Quelques applications des contraintes l^∞ en traitement d'images

Pierre Weiss, Laure Blanc-Féraud, Gilles Aubert

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Some applications of l^∞-constraints in image processing

Pierre Weiss — Gilles Aubert — Laure Blanc-Féraud

N° ????

November 2006

Thème COG

ISSN 0249-6399 ISRN INRIA/RR--????--FR+ENG

Some applications of l^∞-constraints in image processing

Pierre Weiss , Gilles Aubert , Laure Blanc-Féraud

Thème COG — Systèmes cognitifs
Projets Ariana

Rapport de recherche n° ???? — November 2006 — 33 pages

Abstract: Our goal in this paper is to give algorithms for minimizing generic regularizing functionals under a l^∞-constraint. We show that many classical models using total variation can be stated under this formalism. Among others are the Rudin-Osher-Fatemi model, the $BV - l^1$ model, the $BV - l^\infty$ model and Meyer's cartoon+texture decomposition model. Then we describe a general convergent algorithm to solve such problems. This algorithm is the projected subgradient descent. We finally give numerical results that show the qualities and limits of our models, and we tackle the question of the use of the total variation to treat bounded noises such as quantization noise.

Key-words: l^∞-norm, total variation minimization, bounded noises, duality, quantization noise, compression noise, projected subgradient descent.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Quelques applications des contraintes l^∞ en traitement d'images

Résumé : Dans ce rapport de recherche, nous présentons plusieurs modèles de restauration ou de décomposition d'images qui ont la particularité de faire apparaître une contrainte l^∞. Nous montrons que plusieurs modèles classiques utilisant la variation totale peuvent s'exprimer sous cette forme. Parmi d'autres, nous retrouvons le modèle de Rudin-Osher-Fatemi, les modèles $BV - L^1$ et $BV - L^\infty$, ainsi que le modèle de décomposition d'une image en géométrie et texture de Y. Meyer. Nous minimisons une fonctionnelle convexe de régularisation, sous contraintes l^∞. Nous montrons qu'une descente de sous-gradient projeté permet de résoudre ce problème de manière convergente. Nous donnons finalement plusieurs résultats numériques qui montrent quelques qualités et limites des modèles étudiés. Nous finissons par montrer que la variation totale n'est pas toujours la fonctionnelle la plus adaptée à la régularisation d'images entâchées de bruits bornés (bruits de compression), et que des fonctionnelles lissantes peuvent donner des résultats plus satisfaisants.

Mots-clés : norme l^∞, minimisation de la variation totale, bruits bornés, bruits de compression, bruits de quantification, dualité, descente de sous-gradient projeté

Contents

1 Introduction

In this paper we are interested in a class of image restoration and decomposition models that can be written under the general form :

$$\inf_{u \in K} J(u) \tag{1.1}$$

with K defined as :

$$K = \{u \in Z, ||u - f||_\infty \leq \alpha\} \tag{1.2}$$

Here f is a given observed data. Z can represent any of the two discrete spaces : $X = \mathbb{R}^n$ or $Y = X \times X$. X is the discrete space of images (in which case $n = n_x n_y$ is the number of pixels). Y is a discrete space of $2 - D$ vector fields (typically gradients). Finally, $J : Z \to \mathbb{R}$ is a convex function.

In the vast litterature devoted to denoising techniques by optimization, few models incorporate l^∞-norms. One of the reasons is that the l^∞-norm is not differentiable, and thus difficult to handle both numerically and theoretically. We will show throughout this paper, that our formulation actually covers a wide range of useful applications. Let us give a brief overview.

First, when dealing with noises of bounded amplitude, the Maximum a Posteriori (MAP) principle naturally leads to optimization problems with l^∞-constraints. Such noises frequently appear after compression. For instance intensity or wavelet coefficient quantization lead to such bounded noises. This will be the topic of section (7).

The model (1.1) includes another class of applications of great interest in image processing. Many models use the total variation as a prior on the images. This quantity somehow measures the oscillations of an image. It was introduced by Rudin, Osher and Fatemi (ROF) in [26] as a regularizing criterion for image denoising. Its main interest lies in the fact that it regularizes the images without blurring the edges. Using the theory of duality [13], we can show that minimizing the total variation under a l^p-constraint, can be reformulated under form (1.1). This point is detailled in section (6).

Y. Meyer in [18] inspiring from the model of Rudin, Osher and Fatemi [26], proposed a new model to decompose the images into piecewise smooth functions (the cartoon of the images) and oscillating functions (the texture or noise). We will show in section (5) that this model can also be reformulated under form (1.1).

Solving problem (1.1) numerically is not an easy task. The domain K has non smooth boundary and we make no assumptions of differentiability on J. The projected subgradient descent method [25] is a powerful tool to solve such generic problems. We review its properties of convergence and optimality and show how to apply it to the above mentionned problems.

Many of the ideas proposed in this paper are not new. Though we claim that some original contributions are present in this paper. Using a projected subgradient descent to optimize nondifferentiable functionals of image processing was notably proposed in [16, 17, 2]. Our main contribution here might be an easily implementable definition of the discrete subgradients, and a realistic review of the qualities and limits of this method.

A. Chambolle in paper [7], solved the problem of ROF using duality. His numerical method is shown to be convergent, and works in almost real time [8]. Since then some efforts were made to use duality in more general problems. Among other works, let us cite the paper by J. Bect and al. [6] where the authors generalize the algorithm of A. Chambolle for a wide class of l^1-problems including deconvolution, or best basis pursuit. Another generalization was given by T.F. Chan et. al. in [9], were the authors use duality to solve some problems involving second order regularizing functionals. In both works though, the authors are constrained to use l^2-data fidelity terms. Our contribution here, is that we show how to apply duality for any l^p-norm.

Recently some algorithms using total variation were proposed to reduce the artifacts resulting from the compression algorithms. S. Tramini et. al. in [28, 29], proposed to minimize the total variation under a l^∞-constraint to reduce the artifacts of JPEG2000 compression. F. Alter et. al. in [2] proposed to minimize the total variation to reduce the distortions of the JPEG compression. We question the use of the total variation, as a prior when dealing with distortions of bounded amplitude, and show that easier differentiable priors sometimes lead to better results.

It is not the first attempt to solve Y. Meyer's decomposition problem. Acknowledged solutions can be found in [24, 5, 14]. To our knowledge, our formulation of Y. Meyer's decomposition model is new, simple, and could lead to the design of fast algorithms.

Finally, and that might be the main contribution of this paper, we give a general framework for the problems of type (1.1).

2 Notations and definitions

2.1 Specific notations

Let us describe the notations we use throughout this paper.

We first remind that $X = \mathbb{R}^n$, $Y = X \times X$ and Z can be any of those two spaces.

All the theory developed in this paper applies to color images. To simplify the notations, we focus on grayscale images.

$f \in Z$ is a given data, and $J : Z \to \mathbb{R}$ is a convex, proper function.

For $u \in X$, $u_i \in \mathbb{R}$ denotes the i-th component of u.

For $g \in Y$, then $g_i \in \mathbb{R}^2$ denotes the $i-th$ component of g, and $g_i = (g_{i,1}, g_{i,2})$.

$< \cdot, \cdot >_X$ denotes the usual scalar product on X. For $u, v \in X$ we have :

$$< u, v >_X := \sum_{i=1}^{n} u_i v_i \tag{2.3}$$

$< \cdot, \cdot >_Y$ denotes the usual scalar product on Y. For $g, h \in Y$:

$$< g, h >_Y := \sum_{i=1}^{n} \sum_{j=1}^{2} g_{i,j} h_{i,j} \tag{2.4}$$

$| \cdot |_p$, $p \in [1, \infty[$ is the l^p norm on X :

$$|u|_p := (\sum_{i=1}^{n} |u_i|^p)^{1/p} \tag{2.5}$$

$| \cdot |_\infty$ is the l^∞-norm on X :

$$|u|_\infty = \max_{i \in \{1,2,\ldots,n\}} |u_i| \tag{2.6}$$

$|| \cdot ||_p$, for $p \in [1, \infty[$ is a norm on Y defined by :

$$||g||_p := (\sum_{i=1}^{n} |g_i|_2^p)^{1/p} \tag{2.7}$$

And finally :

$$||g||_\infty := \max_{i \in \{1,2,\ldots,n\}} |g_i|_2 \tag{2.8}$$

K is a convex set (a ball in a certain metric). When $Z = X$, K is just an hypercube of width 2α, centered on f.

Finally $\lfloor a \rfloor$, is the integer part of $a \in \mathbb{R}$.

2.2 Recalls of convex analysis

Definition 2.1. [Proper function] A convex function J on Z is proper if and only if J is not identically equal to $+\infty$ and that it does not take the value $-\infty$ on Z.

Definition 2.2. [Coercivity] A function $J : Z \to \mathbb{R}$ is said to be coercive if :

$$||u||_2 \to \infty \Rightarrow J(u) \to \infty \tag{2.9}$$

Definition 2.3. [Indicator function] Let $K \in Z$ be a non empty closed convex subset of Z. The indicator function of K, denoted χ_K, is defined by:

$$\chi_K(x) = \begin{cases} 0 & \text{if } x \in K \\ \infty & \text{otherwise} \end{cases} \tag{2.10}$$

if K is a compact set, then χ_K is coercive.

Definition 2.4. [Subdifferential and subgradient] The subdifferential of J at point $u \in Z$, is defined by :

$$\partial J(u) = \{\eta \in Z, J(u) + < \eta, (x-u) >_Z \leq J(x), \forall x \in Z\} \tag{2.11}$$

$\eta \in \partial J(u)$ is called a subgradient.

The subdifferential can be thought of as the set of hyperplanes passing through point u which underestimate the function J. On points where J is differentiable, the subdifferential reduces to a singleton : the classical gradient.

Definition 2.5. [Legendre-Fenchel Conjugate]
Let G be a convex proper application from Z to $\mathbb{R} \cup \{\infty\}$. The conjugate function of G is defined by :

$$G^*(u) = \sup_{x \in Z}(< x, u >_Z - G(x)) \tag{2.12}$$

G^* is a convex proper function. Moreover, we have : $G^{**} = G$.

2.3 Discretization of the differential operators

To simplify the notations, we denote $u(i,j)$ the value of u on pixel (i,j).

To discretize the gradient we use the classical first order scheme borrowed from [7]. For $u \in X$:

$$(\nabla u)(i,j) = ((\partial_1 u)(i,j), (\partial_2 u)(i,j)) \tag{2.13}$$

∇u is an element of Y.

$$(\partial_1 u)(i,j) = \begin{cases} u(i+1,j) - u(i,j) & \text{if } i < n_x \\ 0 & \text{if } i = n_x \end{cases} \tag{2.14}$$

$$(\partial_2 u)(i,j) = \begin{cases} u(i,j+1) - u(i,j) & \text{if } j < n_y \\ 0 & \text{if } j = n_y \end{cases} \tag{2.15}$$

This definition allows to define the divergence properly by duality, imposing :

$$< \nabla u, p >_Y = - < u, div(p) >_X \tag{2.16}$$

We will see that having such a relation is necessary to compute the exact subgradient of the discrete functional. Simple algebra gives :

$$\begin{aligned}(div(p))(i,j) &= \begin{cases} p^1(i,j) - p^1(i-1,j) & \text{if } 1 < i < n_x \\ p^1(i,j) & \text{if } i = 1 \\ -p^1(i-1,j) & \text{if } i = n_x \end{cases} \\ &+ \begin{cases} p^2(i,j) - p^2(i,j-1) & \text{if } 1 < j < n_y \\ p^2(i,j) & \text{if } j = 1 \\ -p^2(i,j-1) & \text{if } j = n_y \end{cases}\end{aligned} \tag{2.17}$$

With this discretization, we can define the total variation as :

$$J_1(u) := \sum_{i=1}^{n} |(\nabla u)_i|_2 \tag{2.18}$$

Another quantity of interest in this paper is the discretized hypersurface of u :

$$J_2(u) := \sum_{i=1}^{n} \sqrt{|(\nabla u)_i|_2^2 + 1} \tag{2.19}$$

Those two functions are convex.

3 Existence and uniqueness of the solutions of problem (1.1)

3.1 Existence of a solution

Rewriting problem (1.1) under the form :

$$\inf_{u \in Z} \{J(u) + \chi_K(u)\} \tag{3.20}$$

where χ_K is the indicator function of K, it is clear that problem (3.20) is convex, proper and coercive on Z. Thus, we easily get the existence of a minimizer.

3.2 Non uniqueness of the solutions

When J is strictly convex, we have uniqueness of a minimizer, but in the general convex case, we get a convex set of solutions. We illustrate it through an example in $1-D$ that we study later in $2-D$. We take $Z = X$, $n_x = n$, $n_y = 1$ and set :

$$J(u) = \sum_{i=1}^{n-1} |u_{i+1} - u_i| \tag{3.21}$$

Take $f_i = i/(n-1)$, with $i \in \{0, 2, ..., n-1\}$ and $\alpha < 0.5$.

The interpretation of this model is the following : we are looking for the function of minimal discrete total variation, contained in the red tube drawn on Figure (1).

Clearly, all monotonous increasing sequences satisfying the constraint, going from α and arriving to $1-\alpha$ have the same total variation $(1-2\alpha)$, and are solutions of problem (1.1).

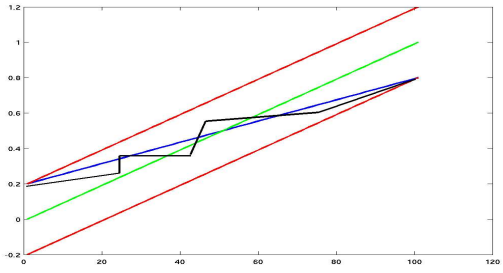

Figure 1: Green : data term - Red : upper and lower constraints - Blue and black : two possible solutions

4 Review of the projected subgradient descent

We go back to the general problem (1.1). J is a general proper convex function on Z, and K is given by (1.2).

In the past few years, many efforts have been dedicated to the design of fast algorithms to solve imaging problems. ROF problem for instance can now be solved in almost real time [8, 11, 12].

Unfortunately, we cannot hope solving (1.1) with a high rate of convergence as we made no assumptions on J except convexity. Imagine that we have an algorithm that generates a sequence $\{u^k\}$ that is supposed to approach a solution of (1.1).

A classical result of convex optimization [20] tells us that no iterative method - only using the sequences $J(u^k)$ and $\partial J(u^k)$ - can achieve a better rate than $\frac{O(1)}{\epsilon^2}$, for general convex minimization problem (ϵ is the desired precision).

If we note $\bar{J}$ the minimum of (1.1), and $\bar{J}^k = \min_{i\in\{1,2,...,k\}}(J(u^i))$, it means that to get a precision $|\bar{J}^k - \bar{J}| \leq \epsilon$, any algorithm will require k to be of order $\lfloor \frac{O(1)}{\epsilon^2} \rfloor$ (see [20] for a more precise statement).

We propose in this paper to use the projected subgradient method for which the complexity is shown to be $\frac{O(1)}{\epsilon^2}$. So that, no method can do better than it for our general

problem, up to a multiplicative factor. To use this technique we need to be able to compute a subgradient and the Euclidean projector on K. In the next sections we show how to compute subgradients. Moreover, the Euclidean projector on K can be computed easily. It can be shown that :

$$(\Pi_K(u))_i = \begin{cases} u_i & \text{if } |u_i - f_i|_2 \leq \alpha \\ f_i + \alpha \frac{u_i - f_i}{|u_i - f_i|_2} & \text{otherwise} \end{cases} \tag{4.22}$$

This algorithm is well known in the optimization community (see for instance [25]). It was proposed recently in image processing in some papers [2, 17, 16]. We describe it briefly below.

4.1 The principle of the algorithm

$\bar{u}$ is a solution of (1.1), if it satisfies for any $t > 0$:

$$\bar{u} = \Pi_K(\bar{u} - t\eta) \tag{4.23}$$

$\Pi_K(\cdot)$ denotes the Euclidean projector on K, and η is any element of $\partial J(\bar{u})$.

To compute $\bar{u}$, we use the projected subgradient algorithm defined by :

$$\begin{cases} u^0 \in K \\ u^{k+1} = \Pi_K(u^k - t_k \frac{\eta^k}{||\eta^k||_2}) \end{cases} \tag{4.24}$$

Here, $t_k > 0$ for any k and η^k is any element of $\partial J(u^k)$.

4.2 How to chose the sequence $\{t_k\}$?

The convergence of this iterative scheme depends on the behaviour of the sequence t_k. Let U be the set of solutions of (1.1). The following theorem proves that for a well chosen sequence t_k, we have convergence [25] :

Theorem 4.1. *If* $\lim_{k\to\infty} t_k = 0$ *and* $\sum_{k=0}^{\infty} t_k = \infty$, *then* $\lim_{k\to\infty} J(u^k) = \inf_{u\in K} J(u)$ *and* $\lim_{k\to\infty} d(u^k, U) = 0$

where $d(\cdot,\cdot)$ *is the Euclidean distance from a point to a set.*

Many sequences satisfy the hypothesis of this theorem. For instance $t_k = \frac{1}{(k+1)^p}$ with $p \in]0, 1]$ or $t_k = \frac{1}{(k+1)\log((k+1))}$. This wide choice makes the theorem difficult to use for practical computations : among all possible sequences, many could lead to very slow convergence. Some ideas were proposed to tackle this problem and the theory is now well established (see [15, 17] for some references).

If J is Lipschitz continous with Lipschitz constant L i.e. :

$$|J(u) - J(v)| \leq L||u - v||_2 \tag{4.25}$$

we can find a parameterless optimal sequence (cf. [27]) :

Theorem 4.2.
The projected subgradient descent with step :

$$t_k = \frac{D}{\sqrt{k}} \tag{4.26}$$

ensures that :

$$\epsilon_k = \bar{J}^k - \bar{J} \leq O(1)\frac{LD}{\sqrt{k}} \tag{4.27}$$

where D is the Euclidean diameter of the set K.

This is an easy choice to ensure a good rate of convergence. More sophisticated techniques can be designed based on iterative estimates of $\bar{J}$ (see [16] for instance). To our knowledge, to really gain efficiency with those techniques, one often needs extra informations about the solution, like the approximate distance from u^k to the set of solutions. It can thus become really difficult to apply in practice.

Finally, we recall that if J is convex, differentiable, with Lipschitz gradient, then we do not need to normalize the gradient, and a constant step size ensures convergence :

Theorem 4.3. *If :*

$$||\nabla J(u) - \nabla J(v)||_2 \leq L'||u - v||_2 \tag{4.28}$$

then the projected gradient descent :

$$\begin{cases} u^0 \in K \\ u^{k+1} = \Pi_K(u^k - t\nabla J(u^k)) \end{cases} \tag{4.29}$$

with constant step $t = \frac{2}{L'}$ ensures that :

$$d(u^k, U) \to 0 \tag{4.30}$$

4.3 Programming aspects and computing time

If we proceed with care, the memory requirements of this algorithm are no more than twice the size of the image and the lengths of our c-codes do not exceed 300 lines. One important advantage of these algorithms is thus to allow fast implementation of a model. The big issue is that computing times can be large.

We can see it when trying to find a stopping criterion. It would be nonsense to assess the condition $0 \in \partial J(u^k)$. Indeed, J might be non-smooth so that the norm of the subgradient might be large even very close to the solution (consider $J(x) = ||x||_2$ for instance). Thus a simple choice is to set the number of iterations.

We will see in the applications that the relation $\epsilon_k \leq O(1)\frac{LD}{\sqrt{k}}$ allows to find the number of iterations needed to get a desired accuracy.

For instance we get 'perceptually stable' solutions pretty fast for the $BV - L^\infty$ problem (45 seconds for Lena), but for the more complex problem of Y. Meyer, we need around 5 minutes for Barbara image. We think that this generic algorithm should be used when willing to get a quick overview of the abilities of a model. Then more refined techniques based on an analysis of the structure of J should be thought of.

Among the concurrent algorithms, let us mention bundle methods [15], which are known to converge faster, but are more difficult to implement in practice, require higher memory space, and have the same theoretical convergence rate. Recently, Yurii Nesterov in [21] proposed a new method with a complexity of $\frac{O(1)}{\epsilon}$. This method applies to a smaller set of problems, but it seems that it might be applied to the problems we present in this paper. It will be the purpose of future investigations.

5 A new solution to Meyer's cartoon + texture decomposition problem

5.1 A brief presentation of the model

This paragraph and this paragraph only is presented in the continuous setting. Ω denotes a bounded open set of $\mathbb{R}^2$. We recall that the space of functions of bounded variation (see [3] for a complete reference) is defined by :

$$BV(\Omega) = \{u \in L^1(\Omega), \sup_{\{\xi \in C_c^1(\Omega;\mathbb{R}^2), \|\xi\|_{L^\infty(\Omega)} \leq 1\}} \int_\Omega u(x) div(\xi(x)) dx < \infty\} \tag{5.31}$$

In [18], Y. Meyer studied theoretically the Rudin-Osher-Fatemi model, and figured out its limitation to discriminate well a cartoon in a noise or a texture. He observed that this limitation could be overpassed using a different data term than the rather uninformative L^2-distance to the data.

He defined a norm :

$$||v||_G = \inf_g \{||g||_\infty, div(g) = v, g = (g_1, g_2), |g| = \sqrt{g_1^2 + g_2^2}\} \tag{5.32}$$

a space : $G = \{v, ||v||_G < \infty\}$,

and proposed to decompose an image f into a cartoon u and a texture v using the following model :

$$\inf_{u\in BV(\Omega),v\in G,f=u+v}\{\int_\Omega |Du| + \lambda||v||_G\} \tag{5.33}$$

The G-norm of an oscillating function remains small, so that this model permits to better extract oscillating patterns of the images. A good example to illustrate this property is the following :

- $||sin(nx)||_{L^2([0,2\pi])} = \pi \quad \forall n \in \mathbb{N}$
- $||sin(nx)||_{G([0,2\pi])} = 1/n \quad \forall n \in \mathbb{N}$

5.2 Current numerical approaches

Y. Meyer did not propose any numerical method to solve his model. The first authors who tried to compute a solution were L. Vese and S. Osher in [24]. They approximate the L^∞-norm in the G-norm, by a L^p-norm to preserve differentiability, and impose that $f = u + v$ by a penalization method. Such an approximation implies some numerical issues, and constrains to use values of p around 2.

J.F. Aujol et. al. in [4] propose a better approximation of the model. They show that the discrete G-semi-norm is the dual of the TV functional and that it can be computed using A. Chambolle algorithm [7]. Nevertheless they also impose that $f = u + v$ by penalization. Their discrete approximation writes :

$$\inf_{u,v}\{TV(u) + \mu\sum_{i,j}|f-u-v|^2_{i,j} + TV^*(\frac{v}{\gamma})\}$$

They propose a convergent algorithm to solve it [4] that was generalized by C. Chaux et al. in [10].

A new approach has recently been developped by D. Goldfarb et. al. in [14]. To our knowledge this is currently the best numerical approximation to Y. Meyer's problem, as no approximation is needed to impose $f = u + v$. The authors solve it using a Second Order Cone Programming (SOCP) approach. This algorithm is also known to converge in polynomial time. The main issue remaining might be the complexity of the codes (the authors use the MOSEK library) and the computing time. In [14], they indicate that the PDE approach to solve Rudin-Osher-Fatemi problem is around twice faster than the SOCP approach.

5.3 Principle and implementation of our method

In the approaches proposed in [24, 4] the user has to give an extra parameter, as $f = u + v$ is imposed by penalization. Taking it too large leads to high computing time, and taking it too small does not ensure that the constraint is taken well into account. In the following, we present a method that permits to avoid this penalization.

Y. Meyer's discretized problem writes :

$$\inf_{u \in X} \{J_1(u) + \lambda \inf_{g \in Y, div(g) = f-u} \{||g||_\infty)\}\} \tag{5.34}$$

Where J_1 is defined in (2.18). We have the following result :

Proposition 5.1.
Problem (5.34) can be reformulated as follows :

$$\inf_{g \in Y, ||g||_\infty \leq \alpha} \{J_1(f - div(g))\} \tag{5.35}$$

Proof: The first idea is that we can use the change of variable :

$$u = f - div(g) \tag{5.36}$$

to get an optimization problem that depends only of one variable g. This is possible because the operator div is surjective from Y to X.

$$\inf_{u \in X} \{J_1(u) + \lambda \inf_{g \in Y, div(g) = f-u} \{||g||_\infty\}\} = \inf_{g \in Y} \{J_1(f - div(g)) + \lambda ||g||_\infty\} \tag{5.37}$$

This problem could be solved using a subgradient descent. Turning the lagrange multiplier λ into a constraint, we get the following minimization problem (we refer to the appendix for the link between the two models) :

$$\inf_{g \in Y, ||g||_\infty \leq \alpha} \{J_1(f - div(g))\} \tag{5.38}$$

■

Notice that if we replace the L^∞-norm by a L^2-norm in (5.34), we get the model of Osher-Solé-Vese [23].

5.4 Numerical implementation

We have shown that the problem of Y. Meyer can be stated as follows :

$$\inf_{g \in Y, ||g||_\infty \leq \alpha} \{J(g)\} \tag{5.39}$$

With :

$$J(g) := \sum_{i=1}^{n} |(\nabla(f - div(g)))_i|_2 \tag{5.40}$$

To use the subgradient descent stated in (4.24) the main difficulty that remains is to compute numerically a subgradient.

To do so, we use the following result on the subgradient of composed functions [30] :

Lemma 5.1. *If $\phi : Y \to \mathbb{R}$ is convex and proper, A is a linear operator from X to Y, that can be decomposed under the form $Au = A_0 u + b$ then :*

$$\partial(\phi \circ A)(u) = A_0^*(\partial\phi(Au)) \tag{5.41}$$

Choosing :

$$\phi(x) = ||x||_1 \tag{5.42}$$

We can write :

$$J(g) = \phi(\nabla(f - div(g))) \tag{5.43}$$

Moreover, $\nabla(f - div(g)) = A_0 g + b$, with $b = \nabla f$ and :

$$A_0(\cdot) = A_0^*(\cdot) = -\nabla div(\cdot) \tag{5.44}$$

Note that this operator is self-adjoint. We can show easily that :

$$(\partial\phi(g))_i = \begin{cases} \frac{g_i}{|g_i|_2} & \text{if } |g_i|_2 > 0 \\ \{\eta_i \in \mathbb{R}^2, |\eta_i|_2 \leq 1\} & \text{otherwise} \end{cases} \tag{5.45}$$

An element of the subgradient of J can thus be given by $A_0^*(\Psi)$, with $\Psi \in Y$ such that :

$$(\Psi)_i = \begin{cases} \frac{(Ag)_i}{|(Ag)_i|_2} & \text{if } |(Ag)_i|_2 > 0 \\ 0 & \text{otherwise} \end{cases} \tag{5.46}$$

On the pixels where $|\nabla(f - div(g^n))|_2$ do not vanish, the gradient of the J functional is :

$$\nabla(div(\frac{\nabla(f - div(g^n))}{|\nabla(f - div(g^n))|_2})) \tag{5.47}$$

which we do not need to regularize.

The last task remaining to use the projected subgradient descent method is to design the sequence t_k, i.e. to find D and to show $L < \infty$.

It is trivial to show that :

$$D = 2\alpha\sqrt{n} \tag{5.48}$$

It is the Euclidean diameter of the l^∞-ball.

Finally, we can show that J is L-Lipschitz with :

$$L \leq 16\sqrt{n} \tag{5.49}$$

Indeed, $\partial J(g)$ can be expressed as $\nabla(div(p))$, for some p with $|p_i|_2 \leq 1 \ \forall i$. So that $|(div(p))_i| = |a+b+c+d| \leq 4$ for some a, b, c, d of norm less than 1. Using the same trick, $\partial J(u) = \nabla v$, with $|v_i|_2 \leq 4 \ \forall i$ ensures that $|(\partial J(u))_i|_2 \leq 16$. So that finally $||\partial J(u)||_2 \leq 16\sqrt{n}$.

Note that the overall complexity of this algorithm is less than :

$$O(1)\frac{16\alpha n}{\sqrt{k}} \tag{5.50}$$

The complexity increases linearly with the number of pixels and with α.

5.5 Results

Finally we present the numerical results on figure (3). The Barbara image is rescaled in $[0, 1]$, and we take $\alpha = 0.1$.

After 4 seconds (100 iterations), we get the middle pictures. The model seems to achieve well what it is supposed to do. The stripes in Barbara clothes are removed, the cartoon part is not visible in the textured part. Moreover, this is done fast.

After 5 minutes (7500 iterations), we reach stability. The result is pretty deceiving. We can see that not all stripes are removed in the cartoon part, and that some cartoon patterns appear in the texture part (see BR (3)). Even worst : the top images represent the results with Rudin-Osher-Fatemi model. The l^2-norms of the residuals in the Top-Right (TR) image and in the Bottom-Right (BR) image are the same. It seems that ROF model performs even better than Y. Meyer's model.

This result is different from the conclusions of JF. Aujol et. al. in [5], where the algorithm gave less geometry in v. We think that the approach proposed in our paper approximates better Y. Meyer's model, as no extra parameter is needed, and that we computed explicitly the rate of convergence. Moreover, the results we obtain are concordant with the conclusions of Wotao Yin et. al. in [31]. The authors show that the Vese-Osher model and the simpler $BV - L^1$ model both outperform Y. Meyer's model for some tasks of image decomposition.

Y. Meyer's discretized model does not seem as promizing as in the continuous setting. We will lead further investigations to figure out the reason for that.

Figure 2: Original Barbara image

6 Using duality to solve l^p problems

6.1 The $BV - l^p$ model

In this section, we look at a second application of our general framework. We are interested in the following problem :

$$\inf_{u \in X} \lambda |u - f|_p^p + J_1(u) \tag{6.51}$$

for $p \in]1, \infty[$. We minimize the total variation subject to a l^p constraint. When $p = 2$ this model is the ROF model [26]. We treat the case $p = 1$ separately.

Note that the solution of (6.51) is unique due to strict convexity and coercivity of the l^p norm.

We first recall some facts of convex analysis (see [13], for a complete reference).

$\Lambda : X \to Y$ denotes a linear differential operator. $\Lambda^* : Y \to X$ is its dual operator. By definition, for any $u \in X$, $q \in Y$, we have :

$$< \Lambda u, q >_Y = < u, \Lambda^* q >_X \tag{6.52}$$

Figure 3: TL : Cartoon part for ROF model - TR : Texture part for ROF model - ML : Cartoon part after 4 seconds - MR : Texture part after 4 seconds - BL : Cartoon part after 5 minutes - BR : Texture part after 5 minutes

Let $F : X \to \mathbb{R}$ and $G : Y \to \mathbb{R}$ be two convex proper functions. Let $\mathcal{P}$, be the primal problem :

$$\inf_{u \in X} G(\Lambda u) + F(u) \tag{6.53}$$

The dual problem $\mathcal{P}^*$ is then defined by :

$$\inf_{q \in Y} G^*(-q) + F^*(\Lambda^* q) \tag{6.54}$$

Let $\bar{u}$ and $\bar{q}$ be the solutions of $\mathcal{P}$ and $\mathcal{P}^*$ respectively. Those solutions are related through the extremality relations :

$$F(\bar{u}) + F^*(\Lambda^* \bar{q}) =< \Lambda^* \bar{q}, \bar{u} >_X \tag{6.55}$$

$$G(\Lambda \bar{u}) + G^*(-\bar{q}) =< -\bar{q}, \Lambda \bar{u} >_Y \tag{6.56}$$

In the following p' is the conjuguate of p. When $p \in]1, \infty[$, $p' = \frac{p}{p-1}$. When $p = 1$, $p' = \infty$. When $p = \infty$, $p' = 1$.

Proposition 6.2. *The dual problem of (6.51) is defined by :*

$$\inf_{\{q \in Y, ||q||_\infty \leq 1\}} < -div(q), f >_X - \beta |div(q)|_{p'}^{p'} \tag{6.57}$$

where :

$$\beta = (\lambda p)^{-1/(p-1)} - \lambda (\lambda p)^{-p'} \tag{6.58}$$

Proof: In our case $\Lambda = \nabla$, $\Lambda^* = -div$, $F(u) = \lambda |u - f|_p^p$, and $G(q) = ||q||_1$.

First, we recall that from Hölder inequality we have for $p \in [1, \infty[\cup \{\infty\}$:

$$\sup_{\{r \in Y, ||r||_p = t\}} q \cdot r = t ||q||_{p'} \tag{6.59}$$

Thus,

$$\begin{aligned} G^*(-q) &:= \sup_{r \in Y} < -q, r >_Y - ||r||_1 && (6.60)\\ &= \sup_{t>0} \sup_{||r||_1 = t} < -q, r >_Y - t && (6.61)\\ &= \sup_{t>0} t ||q||_\infty - t && (6.62)\\ &= \chi_K(q) && (6.63) \end{aligned}$$

Where K, is the set $\{q \in Y, ||q||_\infty \leq 1\}$.

For $1 < p < \infty$:

$$\begin{aligned} F^*(-div(q)) &:= \sup_{u\in X} < -div(q), u >_X -\lambda|u-f|_p^p & (6.64)\\ &= \sup_{v\in X} < -div(q), (v+f) >_X -\lambda|v|_p^p & (6.65)\\ &= < -div(q), f >_X + \sup_{t>0} \sup_{|v|_p=t} - < div(q), v >_X -\lambda t^p & (6.66)\\ &= < -div(q), f >_X + \sup_{t>0} t|div(q)|_{p'} - \lambda t^p & (6.67)\\ &= < -div(q), f >_X + \beta|div(q)|_{p'}^{p'} & (6.68) \end{aligned}$$

The last equation is obtained by canceling the derivative of the function $t \to t|div(q)|_{p'} - \lambda t^p$. Doing so, we find $\beta = (\lambda p)^{-1/(p-1)} - \lambda(\lambda p)^{-p'}$.

■

Proposition 6.3. *The extremality relations allow to show :*

$$\bar{u} = f - \beta p' |div(q)|^{p'-2} div(\bar{q}) \tag{6.69}$$

Proof: First we recall that $F^{**}(u) = F(u)$.

So that $F(\bar{u}) = \sup_{v\in X} < \bar{u}, v >_X -F^*(v)$.

The first extremality relation also gives $F(\bar{u}) =< -div(\bar{q}), \bar{u} >_X -F^*(-div(\bar{q}))$.

Those two equations imply that $-div(\bar{q})$ cancels the derivative of $v \to < \bar{u}, v >_X -F^*(v)$.

So that :

$$(\bar{u} - f) + \beta p' |div(\bar{q})|^{p'-2} div(\bar{q}) = 0 \tag{6.70}$$

Which ends the proof.

■

The numerical interest of duality is pretty clear. We now have to solve a differentiable problem under a convex constraint, instead of having a non differentiable problem. We can use the constant step projected gradient descent to solve it. For $p = 2$, A. Chambolle in [7], proposed a convergent algorithm to solve problem (6.57). In [8] he remarks that the performances of this algorithm seem to be slightly inferior, to the projected gradient descent.

6.2 The $BV - l^1$ model

We now study the problem :

$$\inf_{u \in X, |u-f|_1 \leq \lambda} J_1(u) \tag{6.71}$$

This model was initially proposed to remove impulse noise. We refer to [1, 22, 11, 12] for a good overview of the nice properties of this model. We just recall that it is a simply convex function, so that there should not be uniqueness of the solution in general. In the following, we show how we can apply duality to transform it into form (1.1).

Proposition 6.4. *The dual problem of (6.71) is defined by :*

$$\inf_{\{q \in Y, ||q||_\infty \leq 1\}} < -div(q), f >_X + \lambda |div(q)|_\infty \tag{6.72}$$

We leave the proof for the reader, as it is very similar to the l^p case.

The relations between the dual and the primal problem are more complicated than in the differentiable case.

Proposition 6.5. *If $div(\bar{q}) \neq 0$, then the extremality relations lead to :*

$$\bar{u}_i = f_i + \lambda \gamma_i \frac{(-div(\bar{q}))_i}{|(div(\bar{q}))_i|} \tag{6.73}$$

With $\gamma = (\gamma_1, \gamma_2, ..., \gamma_n) \in \mathbb{R}^n$ such that :

$$\left\{ \begin{array}{rcl} \gamma_i & \geq & 0 \quad \forall i \in \{1, 2, ..., n\} \\ |\gamma|_1 & = & 1 \\ \gamma_i & = & 0 \text{ if } |(div(\bar{q}))_i| < |div(\bar{q})|_\infty \end{array} \right. \tag{6.74}$$

Proof: Using the first extremality relation, and the fact that $F^{**} = F$, we get :

$$F(\bar{u}) = < -div(\bar{q}), \bar{u} >_X - F^*(-div(\bar{q})) \tag{6.75}$$

$$= \sup_{u \in X} < \bar{u}, u >_X - F^*(u) \tag{6.76}$$

This implies that :

$$0 \in \bar{u} - \partial F^*(-div(\bar{q})) \tag{6.77}$$

Using lemma 11, chapter 5 in [25], we can show easily that if $v \neq 0$, then :

$$(\partial F^*(v))_i = f_i + \lambda \gamma_i \frac{v_i}{|v_i|} \tag{6.78}$$

With γ_i satisfying :

$$\left\{ \begin{array}{rcl} \gamma_i & \geq & 0 \ \ \forall i \in \{1,2,...,n\} \\ |\gamma|_1 & = & 1 \\ \gamma_i & = & 0 \text{ if } |v_i|_2 < |v|_\infty \end{array} \right. \tag{6.79}$$

■

This relation does not allow to retrieve $\bar{u}$ from $\bar{q}$. Nevertheless it gives some information on $\bar{u}$.

We can notice that on pixels where $|(div(\bar{q}))_i| \neq |div(\bar{q})|_\infty$, $u_i = f_i$. This means that many pixels will remain unchanged. This is in agreement with the fact that this model is adapted to remove impulsive noise : just the noisy pixels should be changed.

We have not used the second extremality relation yet. It leads to the following result :

Proposition 6.6.

$$(\nabla \bar{u})_i = |(\nabla \bar{u})_i|_2 \bar{q}_i \tag{6.80}$$

Proof:

$$\begin{array}{rcll} G(\nabla \bar{u}) & = & G^{**}(\nabla \bar{u}) & (6.81) \\ & = & \sup\limits_{q \in Y} < \nabla \bar{u}, q >_Y - G^*(q) & (6.82) \\ & = & < -\bar{q}, \nabla \bar{u} >_Y - G^*(-\bar{q}) & (6.83) \end{array}$$

Thus $-\bar{q}$ solves problem (6.82). This yields the existence (cf. [30]) of multipliers μ_i such that :

$$(\nabla u)_i - \mu_i \bar{q}_i = 0 \tag{6.84}$$

With $\mu_i = 0$ if $|\bar{q}_i|_2 < 1$, or $\mu_i > 0$ if $|\bar{q}_i|_2 = 1$. In both cases we get $\mu_i = |(\nabla u)_i|_2$.

■

(6.80) tells us that $\bar{q}$ represents the orientation of the level lines of $\bar{u}$.

The numerical interest of duality for the $BV - l^1$ problem is not clear. Problem (6.72) can be solved with our algorithm, but does not seem especially simpler than the primal problem.

7 Which prior for bounded noises?

The model we study here writes :

$$\inf_{\{u\in X, |u-f|_\infty \leq \alpha\}} J(u) \tag{7.85}$$

With $X = \mathbb{R}^n$. We will study this model with two different priors J. The first one is the total variation :

$$TV(u) = J_1(u) := \sum_{i=1}^{n} |\nabla u|_2 \tag{7.86}$$

The second is the discretized hypersurface of u :

$$J_2(u) := \sum_{i=1}^{n} \sqrt{|\nabla u|_2^2 + 1} \tag{7.87}$$

The problem with $J = J_1$ is denoted $\mathcal{P}^1$, and the one with $J = J_2$ is denoted $\mathcal{P}^2$. They have respective sets of solutions U^1 and U^2.

7.1 Uniqueness of the solution

A difficulty that we encounter using TV as a prior is the non uniqueness of the solution. Actually, there even is a huge set of solutions in $1-D$, as shown in the introductory example. U_1 is thus a convex set, generally not reduced to a singleton.

On the contrary, the solutions of $\mathcal{P}^2$ are generally unique except in non relevant cases.

Proposition 7.7. *If*

$$amplitude(f) := \max_i(f_i) - \min_i(f_i) \geq 2\alpha \tag{7.88}$$

then the solution of $\mathcal{P}^2$ is unique.

Proof : To prove this proposition, first remark that if $amplitude(f) < 2\alpha$ then the solutions are the constants of the interval $[\min_i(f_i) + \alpha, \max_i(f_i) - \alpha]$. In the other cases uniqueness results from the fact that J_2 is stricly convex in all directions except the constants. Thus the solution is unique up to a constant. When $amplitude(f) \geq 2\alpha$ it is easy to prove that the minimum of the solution must be equal to $\min_i f_i + \alpha$, so that the constant is actually fixed. This ends the proof.

■

Let us justify the interest of those models.

7.2 Uniform white noise : the Bayesian justification

We can justify properly these models by using the maximum a posteriori (MAP) formalism.

Suppose that $f = u + b$, where b is a uniform white noise on $[-\alpha, \alpha]$. Further suppose that we have a probability $P(u)$ over the images that is proportional to $exp(-J(u))$ [1]. Then it is reasonable to recover u from f solving this problem :

$$\sup_u\{P(u|f)\} = \sup_u\{P(f|u)\frac{P(u)}{P(f)}\} \tag{7.89}$$

Which is equivalent to :

$$\inf_u\{-log(P(f|u)) - log(P(u))\} \tag{7.90}$$

Note that $P(f_i|u_i) = \frac{1}{2\alpha}\mathbb{1}_{[u_i-\alpha,u_i+\alpha]}$. As the noise is pixelwise independant, $P(f|u)$ is given by :

$$P(f|u) = \begin{cases} 0 & \text{if } |f-u|_\infty > \alpha \\ \frac{1}{(2\alpha)^n} & \text{otherwise (n is the number of pixels)} \end{cases} \tag{7.91}$$

so that the MAP solution is given by :

$$\inf_{u,|u-f|_\infty \leq \alpha}\{J(u)\} \tag{7.92}$$

This model is not new. S. Tramini et. al. in [28, 29] introduced it for the removal of the artifacts appearing in the JPEG2000 image compression. Those artifacts are bounded and might be the principal application of this model.

7.3 Quantization noise removal

All digital images are quantized, typically on 256 grey levels. It is not impeding for the perception of the images, but when computers try to evaluate precisely some image features, like the orientation of the gradients, quantization induces large errors that prevent algorithms to work properly (see [19] for instance). Pre-processings are then crucially needed.

Let Q be the quantization operator, defined as :

$$\begin{array}{rccc} Q & : & \mathbb{R} & \rightarrow & 2\alpha\mathbb{N} \\ & & u_i & \rightarrow & 2\alpha\lfloor\frac{u}{2\alpha}\rfloor + \alpha \end{array} \tag{7.93}$$

Given a prior J, a natural idea to restore the images is to look for the image which minimizes this prior in the set of antecedents $Q^{-1}(f)$. It can be easily shown that :

$$Q^{-1}(f) = \{u, f = Q(u)\} = \{u, |u-f|_\infty \leq \alpha\} \tag{7.94}$$

[1] This is possible if we suppose that images have a bounded amplitude.

Figure 4: TL : lena image on $[0,1]$ - TR : lena image + $U([-0.2,0.2])$ BL : Denoised with $BV-L^\infty$ BR : Denoised with $MinSurface-L^\infty$

Thus, we are again looking for the solution of the following optimization problem :

$$\inf_{u,|u-f|_\infty \leq \alpha} \{J(u)\} \tag{7.95}$$

In the following, we show through some simple examples, that using the total variation as a prior is a bad idea for these kinds of distortions.

The main justification for the use of TV as a prior, is that it regularizes the images, without blurring the edges. In the case of an l^∞-constraint, this advantage is not clear anymore.

Indeed, imagine that $f(i+1,j) - f(i,j) > C > 2\alpha$ then the set $\{u \in X, |u-f|_\infty \leq \alpha\}$ only contains images which satisfy $u(i+1,j) - u(i,j) > C - 2\alpha$. Thus the l^∞-constraint ensures that the jumps greater than 2α are preserved, with a maximal distortion of 2α.

Another remark of interest is that :

Proposition 7.8. *In 1-D (i.e. when $n_y = 1$), $U_2 \subset U_1$*

This result is quite interesting, it shows that an easy way to pick a solution of U_1 is to solve $\mathcal{P}^2$ which is supposed easier as J_2 is differentiable. We skip the proof as the 1-D problem is not the topic of this paper.

The $2-D$ case leads to new surprises. The non uniqueness is not as pronounced as in the $1-D$ case, and proposition (7.8) do not hold.

We show it through a simple illustrated example. We chose f to be a truncated cone as shown on the Top-Left (TL) figure (5). Then we quantize it (Top-Right figure). It gives a superposition of cylinders. We note this function f. If we refer to the introductory example, we naturally come to think that the solution of $\mathcal{P}^1$ might be any radial increasing function going from $\min(f) + \alpha$ to $\max(f) - \alpha$. Actually the numerical solution we get is a new superposition of cylinders (cf. Bottom-left (BL) figure).

An intuitive explanation, is that in 2-D we try to minimize the integral of the length of the level lines over all levels (this is known as Coarea formula cf. [3]). This integral is clearly less for a cylinder than a truncated cone.

This simple example shows that TV is unadapted for dequantization, as the result is a new quantized image. The Bottom-Right (BR) picture in (5) is the solution of $\mathcal{P}^2$. This function has the minimal surface in the set $\{u \in X, |u-f|_\infty \leq \alpha\}$. It is much closer to the original cone.

We come to the same conclusion with a natural image. See figure (6).

Finally, one can wonder why the MAP solution which is theoretically optimal for both quantization noise and uniform white noise works so badly in one case, and so well in the other. This becomes pretty clear when we picture the different noises (7).

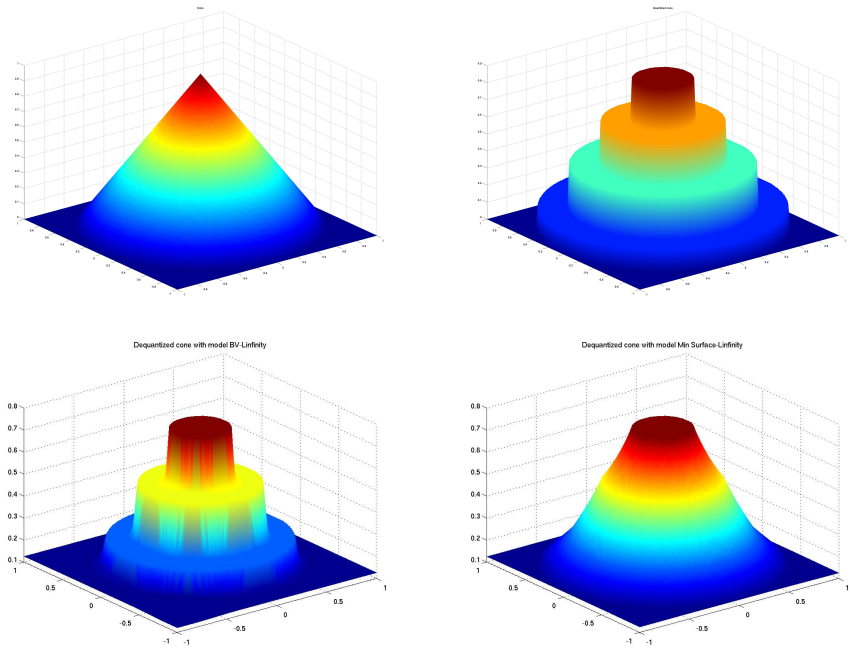

Figure 5: TL : Cone - TR : Quantized cone BL : 'Dequantized' with $BV - L^\infty$ BR : Dequantized with $MinSurface - L^\infty$

Figure 6: TL : original image - TR : Quantized Lena on 10 levels BL : 'Dequantized' with $BV - L^\infty$ BR : Dequantized with $MinSurface - L^\infty$

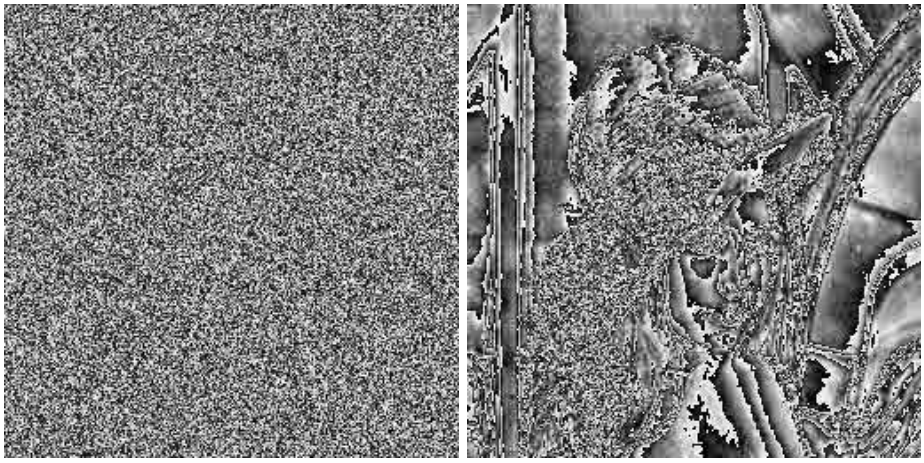

Figure 7: Left : Uniform white noise - Right : Quantization noise

Quantization noise has a small TV, while uniform noise has a large TV. In our MAP approach, the l^∞-norm is supposed to model the noise, and TV the images. Actually, the reason why it works in the case of white noise, is that TV not only models the image but also what is not noise.

7.4 Numerical details

To solve those problems, we used the projected subgradient approach described earlier.

The subgradient of the total variation is given by :

$$div(\Psi) \text{ with } \Psi = \begin{cases} \frac{\nabla u}{|\nabla u|_2} & \text{if } |\nabla u|_2 > 0 \\ 0 & \text{otherwise} \end{cases} \tag{7.96}$$

And it can be shown that $L_1 = ||\partial J_1||_2 \leq 4\sqrt{n}$

Problem $\mathcal{P}^2$ is differentiable. We can use a classical projected gradient descent with constant step to solve it.

8 Conclusion

In this paper, we showed that many models of image processing can be stated under the same formalism (1.1). We recalled the theory of the projected sugradient descent, and successfully applied it to some problems. Those problems include Rudin-Osher-Fatemi model, Meyer's model, and some models for bounded noise denoising. We think that we could reduce

substantially the computing times using more advanced techniques of convex optimization. This will be the topic of further investigations.

Acknowledgement: The first author would like to thank Alexis Baudour for useful mathematical discussions.

A Appendix

Theorem 1.4. *If F and G are two convex functionals of $\mathbb{R}^n$ into $\mathbb{R}^+$, and λ is a real of $\mathbb{R}^+$. The problem :*

$$\inf_{\{u\in\mathbb{R}^n\}}\{F(u)+\lambda G(u)\} \tag{1.97}$$

has a set of solutions U_λ
The problem :

$$\inf_{\{u,\ G(u)\leq\alpha\}}\{F(u)\} \tag{1.98}$$

has a set of solutions U_α
Then :

$$\forall\lambda,\forall u\in U_\lambda,\exists\alpha\in\mathbb{R}^+,u\in U_\alpha \tag{1.99}$$

And conversely :

$$\forall\alpha,\forall u\in U_\alpha,\exists\lambda\in\mathbb{R}^+,u\in U_\lambda \tag{1.100}$$

Here U_λ represents the set of solutions of problem (1.97), and U_α represents the set of solutions of problem (1.98).

We do not give the proof which is pretty long.

References

[1] S. Alliney. A Property of the Minimum Vectors of a Regularizing Functional Defined by Means of the Absolute Norm. *IEEE T. Signal Proces.*, 45:913–917, 1997.

[2] F. Alter, S. Durand, and J. Froment. Adapted Total Variation for Artifact Free Decompression of Jpeg Images. *JMIV*, 23:199–211, 2005.

[3] Ambrosio, Fusco, and Pallara. Functions of Bounded Variation and Free Discontinuity Problems. *Oxford University Press*, 2000.

[4] JF. Aujol. Contribution à l'analyse de textures en traitement d'images par méthodes variationnelles et équations aux dérivées partielles. *Thèse de l'université de Nice-Sophia-Antipolis*, 2004.

[5] J.F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle. Image Decomposition into a Bounded Variation Component and an Oscillating Component. *Journal of Mathematical Imaging and Vision*, 22:71–88, 2005.

[6] J. Bect, L. Blanc-Féraud, G. Aubert, and A. Chambolle. A L^1-unified variational framework for image restoration. *Proc. European Conference on Computer Vision*, pages 1–13, 2004.

[7] A. Chambolle. An Algortihm for Total Variation Minimization and Applications. *JMIV*, 20:89–97, 2004.

[8] A. Chambolle. Total variation minimization and a class of binary MRF models. *Preprint CMAP*, 578, 2005.

[9] Tony F. Chan, Selim Esedoglu, and Frederick E. Park. A Fourth Order Dual Method for Staircase Reduction in Texture Extraction and Image Restoration Problems. *UCLA CAM Report 05-28*, pages 1–13, 2005.

[10] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs. A forward-backward algorithm for image restoration with sparse representations. *Proceedings of the International Conference on Signal Processing with Adaptative Sparse Structured Representations*, pages 49–52, 2005.

[11] J. Darbon and M. Sigelle. Image Restoration with Discrete Constrained Total Variation Part I: Fast and Exact Optimization . *Journal of Mathematical Imaging and Vision*, 2006.

[12] J. Darbon and M. Sigelle. Image Restoration with Discrete Constrained Total Variation Part II: Levelable Functions, Convex and Non-Convex Cases. *Journal of Mathematical Imaging and Vision*, 2006.

[13] I. Ekeland and R. Temam. Analyse convexe et problèmes variationnels. *Dunod Gauthier-Villars*, 1974.

[14] Donald Goldfarb and Wotao Yin. Second-Order Cone Programming Methods for Total Variation Based Image Restoration. *SIAM J. Scientific Computing*, 27:622–645, 2005.

[15] C. Lemarechal, A. Nemirovskii, and Yu. Nesterov. New variants of bundle methods. *Mathematical Programming*, 69:111–147, 1995.

[16] Jian Luo and Patrick Combettes. A Subgradient Projection Algorithm For Non-Differentiable Signal Recovery. *NSIP*, pages 452–456, 1999.

[17] Jian Luo and Patrick Combettes. An adaptive level set method for nondifferentiable constrained image recovery. *IEEE Transactions on Image Processing*, 11:1295–1304, 2002.

[18] Yves Meyer. Oscillating patterns in image processing and in some nonlinear evolution equations. *The Fifteenth Dean Jaqueline B. Lewis Memorial Lectures*, 2001.

[19] Lionel Moisan, Agnès Desolneux, Said Ladjal, and Jean-Michel Morel. Dequantizing image orientation. *IEEE Transactions on Image Processing*, 11:10:1129–1140, 2002.

[20] A.S. Nemirovskii and D.B. Yudin. Problem Complexity and Method Efficiency in Optimization. *Wiley, New-York*, 1983.

[21] Yurii Nesterov. Smooth minimization of non-smooth functions. *Mathematic Programming, Ser. A*, 103:127–152, 2005.

[22] M. Nikolova. A variational approach to remove outliers and impluse noise. *Math. Imag. Vis.*, 20:99–120, 2004.

[23] SJ. Osher, A. Sole, and LA. Vese. Image Decomposition and Restoration using Total Variation Minimization and the H^{-1} Norm. *Multiscale Modeling and Simulation : SIAM*, pages 349–370, 2003.

[24] S.J. Osher and LA. Vese. Modeling Textures with Total Variation Minimization and Oscillating Patterns. *J. Sci. Comput.*, pages 553–572, 2003.

[25] Boris T. Polyak. Introduction to Optimization. *Translation Series in Mathematics and Engineering*, 1987.

[26] L. Rudin, S. Osher, and E. Fatemi. Nonlinear Total Variation Based Noise Removal. *Physica D*, 60:259–268, 1992.

[27] NZ. Shor. Minimization Methods for Nondifferentiable Functions. *Springer-Verlag*, 1985.

[28] S. Tramini, M. Antonini, M. Barlaud, and G. Aubert. Nonlinear Dynamic Filtering for Image Compression. *International Conference on Image Processing*, 3, 1997.

[29] S. Tramini, M. Antonini, M. Barlaud, and G. Aubert. Quantization Noise Removal for Optimal Transform Decoding. *International Conference on Image Processing*, pages 381–385, 1998.

[30] J.B. Hiriart Urruty and C. Lemarechal. Convex Analysis and Minimization Algorithms. *Springer-Verlag*, II, 1996.

[31] Wotao Yin, Donald Goldfarb, and Stanley Osher. A Comparison of Total Variation Based Texture Extraction Models. *submitted to Journal of Visual Communication and Image Representation*, 2006.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399