analogue nous pouvons obtenir les expressions des problèmes de minimisation des fonctionnelles séparées des variables u_2 et v .

Ce faisant, nous obtenons l'itération :

$$\begin{cases}
u_1^{k+1} \in \arg \min_{u_1} \frac{2\gamma}{\mu^k + \nu^k} \|\nabla_p u_1\|_0 + \left\| u_1 - \frac{1}{\mu^k + \nu^k} (\mu^k v^k + \nu^k u_2^k + \lambda_1^k - \rho^k) \right\|_2^2, \\
u_2^{k+1} \in \arg \min_{u_2} \frac{2\gamma}{\mu^k + \nu^k} \|\nabla_p u_2\|_0 + \left\| u_2 - \frac{1}{\mu^k + \nu^k} (\mu^k v^k + \nu^k u_1^{k+1} + \lambda_2^k + \rho^k) \right\|_2^2, \\
v^{k+1} = \arg \min_v \|Av - f\|_2^2 + \frac{\mu^k}{2} \left\| v - \frac{1}{2\mu^k} (\mu^k u_1^{k+1} + \mu^k u_2^{k+1} - \lambda_1^k - \lambda_2^k) \right\|_2^2, \\
\lambda_1^{k+1} = \lambda_1^k + \mu^k (v^{k+1} - u_1^{k+1}), \\
\lambda_2^{k+1} = \lambda_2^k + \mu^k (v^{k+1} - u_2^{k+1}), \\
\rho^{k+1} = \rho^k + \nu^k (u_1^{k+1} - u_2^{k+1}).
\end{cases} \tag{33}$$

Comme paramètre de couplage, nous utilisons une suite croissante $(\mu^k)_{k \in \mathbb{N}}$. Il s'agit d'un moyen permettant d'améliorer l'ADMM standard.

L'observation cruciale est que nous pouvons résoudre efficacement tous les sous-problèmes de l'Équation 12. La première ligne se décompose en n problèmes de Potts univariés de la forme

$$(u_1^{k+1})_{\cdot, j} \in \arg \min_{g \in \mathbb{R}^m} \frac{2\gamma}{\mu^k + \nu^k} \|\nabla g\|_0 + \left\| g - \frac{1}{\mu^k + \nu^k} (\mu^k v_{\cdot, j}^k + \nu^k (u_2)_{\cdot, j}^k + (\lambda_1)_{\cdot, j}^k - \rho_{\cdot, j}^k) \right\|_2^2 \tag{34}$$

où nous utilisons la notation par indice $x_{\cdot, j}$ pour désigner la j -ième ligne de l'image $m \times n$, c'est-à-dire $x_{\cdot, j} = (x_{ij})_{i=1, \dots, m}$. De manière analogue, nous obtenons une décomposition pour la deuxième ligne de l'Équation 33 en sous-problèmes

$$(u_2^{k+1})_{\cdot, j} \in \arg \min_{g \in \mathbb{R}^m} \frac{2\gamma}{\mu^k + \nu^k} \|\nabla g\|_0 + \left\| g - \frac{1}{\mu^k + \nu^k} (\mu^k v_{\cdot, j}^k + \nu^k (u_1)_{\cdot, j}^k + (\lambda_2)_{\cdot, j}^k - \rho_{\cdot, j}^k) \right\|_2^2 \tag{35}$$

Chaque sous-problème est conçu pour être traité une variable à la fois, réduisant ainsi la complexité du problème global. Les trois dernières lignes sont de simples étapes de montée de gradient dans les multiplicateurs de Lagrange.

3.4 Discrétisation isotropique

La discrétisation isotropique lors de l'utilisation du modèle de Potts, est essentielle pour plusieurs raisons techniques et pratiques telles que :

L'uniformité dans toutes les directions. La discrétisation isotropique vise à traiter toutes les directions de l'espace de manière égale. Cela est crucial pour éviter les biais directionnels dans

les reconstructions d'images, où certaines orientations pourraient autrement être privilégiées ou désavantagées. Par exemple, dans une discrétisation anisotrope standard, les directions horizontale et verticale sont souvent mieux résolues que les diagonales, ce qui peut conduire à des artefacts visuels tels que le staircasing (effet d'escalier).

L'approximation précise des distances euclidiennes. L'objectif d'une discrétisation isotrope est d'approximer plus fidèlement la véritable distance euclidienne entre les points. Cela est important non seulement pour l'aspect visuel mais aussi pour le calcul précis des caractéristiques géométriques dans les images, telles que les longueurs et les surfaces.

Initialement, le système de voisinage se compose uniquement des vecteurs $(1, 0)$ et $(0, 1)$, ce qui mène à des reconstructions anisotropes et favorise les discontinuités axiales. Pour atténuer ces effets, des vecteurs supplémentaires sont ajoutés, comme $(1, 1)$ pour les directions diagonales et par symétrie le vecteur orthogonal $(1, -1)$. Ainsi, nous obtenons la base de voisinages

$$\mathcal{N}_1 = \{(1, 0), (0, 1), (1, 1), (1, -1)\}. \quad (36)$$

Les vecteurs suivants à inclure dans le système de voisinage sont les quatre mouvements de cavalier, cela mène à la base de voisinages suivante :

$$\mathcal{N}_2 = \{(1, 0), (0, 1), (1, 1), (1, -1), (-2, -1), (-2, 1), (2, 1), (2, -1)\}. \quad (37)$$

Le schéma général d'ajout de nouveaux vecteurs correspond à l'énumération standard des nombres rationnels.

Le théorème de Thalès affirme que si deux droites parallèles coupent deux droites sécantes alors elles déterminent deux triangles dont les côtés correspondants ont des longueurs proportionnelles. Cela revient à dire que les triangles formés sont semblables.

Pour appliquer le théorème de Thalès, imaginons une situation où nous avons une ligne de pente y/x traversant un carré de côté n (la taille de l'image binaire). Pour simplifier, nous pouvons supposer que la ligne commence à l'origine $(0, 0)$ et se termine à un point $(n, n \cdot y/x)$.

Pour appliquer le théorème de Thalès, nous allons imaginer que cette ligne forme un triangle rectangle avec les axes x et y de l'image.

1. Triangle formé :

- Considérons un triangle rectangle où la ligne traverse de $(0, 0)$ à (n, y') , où $y' = n \cdot \frac{y}{x}$.
- Les côtés de ce triangle sont n sur l'axe des x et $y' = n \cdot \frac{y}{x}$ sur l'axe des y .

2. Application du théorème de Thalès :

- Nous considérons un petit segment de cette ligne qui forme un sous-triangle semblable au triangle entier.
- Le théorème de Thalès nous dit que les segments correspondants dans les triangles semblables sont proportionnels.

3. Calcul de la longueur euclidienne :

- La longueur de l'hypoténuse du triangle rectangle original est :

$$\begin{aligned}
L_E &= \sqrt{n^2 + \left(n \cdot \frac{y}{x}\right)^2} \\
L_E &= \sqrt{n^2 \left(1 + \frac{y^2}{x^2}\right)} \\
L_E &= n \cdot \frac{\sqrt{x^2 + y^2}}{x}
\end{aligned} \tag{38}$$

Cette approximation est essentielle pour établir les poids appropriés dans le système de voisinage afin d'assurer que la somme des sauts pondérés soit égale à la longueur euclidienne réelle.

Les poids appropriés peuvent être déterminés comme suit :

$$\sum_{s=1}^S \omega_s \|\nabla_{p_s} u\|_0 = n \frac{\sqrt{x^2 + y^2}}{x}. \tag{39}$$

où $\|\nabla_{p_s} u\|_0$ est donné par

$$\|\nabla u\|_0 = |\{(i, j) : u(i, j) + p_s \neq u(i, j)\}| \tag{40}$$

Il reste à évaluer le côté gauche de l'équation pour l'image binaire u . Cela peut être fait soit manuellement pour de petits systèmes de voisinage, soit avec l'aide d'un programme informatique pour de plus grands systèmes de voisinage. Lorsque nous comptons les entrées non nulles de ∇u , nous supposons n grand afin que les effets de bord soient négligeables. Nous obtenons un système de S équations pour les S inconnues. Pour le système de voisinage diagonal \mathcal{N} , (11) donne les conditions

$$\begin{aligned}
\omega_1 + \omega_3 + \omega_4 &= 1, \\
\omega_2 + \omega_3 + \omega_4 &= 1, \\
\omega_1 + \omega_2 + 2\omega_3 &= \sqrt{2}, \\
\omega_1 + \omega_2 + 2\omega_4 &= \sqrt{2}.
\end{aligned} \tag{41}$$

En résolvant ce système linéaire, nous obtenons les poids

$$\omega_1 = \omega_2 = \sqrt{2} - 1 \text{ et } \omega_3 = \omega_4 = \frac{1 - \sqrt{2}}{2}. \tag{42}$$

Pour le système de voisinage \mathcal{N}_2 , nous obtenons un système analogue d'équations dans $S = 8$ inconnues qui nous donne les poids

$$\omega_s = \begin{cases} \frac{\sqrt{5}-2}{2}, & \text{pour } s = 1, 2 \\ \frac{\sqrt{5}-\sqrt{2}}{2}, & \text{pour } s = 3, 4 \\ \frac{1}{2}(1 + \sqrt{2} - \sqrt{5}), & \text{pour } s = 5, \dots, 8. \end{cases} \tag{43}$$

Nous nous tournons maintenant vers la question de savoir comment nous approximations la longueur euclidienne avec les discrétisations ci-dessus. Les systèmes de voisinage donnent naissance à une norme $\|\cdot\|_{\mathcal{N}}$ définie pour $p \in \mathbb{R}^2$ par

$$\|p\|_{\mathcal{N}} = \sum_{s=1}^S \omega_s |\langle p, p_s \rangle|. \tag{44}$$

Cette norme est conçue pour que la longueur mesurée $\|p_s\|_{\mathcal{N}}$ coïncide exactement avec la longueur euclidienne $\|p_s\|_2$ pour tous les vecteurs p_s utilisés dans la base de voisinages. Cela permet de s'assurer que la discrétisation n'introduit pas de biais anisotrope significatif dans la mesure des distances. Pour mesurer l'isotropie d'un système de différences finies, on utilise le rapport E entre le plus long et le plus court vecteur unitaire par rapport à cette longueur :

$$E = \frac{\max_{\|p\|_2=1} \|p\|_{\mathcal{N}}}{\min_{\|p\|_2=1} \|p\|_{\mathcal{N}}} \quad (45)$$

Plus la valeur de E est proche de 1, plus le système est isotrope. Pour les systèmes \mathcal{N}_0 , \mathcal{N}_1 et \mathcal{N}_2 , les valeurs de E sont respectivement 1.41, 1.08 et 1.03. Cela montre une amélioration significative de l'isotropie en ajoutant des directions diagonales et du coup du cavalier.

3.5 Discrétisation générale

Nous déterminons maintenant une stratégie de minimisation pour la discrétisation générale de l'Équation 10. Notons le système de voisinage par $\mathcal{N} = \{p_1, \dots, p_S\}$ et supposons que $\omega_1, \dots, \omega_S > 0$ où $S \geq 2$. Nous réécrivons d'abord l'Équation 10 comme le problème d'optimisation sous contrainte suivant :

$$\begin{aligned} \min_{u_1, \dots, u_S, \nu} \quad & \gamma \sum_{s=1}^S \omega_s \|\nabla_{p_s} u_s\|_0 + \|Av - f\|_2^2, \\ u_r - u_t = 0, \quad & \text{pour tout } 1 \leq r < t \leq S, \\ v_s - u_s = 0, \quad & \text{pour tout } 1 \leq s \leq S. \end{aligned} \quad (46)$$

Le Lagrangien augmenté de ce problème d'optimisation s'écrit :

$$\begin{aligned} \mathcal{L}(u, \lambda, \rho) = \quad & \gamma \sum_{s=1}^S \omega_s \|\nabla_{p_s} u_s\|_0 + \|Av - f\|_2^2 + \frac{\mu}{2} \left\| v - u_s + \frac{\lambda_s}{\mu} \right\|_2^2 \\ & + \frac{\nu}{2} \sum_{1 \leq r < t \leq S} \left\| u_r - u_t + \frac{\rho_{r,t}}{\nu} \right\|_2^2. \end{aligned} \quad (47)$$

où \mathcal{L} dépend des variables $\{u_s\}_{1 \leq s \leq S}$, $\{\lambda_s\}_{1 \leq s \leq S}$ et $\{\rho_{r,t}\}_{1 \leq r < t \leq S}$. Le paramètre $\nu > 0$ contrôle la des variables séparées u_1, \dots, u_S , et $\mu > 0$ leurs couplage à v . Les variables $\lambda_s, \rho_{r,t} \in \mathbb{R}^{n \times m}$ sont les multiplicateurs de Lagrange. Dans l'itération ADMM, nous minimisons \mathcal{L} séquentiellement par rapport à v, u_1, \dots, u_S , suivie d'une étape d'ascension par gradient dans les multiplicateurs de Lagrange.

La minimisation de \mathcal{L} par rapport à u_s s'écrit :

$$\begin{aligned} \arg \min_{u_s} \mathcal{L}(u, \lambda, \rho) = \quad & \arg \min_{u_s} \gamma \omega_s \|\nabla_{p_s} u_s\|_0 + \frac{\mu}{2} \left\| v - u_s + \frac{\lambda_s}{\mu} \right\|_2^2 \\ & + \frac{\nu}{2} \sum_{1 \leq r < S} \left\| u_r - u_s + \frac{\rho_{r,s}}{\nu} \right\|_2^2 + \frac{\nu}{2} \sum_{s < t \leq S} \left\| u_s - u_t + \frac{\rho_{s,t}}{\nu} \right\|_2^2. \end{aligned} \quad (48)$$

Nous pouvons simplifier l'expression de la fonctionnelle \mathcal{L} grace au lemme :

$$\begin{aligned}
& \gamma\omega_s \left\| \nabla_{p_s} u_s \right\|_0 + \frac{\mu}{2} \left\| v - u_s + \frac{\lambda_s}{\mu} \right\|_2^2 \\
& + \frac{\nu}{2} \sum_{1 \leq r < S} \left\| u_r - u_s + \frac{\rho_{r,s}}{\nu} \right\|_2^2 + \frac{\nu}{2} \sum_{s < t \leq S} \left\| u_s - u_t + \frac{\rho_{s,t}}{\nu} \right\|_2^2 \\
& \Leftrightarrow 2\gamma\omega_s \left\| \nabla_{p_s} u_s \right\|_0 + \mu \left\| u_s - \left(v + \frac{\lambda_s}{\mu} \right) \right\|_2^2 \\
& + \nu \sum_{1 \leq r < S} \left\| u_s - \left(u_r + \frac{\rho_{r,s}}{\nu} \right) \right\|_2^2 + \nu \sum_{s < t \leq S} \left\| u_s - \left(u_t - \frac{\rho_{s,t}}{\nu} \right) \right\|_2^2 \\
& \Leftrightarrow 2\gamma\omega_s \left\| \nabla_{p_s} u_s \right\|_0 \\
& + (\mu + s\nu + (S - (s + 1))\nu) \left\| u_s - \frac{\mu\left(\nu + \frac{\lambda_s}{\mu}\right) + \nu\left(\sum_{1 \leq r \leq s} u_r + \frac{\rho_{r,s}}{\nu}\right) + \nu \sum_{s \leq t \leq S} u_t - \frac{\rho_{s,t}}{\nu}}{\mu + \nu(S - 1)} \right\|_2^2 \\
& \Leftrightarrow \frac{2\gamma\omega_s}{\mu + \nu(S - 1)} \left\| \nabla_{p_s} u_s \right\|_0 + \|u_s - w_s\|_2^2
\end{aligned} \tag{49}$$

avec

$$w_s = \frac{\mu\nu + \lambda_s + \sum_{1 \leq r < s} (\nu u_r + \rho_{r,s}) + \sum_{s \leq t \leq S} (\nu u_t - \rho_{s,t})}{\mu + \nu(S - 1)} \tag{50}$$

Ainsi l'expression devient :

$$\arg \min_{u_s} \mathcal{L}(u, \lambda, \rho) = \arg \min_{u_s} \frac{2\gamma\omega_s}{\mu + \nu(S - 1)} \left\| \nabla_{p_s} u_s \right\|_0 + \|u_s - w_s\|_2^2 \tag{51}$$

De manière analogue en considérant le lagrangien augmenté selon la variable v uniquement, et en utilisant encore le lemme, on a :

$$\arg \min_v \mathcal{L}(u, \lambda, \rho) = \arg \min_v \|Av - f\|_2^2 + \frac{\mu S}{2} \left\| v - \frac{1}{S} \sum_{s=1}^S \left(u_s - \frac{\lambda_s}{\mu} \right) \right\|_2^2 \tag{52}$$

Finalement en ayant implémenté les minimiseurs pour chaque variable, on obtient par l'ADMM les itérations suivantes :

$$\begin{cases}
u_1^{k+1} \in \arg \min_{u_1} \frac{2\gamma\omega_1}{\mu^k + \nu^k(S-1)} \left\| \nabla_{p_1} u_1 \right\|_0 + \|u_1 - w_1^k\|_2^2, \\
\vdots \\
u_s^{k+1} \in \arg \min_{u_s} \frac{2\gamma\omega_s}{\mu^k + \nu^k(S-1)} \left\| \nabla_{p_s} u_s \right\|_0 + \|u_s - w_s^k\|_2^2, \\
v^{k+1} = \arg \min_v \|Av - f\|_2^2 + \frac{\mu^k S}{2} \|v - z^k\|_2^2, \\
\lambda_s^{k+1} = \lambda_s^k + \mu^k (v^{k+1} - u_s^{k+1}), \text{ pour tout } 1 \leq s \leq S \\
\rho_{r,t}^{k+1} = \rho_{r,t}^k + \nu^k (u_r^{k+1} - u_t^{k+1}), \text{ pour tout } 1 \leq r < t \leq S.
\end{cases} \tag{53}$$

avec w_s^k donné par :

$$w_s^k = \frac{\mu^k \nu^k + \lambda_s^k + \sum_{1 \leq r < s} (\nu^k u_r^{k+1} + \rho_{r,s}^k) + \sum_{s \leq t \leq S} (\nu^k u_t^k - \rho_{s,t}^k)}{\mu^k + \nu^k(S - 1)} \tag{54}$$

et z^k par

$$z^k = \frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} \right) \quad (55)$$

L'observation clé est que chaque sous-problème dans l'itération ADMM peut être résolu efficacement. Les problèmes de minimisation par rapport à u_1, \dots, u_s se décomposent en problèmes de Potts univariés par rapport aux chemins induits par les vecteurs de différences finies p_s . L'itération sur v^{k+1} est résolue comme dans le cas anisotrope, il s'agit d'une régularisation classique de type Tikhonov. Enfin nous notons que la discrétisation anisotrope est un cas est un cas particulier de la discrétisation isotrope si nous choisissons de faire les différences finie suivant la base de voisinage \mathcal{N}_0 .

4 Solutions des sous-problèmes

Les méthodes de résolution pour les problèmes de Potts univariés et de régularisation classique de Tikhonov sont bien documentées. Étant donné qu'elles constituent les éléments essentiels de notre algorithme Équation 53, nous allons brièvement rappeler le principe de ces algorithmes.

4.1 Fonctionnelle de Potts univariée

La fonctionnelle univariée de Potts s'énonce comme suit :

$$\mathcal{P}_\gamma(g) = \gamma \|\nabla g\|_0 + \|g - f\|_2^2 \rightarrow \min, \quad (56)$$

On a que $g, f \in \mathbb{R}^n$ et le terme de saut est donné par $\|\nabla g\|_0 = |\{i : g_i \neq g_{i+1}\}|$

La programmation dynamique étant une méthode de résolution de problèmes complexes en les décomposant en sous-problèmes plus simples, en résolvant chacun de ces sous-problèmes une seule fois et en stockant leurs solutions, elle s'illustre assez bien dans notre cas.

En effet, vu que nous voulons minimiser $\mathcal{P}_\gamma(\cdot)$, le terme de fidélité $\|g - f\|_2^2$ pour une donnée $f \in \mathbb{R}^n$, nous allons d'abord essayer de minimiser $\mathcal{P}_\gamma(\cdot)$ pour des données partielles $f_r = (f_1, \dots, f_r)$, avec $r \leq n$. Pour ce faire nous devons avoir à notre disposition les minimiseurs g^1, \dots, g^{r-1} associées aux données partielles $(f_1), \dots, (f_1, \dots, f_{r-1})$ respectivement.

Les étapes suivantes nous permettent alors d'obtenir un minimiseur :

1. Création de r minimiseurs candidats : h^1, \dots, h^r , chacun de longueur r données par :

$$h^\ell = (g^{\ell-1}, \mu_{[\ell, r]}, \dots, \mu_{[\ell, r]}) \quad (57)$$

avec $\mu_{[\ell, r]}$ qui correspond à la moyenne des valeurs des données partielles

$$f_{[\ell, r]} = (f_\ell, \dots, f_r) \quad (58)$$

2. On choisit parmi les r candidats, celui qui minimise la fonctionnelle de Potts associées aux données $f_{[1, r]}$

Par ailleurs Friedrich F, Kempe A, Liebscher V et Winkler G, proposent dans leurs travaux « Complexity penalized M-estimation » de 2008, une résolution analogue à celle introduite ici, une remarque importante est que les minimiseurs des fonctionnelles univariées de Potts ne sont pas forcément unique, par contre les données associés amenant à des solutions non-unique forment un ensemble négligeable.

4.2 Cas d'application de régularisation de Tikhonov

Ensuite la solution du sous-problème :

$$v^{k+1} = \arg \min_v \|Av - f\|_2^2 + \frac{\mu^k S}{2} \|v - z^k\|_2^2, \quad (59)$$

est donnée par une résolution classique de régularisation de Tikhonov. Ainsi en posant $\mu^k S = \mu$ et en définissant h la fonctionnelle :

$$v \mapsto \|Av - f\|_2^2 + \frac{\mu}{2} \|v - z^k\|_2^2 \quad (60)$$

On peut montrer que h est une fonctionnelle quadratique, en effet :

$$\begin{aligned} h(v) &= \langle Av - f, Av - f \rangle + \frac{\mu}{2} \langle v - z, v - z \rangle \\ &= \langle Av, Av \rangle - 2\langle Av, f \rangle + \langle f, f \rangle + \frac{\mu}{2} (\langle v, v \rangle - 2\langle v, z \rangle + \langle z, z \rangle) \\ &= \langle A^* Av, v \rangle - 2\langle A^* f, v \rangle + \|f\|_2^2 + \frac{\mu}{2} (\|v\|_2^2 - 2\langle v, z \rangle + \|z\|_2^2) \\ &= \left\langle \left(A^* A + \frac{\mu}{2} I \right) v, v \right\rangle - 2 \left\langle A^* f + \frac{\mu z}{2}, v \right\rangle + \|f\|_2^2 + \frac{\mu}{2} \|z\|_2^2 \end{aligned} \quad (61)$$

Ainsi le sous-problème admet une unique solution, qui vérifie les équations normales :

$$\left(A^* A + \frac{\mu}{2} I \right) v = A^* f + \frac{\mu}{2} z. \quad (62)$$

avec $z \in \mathbb{R}^{m \times n}$, et A^* l'opérateur adjoint de A .

5 Convergence

Dans cette partie nous voulons montrer que dans le cas typique où $\nu^k = 0$ pour tout k , (ce qui implique $\rho^k = 0$ pour tout k), on a la convergence de l'algorithme Équation 53 (nous n'aborderons pas le cas général ici). Pour cela nous allons énoncer le théorème suivant :

Théorème

Soit $(\mu^k)_{k \in \mathbb{N}}$ une suite croissante telle que $\sum_k \mu_k^{-1/2} < +\infty$, de plus on suppose que $\nu^k = 0$ pour tout k , alors l'algorithme Équation 53 converge au sens :

$$(u_1^k, \dots, u_S^k, v^k) \rightarrow (u_1^*, \dots, u_S^*, v^*), \text{ avec } u_1^* = \dots = u_S^* = v^*, \quad (63)$$

$$\frac{\lambda_s^k}{\mu^k} \rightarrow 0 \text{ pour tout } s \in \{1, \dots, S\} \quad (64)$$

Preuve

Tout d'abord montrons qu'on peut reformuler les S premières équations de Équation 53 grace à la fonctionnelle F_s^k :

$$F_s^k(u_s) = \frac{2\gamma\omega_s}{\mu^k} \|\nabla_{p_s} u_s\|_0 + \left\| u_s - \left(v^k + \frac{\lambda_s^k}{\mu^k} \right) \right\|_2^2. \quad (65)$$

En effet comme $\nu^k = 0$ pour tout k (et $\rho^k = 0$ pour tout k), le terme w_s^k dans Équation 53 est donné par :

$$w_s^k = \frac{\mu^k v^k + \lambda_s^k}{\mu^k} = v^k + \frac{\lambda_s^k}{\mu^k} \quad (66)$$

Ainsi on a bien que les S premières ligne de Équation 53 coïncident avec le problème :

$$u_s^{k+1} \in \arg \min_u F_s^k(u_s) \quad (67)$$

Ensuite nous essayons d'estimer la distance $\left\| u_s - \left(v^k + \frac{\lambda_s^k}{\mu^k} \right) \right\|_2$. Pour cela on note que $F_s^k(u_s^{k+1}) \leq F_s^k\left(v^k + \frac{\lambda_s^k}{\mu^k}\right)$ ce qui est vraie par hypothèse car u_s^{k+1} minimise F_s^k . Ainsi on a :

$$\begin{aligned} \frac{\mu^k}{2} F_s^k(u_s^{k+1}) &= \gamma \omega_s \left\| \nabla_{p_s} u_s^{k+1} \right\|_0 + \frac{\mu^k}{2} \left\| u_s^{k+1} - \left(v^k + \frac{\lambda_s^k}{\mu^k} \right) \right\|_2^2 \\ &\leq \frac{\mu^k}{2} F_s^k\left(v^k + \frac{\lambda_s^k}{\mu^k}\right) = \gamma \omega_s \left\| \nabla_{p_s} \left(v^k + \frac{\lambda_s^k}{\mu^k} \right) \right\|_0 \leq \gamma \omega_s L \end{aligned} \quad (68)$$

où $L = NM$ correspondant au dimensions($N \times M$) de l'image. Ceci est vraie car $\left\| \nabla_{p_s} z \right\|_0 \leq MN$ pour tout p_s et pour toute donnée z . En effet en posant $s = (s_1, s_2) \in \mathcal{N}^S$ avec $S \geq 2$ on a :

$$\left\| \nabla_{p_s} z \right\|_0 = |\{(i, j) : z_{i,j} \neq z_{i+s_1, j+s_2}\}| \leq N |\{j : z_{i,j} \neq z_{i+s_1, j+s_2}\}| \leq NM = L \quad (69)$$

Ensuite comme la norme $\|\cdot\|_0$ est positive, on obtient bien que :

$$\left\| u_s^{k+1} - \left(v^k + \frac{\lambda_s^k}{\mu^k} \right) \right\|_2^2 \leq \frac{2\gamma \omega_s L}{\mu^k}. \quad (70)$$

En particulier, pour tout $s \in \{1, \dots, S\}$, on a :

$$\lim_{k \rightarrow \infty} u_s^{k+1} - \left(v^k + \frac{\lambda_s^k}{\mu^k} \right) = 0. \quad (71)$$

En effet comme $(\mu^k)_{k \in \mathbb{N}}$ est une suite croissante telle que $\sum_k \mu_k^{-1/2} < +\infty$, le terme général de la serie $\sum_k \mu_k^{-1/2}$ converge vers 0, et donc $\mu_k^{-1/2} \rightarrow 0 \iff \mu_k^{1/2} \rightarrow +\infty$ ce qui conclue le fait que $u_s^{k+1} - \left(v^k + \frac{\lambda_s^k}{\mu^k} \right)$ vers 0 dans ℓ^2 .

Maintenant nous allons porter notre attention sur la $(S+1)$ -ème ligne de Équation 53. On définit la fonctionnelle correspondante par G^k :

$$G^k(v) = \|Av - f\|_2^2 + \frac{\mu^k S}{2} \left\| v - \frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} \right) \right\|_2^2. \quad (72)$$

Comme v^{k+1} minimise G^k on a l'inégalité suivante :

$$G^k(v^{k+1}) \leq G^k\left(\frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} \right)\right). \quad (73)$$

Par définition de G^k on a :

$$\begin{aligned}
(b) &= \|Av^{k+1} - f\|_2^2 + \frac{\mu^k S}{2} \left\| v^{k+1} - \left(\frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} \right) \right) \right\|_2^2 \\
&\leq \left\| A \left(\frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} \right) \right) - f \right\|_2^2 \\
&= \left\| A \left(\frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} - v^k + v^k \right) \right) - f \right\|_2^2 \\
&= \left\| A \left(\frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} - v^k \right) \right) + Av^k - f \right\|_2^2
\end{aligned} \tag{74}$$

puis par inégalité triangulaire et par homogénéité de la norme, on a

$$\leq \left(\|A\| \left\| \frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} - v^k \right) \right\|_2 + \|Av^k - f\|_2 \right)^2 \tag{75}$$

Or en se servant de Équation 70, on a :

$$\begin{aligned}
\|A\| \left\| \frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} - v^k \right) \right\|_2 &\leq \|A\| \frac{1}{S} \sum_{s=1}^S \left\| u_s^{k+1} - \left(v^k + \frac{\lambda_s^k}{\mu^k} \right) \right\|_2 \\
&\leq \|A\| \sqrt{\frac{2\gamma\omega_s L}{\mu^k}}
\end{aligned} \tag{76}$$

Ainsi en posant $C = \|A\| \sqrt{2\gamma\omega_s L}$, à partir de Équation 75 on se ramène à :

$$\|Av^{k+1} - f\|_2 \leq \frac{C}{\sqrt{\mu^k}} + \|Av^k - f\|_2, \tag{77}$$

Ici, $\|A\|$ est la norme de l'opérateur A dans ℓ^2 , $C > 0$ dépendant uniquement de γ, ω_s, L , et $\|A\|$. Par récursivité nous avons

$$\|Av^{k+1} - f\|_2 \leq C \sum_{j=1}^k \frac{1}{\sqrt{\mu^j}} + \|Av^0 - f\|_2, \tag{78}$$

ce qui montre que les itérations $\{Av^{k+1} - f\}_{k \in \mathbb{N}}$ sont bornées.

Par ailleurs Équation 75 nous donne également l'inégalité suivante :

$$\begin{aligned}
\frac{\mu^k S}{2} \left\| v^{k+1} - \frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} \right) \right\|_2^2 &\leq \left(\|A\| \left\| \frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} - v^k \right) \right\|_2 + \|Av^k - f\|_2 \right)^2 \\
&\leq \left(\|A\| \frac{1}{S} \sum_{s=1}^S \left\| u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} - v^k \right\|_2 + \|Av^k - f\|_2 \right)^2 \\
&\leq \left(\|A\| \sqrt{\frac{2\gamma\omega_s L}{\mu^k}} + C \sum_{j=1}^k \frac{1}{\sqrt{\mu^j}} + \|Av^0 - f\|_2 \right)^2
\end{aligned} \tag{79}$$

Ainsi on a :

$$\left\| v^{k+1} - \frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} \right) \right\|_2 \leq \frac{2}{\mu^k + S} \left(\|A\| \sqrt{\frac{2\gamma\omega_s L}{\mu^k}} + C \sum_{j=1}^k \frac{1}{\sqrt{\mu^j}} + \|Av^0 - f\|_2 \right) \quad (80)$$

Cela nous permet de montrer la convergence de v^k en montrant que la suite de cauchy : $(v^{k+1} - v^k)$ converge.

On a alors :

$$\begin{aligned} \|v^{k+1} - v^k\| &= \left\| v^{k+1} - v^k - \frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} \right) + \frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} \right) \right\|_2 \\ &\leq \left\| v^{k+1} - \frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} \right) \right\|_2 + \left\| \frac{1}{S} \sum_{s=1}^S \left(u_s^{k+1} - \frac{\lambda_s^k}{\mu^k} - v^k \right) \right\|_2 \\ &\leq \frac{2}{\mu^k + S} \left(\|A\| \sqrt{\frac{2\gamma\omega_s L}{\mu^k}} + C \sum_{j=1}^k \frac{1}{\sqrt{\mu^j}} + \|Av^0 - f\|_2 \right) + \sqrt{\frac{\gamma\omega_s L}{\mu^k}} \\ &= K \end{aligned} \quad (81)$$

Or $K \rightarrow 0$ lorsque $k \rightarrow +\infty$, ceci est une conséquence directe du fait que $\mu^k \rightarrow +\infty$. Par conséquent on la convergence de v^k vers un certain v^* .

Ensuite pour établir que $\frac{\lambda_s^k}{\mu^k} \rightarrow 0$ pour tout $s \in \{1, \dots, S\}$ nous allons reformuler la ligne

$$\lambda_s^{k+1} = \lambda_s^k + \mu^k (v^{k+1} - u_s^{k+1}), \text{ pour tout } 1 \leq s \leq S \quad (82)$$

de l'algorithme Équation 53

Ainsi on aura :

$$\begin{aligned} \frac{\lambda_s^{k+1}}{\mu^k} &= \frac{\lambda_s^k}{\mu^k} - u_s^{k+1} + v^{k+1} \\ -\frac{\lambda_s^{k+1}}{\mu^k} &= \left(u_s^{k+1} - \left(\frac{\lambda_s^k}{\mu^k} + v^k \right) \right) + (v^k - v^{k+1}) \end{aligned} \quad (83)$$

Ainsi on a bien

$$\lim_{k \rightarrow \infty} \frac{\lambda_s^{k+1}}{\mu^k} = 0. \quad (84)$$

Car les deux terme de la somme convergent vers 0. Vu que la suite μ^k est croissante, on a que $\mu^k / \mu^{k+1} \leq 1$ et donc, pour tout $s = 1, \dots, S$

$$\lim_{k \rightarrow \infty} \frac{\lambda_s^k}{\mu^k} = 0. \quad (85)$$

Enfin nous reformulons la dernière ligne de l'algorithme Équation 53 comme suit :

$$v^{k+1} - u_s^{k+1} = (\lambda_s^{k+1} - \lambda_s^k) / \mu^k \quad (86)$$

et on obtient l'inégalité

$$\|u_s^{k+1} - v^{k+1}\| \leq \frac{\|\lambda_s^{k+1}\|}{\mu^k} + \frac{\|\lambda_s^k\|}{\mu^k} \rightarrow 0. \quad (87)$$

cela est équivalent à $u_s^k - v^k \rightarrow 0$ pour tout $s = 1, \dots, S$ et, comme v^k converge, alors chaque u_s^k converge aussi et sa limite correspondante est $u_s^* = v^*$, ce qui complète la preuve. \square

6 Conclusion

Ce projet a permis de réaliser une étude complète du modèle de Potts, depuis ses fondements théoriques jusqu'à l'implémentation d'un algorithme de résolution avancé basé sur l'ADMM. Nous avons établi le cadre mathématique des problèmes inverses mal-posés, justifié la nécessité de la régularisation par une approche variationnelle et bayésienne, et développé une discrétisation isotropique du modèle de Potts sur grille régulière.

Nos expériences ont validé la supériorité du modèle de Potts sur les filtres classiques pour le débruitage d'images présentant des régions homogènes, et démontré sa robustesse face à des niveaux de bruit élevés. Nous avons également mis en lumière le compromis entre la complexité théorique de l'ADMM et son efficacité pratique pour des images de taille raisonnable.

L'algorithme ADMM proposé par Storath et al. offre un cadre rigoureux et performant pour la résolution du modèle de Potts. Bien que le problème soit non-convexe, les garanties de convergence vers un point stationnaire et la qualité empirique des solutions en font une approche de choix pour les applications pratiques.

Les travaux futurs pourraient s'orienter vers plusieurs directions prometteuses. L'implémentation de solveurs 1D en temps quasi-linéaire ($O(n \log n)$) pour le problème de Potts univarié permettrait d'accélérer significativement le traitement des images de grande taille, ouvrant la voie à des applications en temps réel. L'extension de l'algorithme à d'autres problèmes inverses, comme la déconvolution avec noyau de flou connu ou la reconstruction tomographique complète avec opérateur de projection non trivial, constitue également une perspective naturelle de ce travail. Enfin, l'étude de variantes du modèle de Potts intégrant des priors plus sophistiqués (Potts couleur, Potts multicanal, couplage avec des modèles de texture) pourrait améliorer les performances sur des images réelles complexes.

7 Références

- [1] Wikipedia
- [2] Juan Peypouquet : Convex Optimization in Normed Spaces (2015)
- [3] Sur quelques problèmes inverses en traitement d'image par Laure BLANC-FERAUD
- [4] Lukas Kiefer, Martin Storath et Andreas Weinmann : Iterative Potts Minimization for the Recovery of Signals with Discontinuities from Indirect Measurements : The Multivariate Case
- [5] Pierre CHARBONNIER : Reconstruction d'image : régularisation avec prise en compte des discontinuités.
- [6] Jean-Baptiste Hiriart-Urruty : Optimisation et analyse convexe (1998)