J Math Imaging Vis 26: 261–276, 2006

DOI: 10.1007/s10851-006-8803-0

Image Restoration with Discrete Constrained Total Variation Part I: Fast and Exact Optimization

JÉRÔME DARBON
EPITA Research and Development Laboratory (LRDE), 14-16 rue Voltaire F-94276 Le Kremlin-Bicêtre, France
École Nationale Supérieure des Télécommunications (ENST), 46 rue Barrault, F-75013 Paris, France
jerome.darbon@{lrde.epita.fr, enst.fr}

MARC SIGELLE
École Nationale Supérieure des Télécommunications (ENST) / LTCI CNRS UMR 5141, 46 rue Barrault, F-75013 Paris, France
marc.sigelle@enst.fr

Published online: 14 August 2006

Abstract. This paper deals with the total variation minimization problem in image restoration for convex data fidelity functionals. We propose a new and fast algorithm which computes an exact solution in the discrete framework. Our method relies on the decomposition of an image into its level sets. It maps the original problems into independent binary Markov Random Field optimization problems at each level. Exact solutions of these binary problems are found thanks to minimum cost cut techniques in graphs. These binary solutions are proved to be monotone increasing with levels and yield thus an exact solution of the discrete original problem. Furthermore we show that minimization of total variation under L^1 data fidelity term yields a self-dual contrast invariant filter. Finally we present some results.

Keywords: restoration, total variation, level sets, Markov random fields, graph cuts

1. Introduction

Minimization of the total variation (TV) for image reconstruction is of great importance for image processing applications [1, 36, 38, 40, 41]. It has been shown that these minimizers live in the space of bounded variation [23] which preserves edges and allows for sharp boundaries. In this paper we propose a new and fast algorithm which computes an exact solution of discrete TV minimization-based problems along with some new theoretical results.

Assume u is an image defined on Ω then its total variation is $TV(u) = \int_\Omega |\nabla u|$, where the gradient is taken in the distributional sense. In this paper, we assume v is an observed image defined on Ω. We are interested in minimizing the following functional:

$$E_v(u) = \int_\Omega f\,(u(x),\ v(x))\ \mathrm{d}x + \beta \int_\Omega |\nabla u|\,. \quad (1)$$

We assume that the attachment to data term is a convex function of $u(\cdot)$, such as: $f\,(u(x),\ v(x))) = |u(x) - v(x)|^p$ for the L^p case ($p = 1, 2$), and that the regularization parameter β is a positive constant. In sequel the total variation model with L^2 (L^1) data fidelity terms are referred to as $L^2 + TV$ ($L^1 + TV$).

A classical approach to minimize TV is achieved by a gradient descent [44] which yields the following evolution equation $\frac{\partial u}{\partial t} = \mathrm{div}(\frac{\nabla u}{|\nabla u| + \epsilon})$. To avoid division by zero, ϵ is set to a small positive value. In [9], Chambolle reformulates TV minimization problem using duality which enables him to propose a fast algorithm. Pollak *et al.* present a fast algorithm providing the exact solution in one dimension [38]. However only an approximation is available in higher dimensions. After a discretization, TV minimization can be reformulated as a minimization problem involving a Markov Random Field (MRF). Boykov *et al.* present a fast approximation minimization algorithm based on graph cuts for MRFs [7]. Ishikawa presents an exact minimization algorithm also based on graph cuts for convex priors [30].

Our minimization algorithm presents some advantages compared to other algorithms which also compute an exact minimizer. The method proposed by Pollak *et al.* [38] only works for one dimensional signal. The minimum cost approach of Ishikawa [30] requires a lot of memory to construct the graph and is time consuming. Compared to this latter approach, our algorithm requires much less memory and is quite faster because the graph we build are by far smaller. However the Ishikawa's approach deal with non convex data fidelity. In [29], Hochbaum originally reformulates the energy (1) with binary variables. She shows essentially the same results as ours. It includes, in particular our Lemma 1. However, the proof is rather different since she is considering the minimum s-excess problem while we make use of stochastic arguments. In [10], Chambolle also proposed a similar approach for total variation minimization (including Lemma 1). His proof relies on submodular functions. During the revised version of this paper, we also became aware of the pioneer work of Zalesky [47] which also presents a proof of our Lemma 1. Both works propose algorithms to perform exact optimization. The one proposed by Zalesky [47] is similar to our sequential one described in Subsection 4.1. A faster algorithm based on a dichotomic approach, that we present in Subsection 4.2, is also proposed by Chambolle in [10]. These two works present some numerical results. Contrary to Chambolle's algorithm we further improve performance using a divide-and-conquer approach. The latter is similar to the work of Hochbaum in [29]. However she goes further since she is able to give the complexity of her algorithm. However, she does not present numerical results.

The contributions of this paper are the following. We propose a fast algorithm which computes an exact minimizer of problem 1. It relies on reformulating this problem into independent binary MRFs attached to each level set of an image. Exact minimization is performed thanks to a minimum cost cut algorithm. We also prove that minimization of the model $L^1 + TV$ yields a contrast invariant and self-dual filter. The rest of this paper is organized as follows. In section 2 we map the original problem 1 into independent binary Markov Random Field optimizations. In section 3 we shed new lights on TV minimization under the L^1-norm as fidelity term. In section 4, a fast algorithm based on graph cuts is presented. Some numerical results are presented in section 5. Finally we draw some conclusions in section 6.

2. Formulation Using Level Sets and MRF

For the rest of this paper, we assume that u takes values in the discrete integer set $\mathcal{L} = [0, L-1]$ and is defined on a finite discrete lattice S. We denote by u_s the value of the image u at the site $s \in S$. An image is decomposed into its level sets using the decomposition principle [28]. It corresponds to considering all thresholding images u^λ where $u_s^\lambda = \mathbb{1}_{u_s \leq \lambda}$. Note the original image can be reconstructed from its level sets using $u_s = \min\{\lambda, u_s^\lambda = 1\}$ as shown in [28].

2.1. Reformulation into Binary MRFs

For any function u which belongs to the space of bounded variation, the coarea formula states that [23] states that $TV(u) = \int_{I\!R} P(u^\lambda)\, d\lambda$ almost surely, where $P(A)$ is the perimeter of the set A. In the discrete case, we write $TV(u) = \sum_{\lambda=0}^{L-2} P(u^\lambda)$ (note that $u_s^{L-1} = 1$ for every $s \in S$, which explains the summation up to $(L-2)$ only.) Let us denote by $s \sim t$ the neighboring relationship between sites s and t and by (s, t) the related clique of order two. For sake of simplicity we shall note sums on cliques of order one and two by $\sum_s$ and $\sum_{(s,t)}$ respectively.

In [5], Boykov *et al.* justify the local estimation of perimeter. Such an estimation has already been done empirically in [8]. We estimate the perimeter locally with cliques of order two. Thus we have

$$TV(u) = \sum_{\lambda=0}^{L-2} \sum_{(s,t)} w_{st}\; |u_s^\lambda - u_t^\lambda| \;. \tag{2}$$

where w_{st} are positive coefficient. See Section 5 for numerical values of these coefficients. We now reformulate the energy as a summation on gray levels.

Proposition 1. *The discrete version of energy (1) rewrites as follows*

$$E_v(u) = \sum_{\lambda=0}^{L-2} E_v^\lambda(u^\lambda) \;+\; C \;\;, \;\; where \tag{3}$$

$$E_v^\lambda(u^\lambda) = \beta \left[\sum_{(s,t)} w_{st}\; ((1 - 2u_t^\lambda)\, u_s^\lambda + u_t^\lambda) \right] + \sum_s (g_s(\lambda+1) - g_s(\lambda))\, (1 - u_s^\lambda) \tag{4}$$

$$g_s(x) = f(x,\; v_s)\;\; \forall s \in S \;\; and \;\; C = \sum_s g_s(0).$$

Proof: Using the fact that for binary variables d, e: $|d-e| = d+e-2de$, and starting from the previous discrete approximation of the coarea formula, we obtain $TV(u) = \sum_{\lambda=0}^{L-2} \sum_{(s,t)} w_{st}\ ((1-2u_t^\lambda)\ u_s^\lambda + u_t^\lambda)$. Moreover the following decomposition holds for any function g of a single variable:

$$\begin{aligned}\forall k \in [0, L-1]\ \ g(k) &= \sum_{\lambda=0}^{k-1} ((g(\lambda+1) - g(\lambda)) + g(0) \\ &= \sum_{\lambda=0}^{L-2} (g(\lambda+1) - g(\lambda))\ \mathbb{1}_{\lambda<k} + g(0).\end{aligned}$$

(notice that this formula is coherent for both $k = 0$ and $k = L-1$). Thus, by defining $g_s(u_s) = f(u_s,\ v_s)$ and since $\mathbb{1}_{\lambda<u_s} = 1 - u_s^\lambda$, we have

$$\begin{aligned}f(u_s,\ v_s) &= g_s(u_s) \\ &= \sum_{\lambda=0}^{L-2} (g_s(\lambda+1) - g_s(\lambda))\ (1-u_s^\lambda) + g_s(0).\end{aligned}$$

This concludes the proof. □

Note that each $E_v^\lambda(\cdot)$ in (4) is a binary MRF with an Ising prior model. We endow the space of binary configurations by the following order : $a \preceq b$ iff $a_s \le b_s\ \forall s \in \Omega$. In order to minimize $E_v(\cdot)$ one would like to minimize all $E_v^\lambda(\cdot)$ independently. Thus we get a family $\{\hat{u}^\lambda\}$ which are respectively minimizers of $E_v^\lambda(\cdot)$. Suppose we do so, then clearly the summation will be minimized and thus we have a minimizer of $E_v(\cdot)$ provided this family is monotone, i.e,

$$\hat{u}^\lambda \preceq \hat{u}^\mu \quad \Leftrightarrow \quad \hat{u}_s^\lambda \le \hat{u}_s^\mu\ \ \forall \lambda \le \mu\ \ ,\ \ \forall s \in S. \tag{5}$$

If this property holds then the optimal solution is given by the reconstruction formula from level sets [28]: $\hat{u}_s = \min\{\lambda, \hat{u}_s^\lambda = 1\}\ \forall s$. Else the family $\{u^\lambda\}$ does not define a function, and thus our optimization scheme is no more valid.

2.2. *Two Lemmas Based on Coupled Markov Chains*

Since the MRF posterior energy is decomposable on levels, we shall use in the sequel both "local neighborhood configuration" $N_s = \{u_t\}_{t\sim s}$ and its level decomposition $N_s^\lambda = \{u_t^\lambda\}_{t\sim s}$ for a given site $s, \lambda \in [0, L-2]$. The local conditional posterior energy at this site will be noted $E_v(\cdot \mid N_s)$ (our assumptions imply that it depends in fact on v_s only.) Then [16]:

Lemma 1. *If the local conditional posterior energy at each site s can be written as*

$$E_v(u_s \mid N_s) = \sum_{\lambda=0}^{L-2} (\ \Delta\phi_s(\lambda)\ u_s^\lambda\ +\ \chi_s(\lambda)\ ), \tag{6}$$

where $\Delta\phi_s(\lambda)$ is a non-increasing *function of λ and $\chi_s(\lambda)$ is a function which does not depend on u_s^λ, then one can exhibit a "coupled" stochastic algorithm minimizing each total posterior energy $E_v^\lambda(u^\lambda)$ while preserving the monotone condition: $\forall s$, u_s^λ is non-decreasing with λ .*

The Lemma states that given a binary solution $a^\star$ to the problem $E_v^\lambda(\cdot)$, there exists at least one solution $\hat{b}$ to the problem $E_v^\mu(\cdot)$ such that $a^\star \preceq \hat{b}\ \forall \lambda \le \mu$. The proof relies on coupled Markov chains [19, 39].

Proof: From the decomposition (6) the local conditional posterior energy at level value λ is $E_v^\lambda(u_s^\lambda \mid N_s^\lambda) = \Delta\phi_s(\lambda)\ u_s^\lambda +\ \chi_s(\lambda)$. Thus the following Gibbs local conditional posterior probability can be computed:

$$\begin{aligned}P_s(\lambda) = P(u_s^\lambda = 1 \mid N_s^\lambda,\ v_s) &= \frac{\exp\{-\Delta\phi_s(\lambda)\}}{1+\exp\{-\Delta\phi_s(\lambda)\}} \\ &= \frac{1}{1+\exp\{\Delta\phi_s(\lambda)\}}.\end{aligned} \tag{7}$$

With the conditions of the Lemma 1, this latter expression is clearly a monotone non-decreasing function of λ.

Let us now design a "coupled" Gibbs sampler for the $(L-1)$ binary images in the following sense: first consider a visiting order of the sites (tour). When a site s is visited, pick up a *single* random number ρ_s uniformly distributed in [0, 1]. Then, for each value of λ, assign: $u_s^\lambda = 1$ if $0 \le \rho_s \le P_s(\lambda)$ or else $u_s^\lambda = 0$ (this is the usual way to draw a binary value according to its probability, except that we use here the same random number ρ_s for the $(L-1)$ binary images.) From the non-decreasing monotony of (7) it is seen that the set of assigned binary values at site s satisfies $u_s^\lambda = 1 \Rightarrow u_s^\mu = 1\ \forall \mu \ge \lambda$. The monotone property $u^\lambda \preceq u^\mu\ \forall\, \lambda \le \mu$ is thus preserved. Clearly, this property also extends to a family of $(L-1)$ coupled Gibbs samplers having *the same* positive temperature T when visiting a given site s: it suffices to replace $\Delta\phi_s(\lambda)$ by $\Delta\phi_s(\lambda)\ /\ T$

in (7). Hence, this property also holds for a series of $(L-1)$ coupled Simulated Annealing algorithms [25] where a *single* temperature T boils down to 0 (either after each visited site s or at the beginning of each tour [45] .) This concludes the proof. □

Our Lemma gives a *sufficient* condition for the simultaneous "level-by-level" minimization of posterior energies while preserving the monotone property. Let us stress again that other proofs of this Lemma are given by Chambolle in [10], Hochbaum in [29] and Zalesky in [47].

Proposition 2. *Lemma 1 applies for both L^1+ TV and L^2+ TV MRF posterior energies.*

Proof: equation (4) implies that, up to constant C:

$$E_v^\lambda(u_s^\lambda \mid N_s^\lambda) = \sum_{\lambda=0}^{L-2}(\ \Delta\phi_s(\lambda)\, u_s^\lambda \ +\ \chi_s(\lambda)\ ),$$

$$\text{with} \begin{cases} \Delta\phi_s(\lambda) \ = \ \beta \left[\sum_{t\sim s} w_{st}\ (1-2u_t^\lambda) \right] - (g_s(\lambda+1) - g_s(\lambda)) \\ \chi_s(\lambda) \quad = \ \beta \left[\sum_{t\sim s} w_{st}\ u_t^\lambda \right] + g_s(\lambda+1) - g_s(\lambda). \end{cases}$$

The contribution of the total variation term to $\Delta\phi_s(\lambda)$, $\beta\ \sum_{s\sim t} w_{st}\ (1-2u_t^\lambda)$, is clearly a non-increasing function of λ (β and $w_{st} \geq 0$). Now, a similar reasoning for L^1 data fidelity yields directly in this case

$$\begin{aligned} g_s(u_s) = |u_s - v_s| &= \sum_{\lambda=0}^{L} |u_s^\lambda - v_s^\lambda| \\ &= \sum_{\lambda=0}^{L} u_s^\lambda(1-2v_s^\lambda) + v_s^\lambda. \end{aligned}$$

The contribution to $\Delta\phi_s(\lambda)$ of $g_s(u_s)$ is thus $(1-2v_s^\lambda)$, which shares the same non-increasing property. On the other side, one has for the L^2 term:

$$\begin{aligned} -(g_s(\lambda+1) - g_s(\lambda)) &= -((\lambda+1-v_s)^2 - (\lambda - v_s)^2) \\ &= -\ (2\ (\lambda - v_s) + 1), \end{aligned}$$

which is also a decreasing function of λ. □

Thus both L^1 + TV and L^2 + TV posterior energies could be minimized "independently" on levels.

We now precisely characterize requirements of Lemma 1. To that aim we recall that a one-dimensional discrete function f defined on $[A, B]$ is *convex* on $]A, B[$ iff $2f(x) \leq f(x-1) + f(x+1)\ \forall x \in]A, B[$ or equivalently, iff $f(x+1) - f(x)$ is a *non-decreasing* function on $[A, B[$.

Lemma 2 . *The requirements stated by Lemma 1 are equivalent to these: all conditional energies $E_v(u_s \mid N_s)$ are* convex *functions of grey level $u_s \in]0, L-1[$, for any neighborhood configuration and local observed data.*

Proof: Since the total energy is "decomposable" on the levels from (3) , so are the local conditional energies:

$$E_v(u_s \mid N_s) = \sum_{\lambda=0}^{L-2} E_v^\lambda(u_s^\lambda \mid N_s^\lambda).$$

Besides, since the local conditional posterior energy component at site s and for level λ is a function of binary variable u_s^λ, it satisfies:

$$\begin{aligned} &E_v^\lambda(u_s^\lambda \mid N_s^\lambda)\ -\ E_v^\lambda(u_s^\lambda = 0 \mid N_s^\lambda) \\ &\quad = \left(E_v^\lambda(u_s^\lambda = 1 \mid N_s^\lambda) - E_v^\lambda(u_s^\lambda = 0 \mid N_s^\lambda)\right)\ u_s^\lambda, \end{aligned}$$

which yields by identification with (6):

$$\Delta\phi_s(\lambda) = E_v^\lambda(u_s^\lambda = 1 \mid N_s^\lambda) - E_v^\lambda(u_s^\lambda = 0 \mid N_s^\lambda). \quad (8)$$

Now, in the transition $\lambda \to \lambda + 1$, only the following level variable does change: $u_s^\lambda = 1 \to u_s^\lambda = 0$. From the decomposition of conditional energies on levels, this means that only the level component $E_v(u_s \mid N_s)$ does change and thus:

$$\begin{aligned} &E_v(\lambda+1 \mid N_s) - E_v(\lambda \mid N_s) \\ &\quad = E_v^\lambda(u_s^\lambda = 0 \mid N_s^\lambda) - E_v^\lambda(u_s^\lambda = 1 \mid N_s^\lambda) = -\Delta\phi_s(\lambda). \end{aligned}$$

The monotone non-increasing condition on $\phi_s(\lambda)$ is thus equivalent to: $E_v(\lambda+1 \mid N_s) - E_v(\lambda \mid N_s)$ is a non-decreasing function on $[0, L-1]$. This concludes the proof. □

Clearly both L^1 + TV and L^2 + TV models enjoy this convexity property and we find again the results of Proposition 2. In next section we study the specific L^1 + TV case.

3. Theoretical Study of the L^1+ TV Case

The use of total variation with L^1 data fidelity has been studied in [2, 4, 12, 34, 35]. However, the results of this section are new as far as we know. Contrast and self-dual invariance properties of the L^1+ TV energy were first proved in [18] for the continuous formulation of equation (1) both in terms of image support and grey levels. Here we prove them in the discrete framework. Then we study the uniqueness vs. non-uniqueness of energy minimizers in this case.

3.1. Contrast and Self-Dual Invariant Filters

Let h be a discrete change of contrast, i.e., a *non-decreasing* application: $\mathcal{L} = [0, L-1] \overset{h}{\mapsto} \mathcal{L}' = [0, L'-1]$.

Lemma 3. *Assume h to be a discrete change of contrast and u a discrete image defined on $\mathcal{L}^S$. The following holds:*

$$\forall \mu \in \mathcal{L}' \ \ \exists \lambda_\mu \in \mathcal{L} \ \ s.t. \ \ (h(u))^\mu = u^{\lambda_\mu}.$$

In other words, after a discrete change of contrast, the level sets of an image $h(v)$ are some level sets of the image v.

Proof: For each $\mu \in \mathcal{L}'$ let us note

$$\lambda_\mu = \sup(\lambda \in \mathcal{L} \mid h(\lambda) \le \mu)$$

It is clear that

$$\forall u_s \in \mathcal{L}, \ h(u_s) > \mu \Leftrightarrow u_s > \lambda_\mu.$$

Thus:

$$(h(u_s))^\mu = {u_s}^{\lambda_\mu}.$$

This conclues the proof. □

Proposition 3. *Let v be an observed image and h be a discrete change of contrast. Assume u to be a global minimizer of $E_v(\cdot)$. Then $h(u)$ is a global minimizer of $E_{h(v)}(\cdot)$.*

Proof: It is sufficient to prove that for any level μ, a minimizer for $h(v)^\mu$ is $h(u)^\mu$. The key point here is that the $L^1 + TV$ total energy decomposes on levels as:

$$E_v(u) = \sum_{\lambda=0}^{L-2} E^\lambda_{v^\lambda}(u^\lambda)$$

Using lemma 3, there exists $\lambda = \lambda_\mu$ such that $v^\lambda = h(v)^\mu$. A minimizer of $E^\lambda_{v^\lambda}(\cdot)$ is u^λ. Thus, u^λ is a minimizer of $E^\lambda_{h(v)^\mu}(\cdot)$ And we have $u^\lambda = h(u)^\mu$. This concludes the proof for contrast invariance. □

Self-dual invariance is easily obtained. Let us define for this purpose the "discrete inverse contrast" operator $\mathcal{L} \overset{\tau}{\mapsto} \mathcal{L}$ as: $\tau(u_s) = L - 1 - u_s$. The proof of the following proposition is straightforward.

Proposition 4. *Let v be an observed image and assume u is a minimizer of $E(\cdot|v)$, then $\tau(u)$ minimizes $E_{\tau(v)}(\cdot)$.*

3.2. Uniqueness vs. Non-uniqueness of Solutions

We now study the behavior of minimizers for the $L^1 + TV$ model. An approach to look for the existence and uniqueness of discrete minimizer(s) comes from the modern theory of phase transitions [20, 26]. Without going to the detail, it can be shown that sensitivity of minimizers to boundary conditions (i.e., the discrete analog of "Dirichlet conditions" rather than Von Neumann's ones) is the signature of a phase transition.

Let us see what happens for the $L^1 + TV$ case. Due to the usual formula $|u_s - v_s| = \sum_{\lambda=0}^{N-2} |u^\lambda_s - v^\lambda_s|$, the energy component at level $\lambda \in [0, L-2]$ is:

$$E^\lambda_v(u^\lambda) = \sum_s |u^\lambda_s - v^\lambda_s| + \beta \sum_{(s,t)} w_{st} \, |u^\lambda_s - u^\lambda_t|. \quad (9)$$

In the case $w_{st} = 1 \ \forall (s,t)$ this corresponds to an isotropic *ferromagnetic* Ising model with *single* coupling constant $J = \beta/2 > 0$ and magnetic field amplitude $B = 1/2$ (whose local sign at site s depends on v^λ_s) *over all levels*. A particular case of equation (9), the binary "chessboard" model [37, 42], i.e., an isotropic 4-connected ferromagnetic Ising model where the observed data v^λ at level λ is a binary chessboard image, was indeed shown to exhibit a phase transition property. Namely when the basic square cell side A satisfies: $A > 4J/B \ (= 4\beta)$ the unique minimal energy configuration (also called ground state) is the initial binary chessboard itself, whatever boundary conditions. In the opposite case two periodic ground states occur, namely the uniform binary white and black images. Physically speaking any object whose characteristic size (diameter) is greater than 4β is conveniently restored, whereas smaller objects are lost in their "background".

In the general case we consider now a grey level chessboard image with constant minimal and maximal grey level values m and M respectively. The associated binary images v^λ can now take three forms:

$$v^\lambda = \begin{cases} \overline{0} & \text{if } 0 \leq \lambda < m \\ c & \text{if } m \leq \lambda < M \\ \overline{1} & \text{if } M \leq \lambda \leq L-2 \end{cases},$$

where c is the binary chessboard image defined by $c = \{\mathbb{1}_{v_s = M}\}_{s \in S}$.

- Minimizing energies E^λ for levels outside "effective" grey level range $[m, M[$ yields:

$$u^\lambda = \begin{cases} \overline{0} & \text{if } 0 \leq \lambda < m \\ \overline{1} & \text{if } M \leq \lambda \leq L-2, \end{cases}$$

  since it consists in restoring uniform black and white observed binary images. In other words *no grey level value* outside the interval $[m, M]$ is generated. This has to be compared to the continuous approach which generates extra grey levels outside the initial grey level range [11], because of the coefficient ϵ introduced in the numerical scheme in order to avoid division by zero.
- For intermediate levels ($m \leq \lambda < M$) the *same* binary chessboard image $v^\lambda = c$ has to be restored with the *same* energy functional over all these levels while ensuring the monotone condition to hold on $\{u^\lambda\}$. If we fall into the case where $A > 4\beta$ this yields back the original binary image $u^\lambda = c$ for this interval of levels (Fig. 1 in the case $m > 0$), so that our overall restoration scheme yields the original grey level image v, in a perfectly coherent way.

  When the cell size becomes nonstationary previous condition may be no more valid and the analog of a "phase transition" in grey levels can be observed according to the value of β wrt. the characteristic cell sizes (Fig. 2). This property remains to be proved rigorously at the theoretical level.

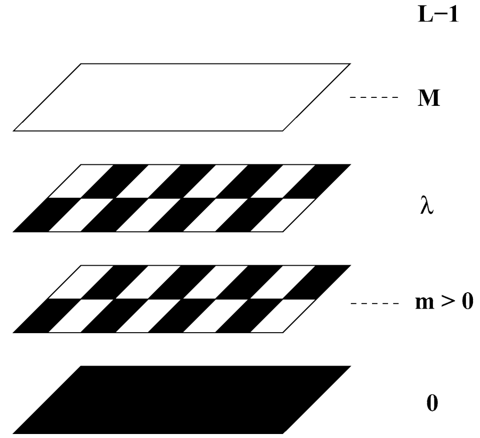


Figure 1. Level-by-level minimum energy configurations for the grey level chessboard model.

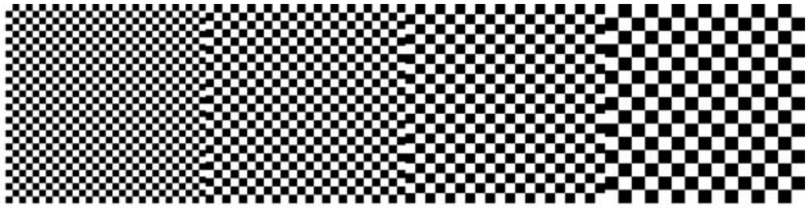
4. Minimization Algorithms: The Convex Case

In this section we propose two alternative algorithms to minimize exactly the total variation with convex data fidelity terms. The first one is a sequential algorithm while the second one relies on a divide-and-conquer approach. Both algorithms rely on results of the previous section. They are based on exact and efficient optimization of binary MRFs thanks to a graph-cut technique. It consists in building a graph such that its minimum cut gives an optimal labelling. The seminal work which describes such an approach is described by Greig *et al.* in [27]. The sequential algorithm requires respectively one cut and $\frac{L}{2}$ cuts for the best and worst case. The divide-and-conquer based algorithm performs $\log_2 L$ minimum cost cuts in any case. We emphasize that both algorithms compute an exact minimizer.

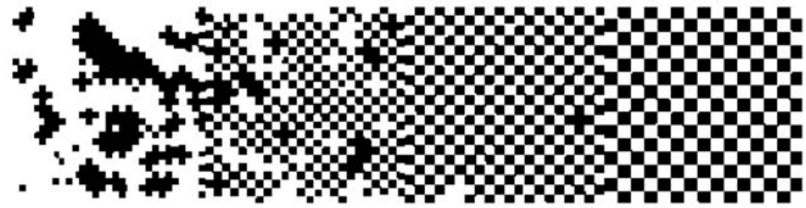
4.1. A Sequential Algorithm

According to the decomposition of the total energy on the level sets of image u and considering the monotone property given in Lemma 1, a straightforward algorithm is to minimize the energy "level by level". Each of these optimizations consists in computing the Maximum *a posteriori* of a binary MRF. This approach leads to perform $(L-1)$ binary optimizations. Assume these optimizations are performed independently, then the monotone property given in equation 5 can be violated if the energy is not strictly convex (like the $L^1 + TV$ model for instance). Thus we need to assure the coherence of solutions.

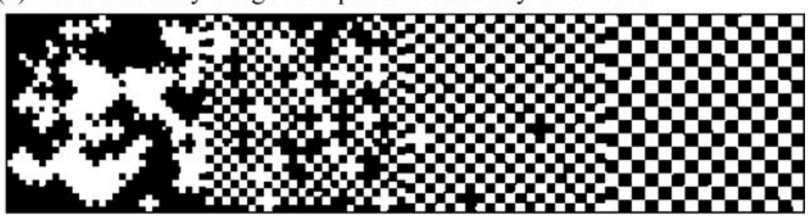
We proceed as follows. We perform optimizations from the lowest, i.e., 0, to the highest grey level, i.e., (L−1). We show the optimality of the solution by induction. Assume we compute a solution $\hat{u}^\lambda$ at a level λ which satisfies the monotone property for levels strictly lower that λ. Let $A_s = \{s \in S | \hat{u}_s^\lambda = 1\}$. The key observation is the following. Recall that Lemma 1



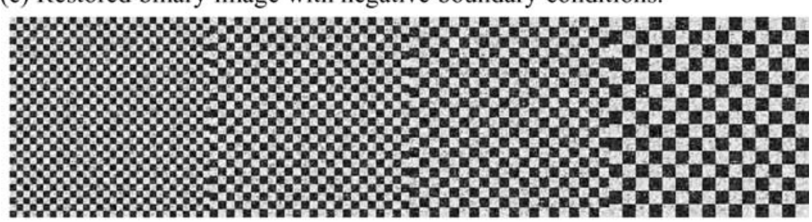
(a) Initial binary image with various cell sizes: 4,5, 6 and 8 (left→right).



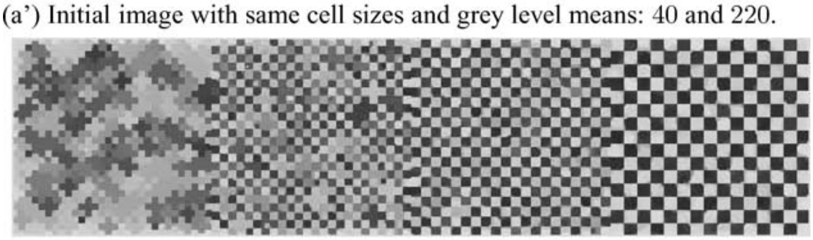
(b) Restored binary image with positive boundary conditions.



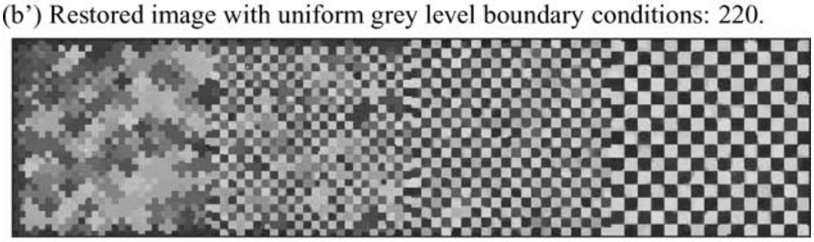
(c) Restored binary image with negative boundary conditions.



(a') Initial image with same cell sizes and grey level means: 40 and 220.



(b') Restored image with uniform grey level boundary conditions: 220.



(c') Restored image with uniform grey level boundary conditions: 40.

Figure 2. Minimal energy configurations obtained by Simulated Annealing. Initial temperature $T_0 = 16$ with decreasing step = 0.98, $\beta = 1.5$ (4-connectivity).

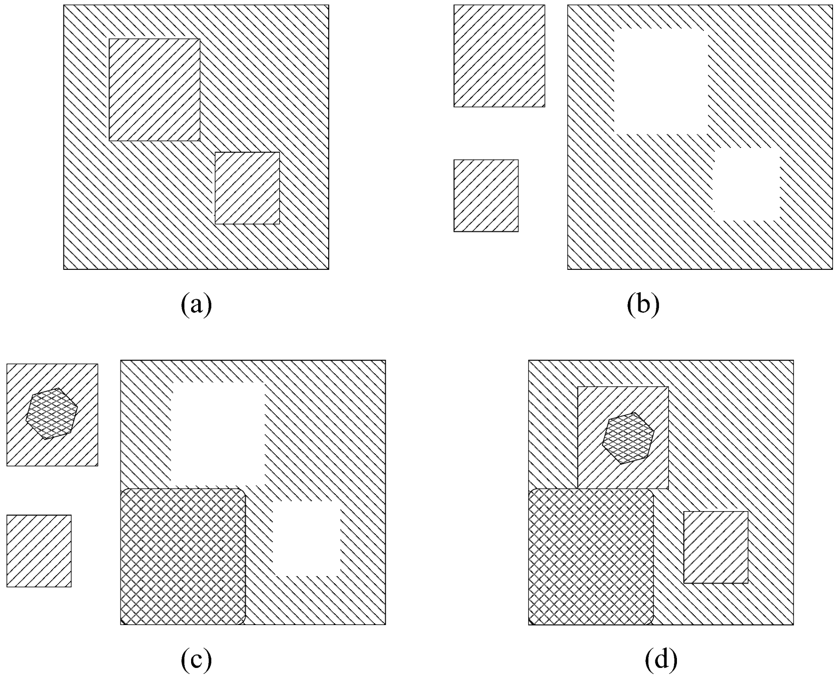


Figure 3. Illustration of our algorithm based on divide-and-conquer technique. The partition of the image after a minimization with respect to some level λ is shown on (a). The connected components of the image (a) are depicted in (b): it corresponds to the decomposition of the problem into subproblems. Each subproblem are solved independently and the result is depicted on (c). Finally the solution of subproblem are recombined and it yields the image (d).

states that given an optimal labelling for the level λ (i.e, $\hat{u}^\lambda$), there exists at least one minimizer for the level $(\lambda + 1)$, referred to as $\hat{u}^{\lambda+1}$, which satisfies the monotone property. The latter is satisfied if $\hat{u}_s^{\lambda+1} = 1 \, \forall s \in A_s$. Such a minimizer for the level $(\lambda + 1)$ is computed by restricting the energy to sites $s \notin A_s$ during the optimization. The obtained minimizer is thus a global minimizer for the level $\lambda + 1$ which also satisfies the monotone property (equation 5). From an implementation point of view, it means that we build the graph such that its minimum cost cut always labels $u_s^{\lambda+1}$ as 1. This algorithm mainly corresponds to the one proposed by Zalesky in [47].

4.2. A Divide-and-Conquer Based Algorithm

We present now another algorithm which takes benefit from inclusion properties of binary solutions in order to increase the performances.

4.2.1. Divide-and-Conquer Assume that $\hat{u}^\lambda$ is an optimal solution for the level λ. Each site s is labelled to a boolean value $\hat{u}_s^\lambda$ which indicates if its optimal grey level value is lesser or equal, or greater than λ. Recall the decomposition of the energy givenby equation (3): $E_v(u) = \sum_{\lambda=0}^{L-2} E_v^\lambda(u^\lambda) \; + \; C$. The terms in the summation only requires the thresholded images u^λ of u (we can drop the constant C since we deal with minimization). The precise value of u is not required. It is useless to take into account pixels which are greater than λ for optimizations which only deal with areas that are already lesser or equal to λ. Obviously, the same observation holds for pixels which are lower than or equal to λ. Consequently, we consider the image connected components (note that they define a partition of the image), and we independently launch optimizations from each others.

These properties lead us to propose an algorithm which rely a divide-and-conquer strategy [14]. Such an approach is as follows:

- first, *decompose* the problem into smaller ones.
- then, *solve* independently each of these subproblems.
- last, *recombine* the solutions of the subproblems in order to get the solution of the global problem.

The decomposition of the problem into smaller ones is performed by computing the connected components of

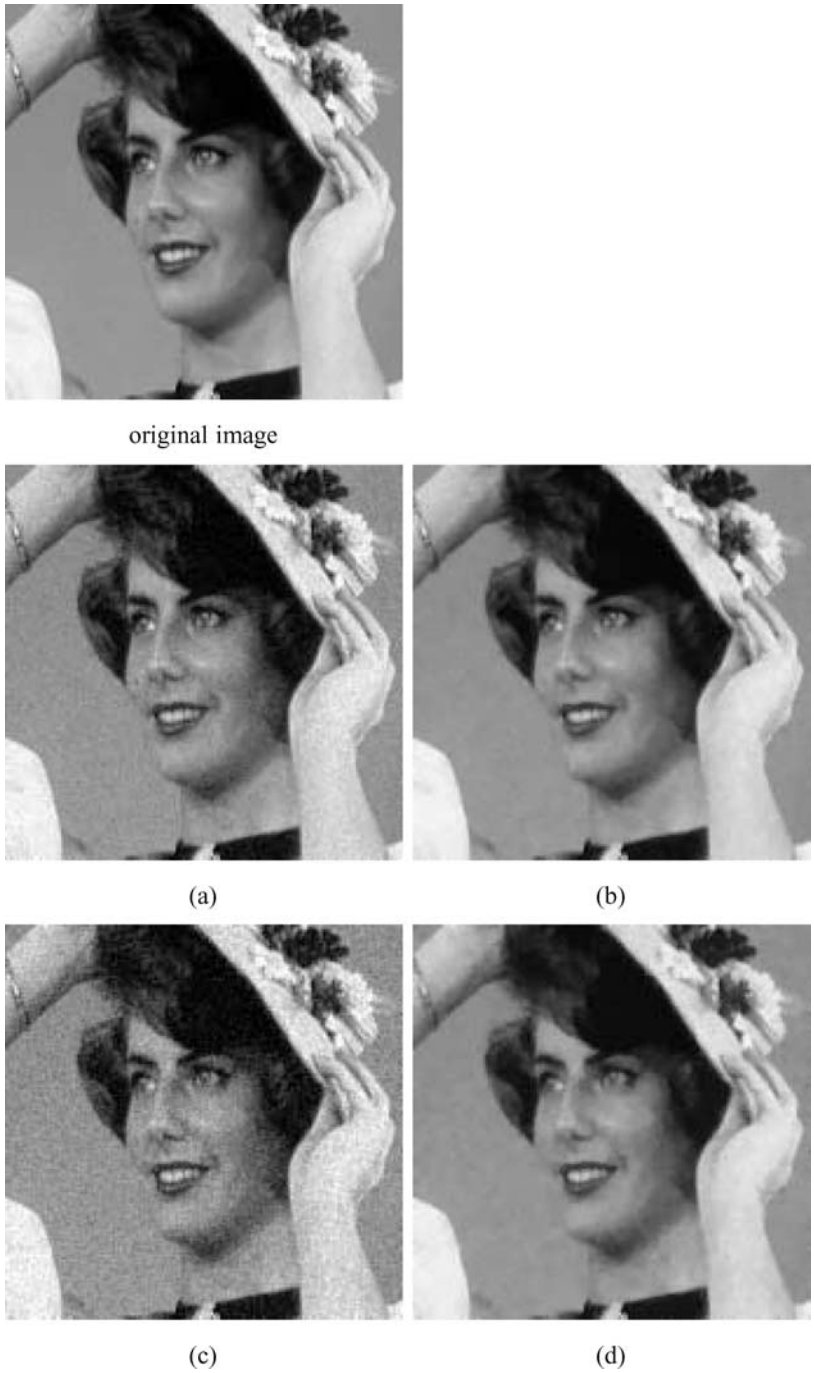


Figure 4. The image *girl* corrupted with an additive Gaussian noise $\sigma = 12$ and $\sigma = 20$ in (a) and (c) respectively. Their $L^2 + TV$ restaurations are shown in (b) and (d) for with $\beta = 23.5$ and $\beta = 44.5$ respectively.

the minimizer at level λ. Combination of solutions is straightforward since the connected components define a partition of the image. This process is depicted on Figure 3.

During the resolution of a subproblem, one has to pay great attention to the local boundary conditions. If the pixels of a connected component are lower than λ, then the neighboring pixels of this component are

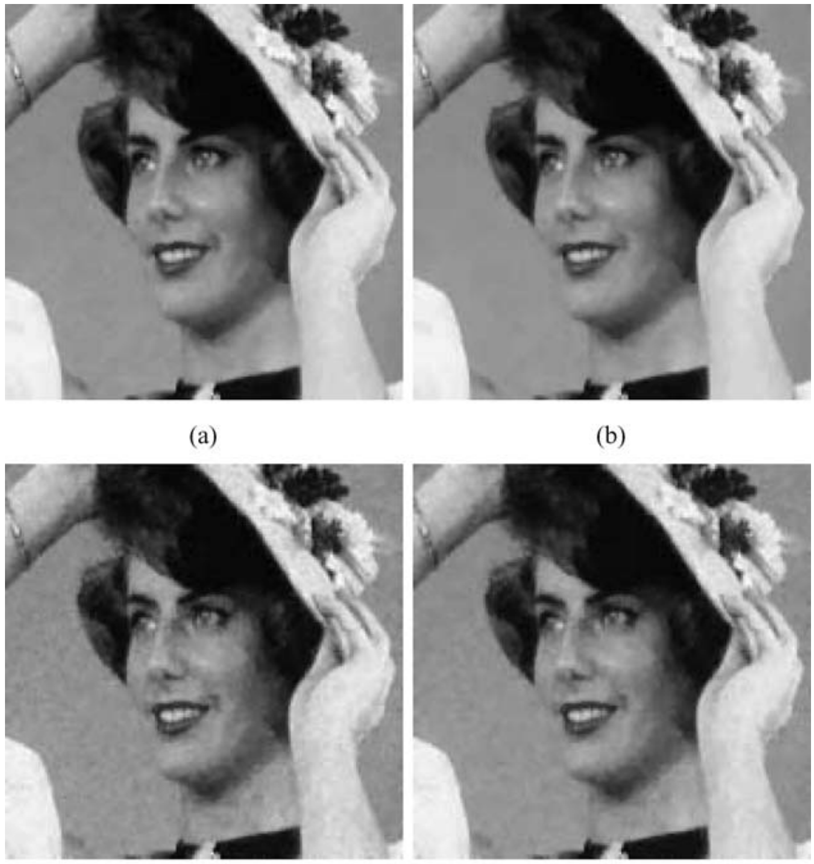


Figure 5. Results of the restoration of the image *girl* with the model $L^2 + TV$ using our and Chambolle algorithms in 4-connectivity. Minimizers for the image *girl* corrupted with a Gaussian noise $\sigma = 12$ are presented in (a) and (b) for Chambolle and our algorithms ($\beta = 16$). Images depicted in (c) and (d) are respectively minimizers using Chambolle and our algorithms ($\beta = 20$) for the image *girl* corrupted with an additive Gaussian noise $\sigma = 20$.

necessarily greater than λ. Indeed, if this was not true, these pixels would be a part of the connected component. A similar reasoning is conducted for connected components whose pixels are greater than λ.

4.2.2. Choice of the Threshold A good strategy for the choice of the level λ at which the optimization has to be performed is to use a dichotomic process. Compared to the sequential algorithm which requires L binary optimizations to compute the solution, this strategy requires only $\log_2(L)$ binary optimizations by pixels. Indeed, a dichotomic strategy requires n comparisons to get a solution with a precision $L2^{-n}$. If one would have an oracle which would give the true value λ for each pixel, then only two minimum cuts for pixel would be needed : one for the level λ and one for the level $(\lambda + 1)$. Finally, the more the size of subproblems are equal, the smaller the complexity of a divide-and-conquer based algorithm is [14].

The dichotomic approach has also been described by Chambolle in [10] however he does not propose the divide-and-conquer improvement. The method of Hochbaum is more similar to ours. She is using an approach similar to the one proposed in [24] to solve the parametric max-flow problem. Due to the inclusion property, she can reduce the size of the graph by merging nodes which are known not to change their labels into a single one, instead of computing the connected components as we do. Besides, her algorithm also takes into account the solution found at the previous step to get an initialization for the new minimum-cut to compute. She shows that the complexity of her algorithm reduces to the complexity of solving a single minimum-cut problem plus the complexity of finding

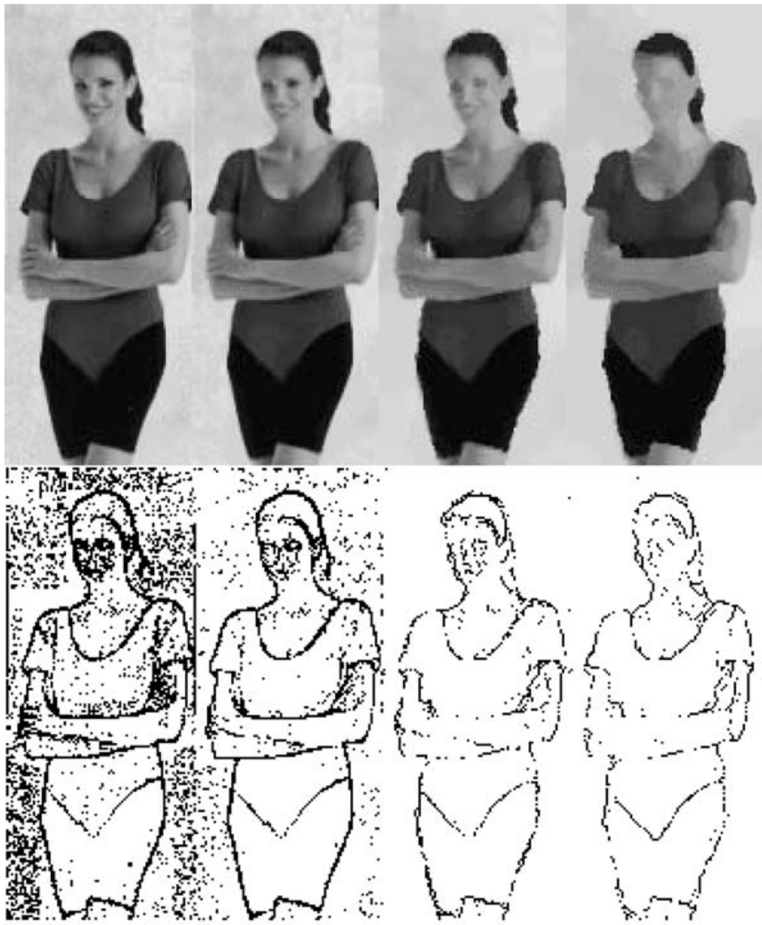


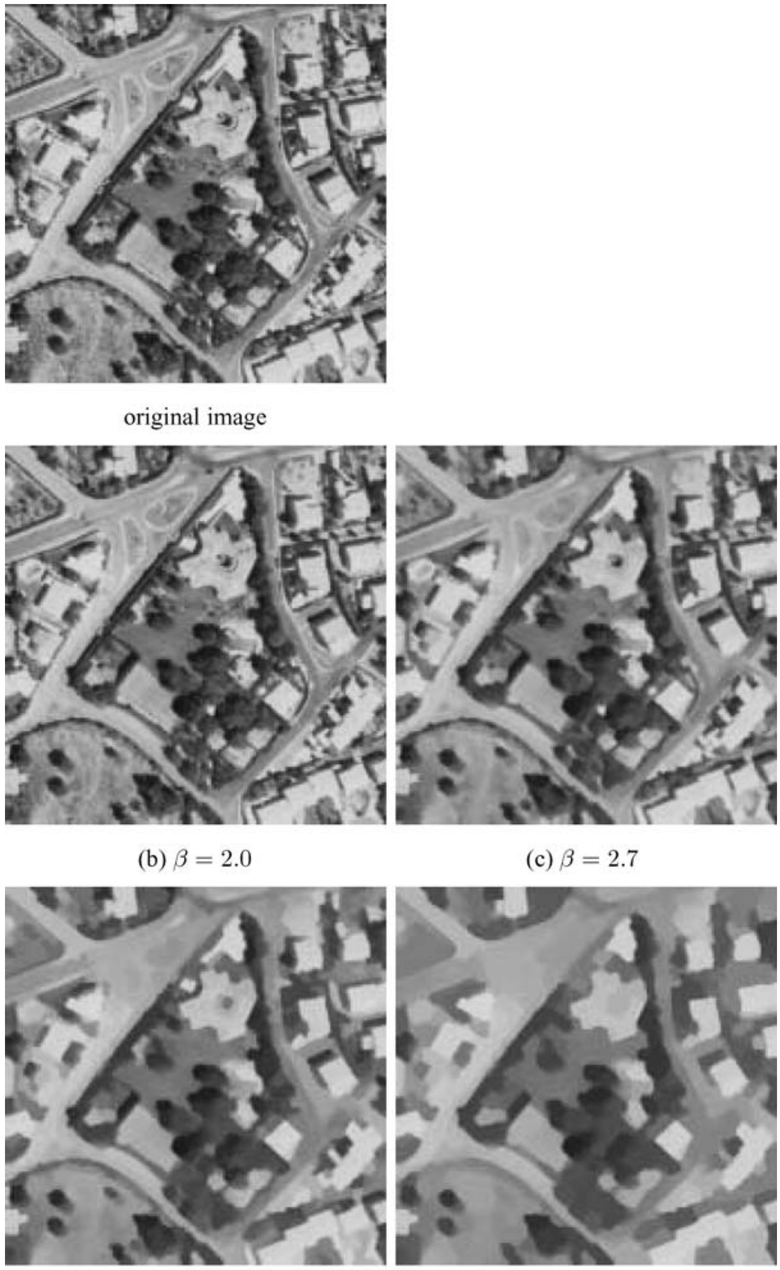
Figure 6. Minimizers of TV with L^1 fidelity for the *woman* image. From left to right: original image, then minimizers for $\beta = 1$, $\beta = 2.1$, $\beta = 3$. Finally, some level lines of the minimizers (in the same order). Only level lines multiples of 10 are displayed.

the integer minima of n convex functions where n is the number of pixels. This complexity is better than the one of our algorithm since we have to compute $log_2(L)$ minimum cuts. Note that we were not able to derive the exact complexity of our algorithm for the average case. However, Hochbaum does not present any numerical results.

5. Experiments and Discussion

Our implementation makes use of the graph construction proposed by Kolmogorov *et al.* in [31] to get an optimal labelling of a binary MRF. Although many minimum cut algorithms are available [14], we used the algorithm described in [6] *et al.* which deals with these energies encountered in computer vision (including our case). We used the approach proposed in [33] for perimeter estimation: it means that in Eq. (2) we set w_{st} to 0.26 and 0.19 for first- and second- nearest neighbors, respectively. For strictly convex energies, we have verified in all experiments that our algorithms and the approach described by Ishikawa in [30] give the same results. For convex energies (but not strictly convex, such as the model $L^1 + TV$), we have obtained different minimizers for each of the three algorithms (sequential, divide-and-conquer, Ishikawa's algorithms), but with the same energy, as predicted by the theory.

Figure 4. depicts our results for the *girl* (256×256) corrupted with additive Gaussian noise of standard deviation $\sigma = 12$ and $\sigma = 20$ with the $L^2 + TV$ model. The results present the stair-case effect. This phenomenon has been already noticed and described



original image

(b) $\beta = 2.0$ (c) $\beta = 2.7$

(d) $\beta = 4.7$ (d) $\beta = 7.0$

Figure 7. Minimizers of TV with L^1 fidelity for the aerial image *Montpellier*. Results for different β are presented.

in [13, 21, 22]. We compare our results with the ones obtained by the duality-based algorithm of Chambolle which is presented in [9]. For a fair comparison, we use the same 4-connectivity as him. It means that we account for the 4 nearest neighbors only, with related coefficients being set to $w_{st} = 1$. Results are depicted on Figure 5. For the Gaussian noise corruption of $\sigma = 12$, our algorithm produces a minimizer whose associated energy is 1.63487×10^7, while the one obtained by Chambolle algorithm is 1.68984×10^7. For the image corrupted with an additive Gaussian noise with $\sigma = 20$, our minimizer has the energy 3.01547×10^7, whereas Chambolle's one is 3.12923×10^7. We also observe a small loss of contrast in the results compared to the

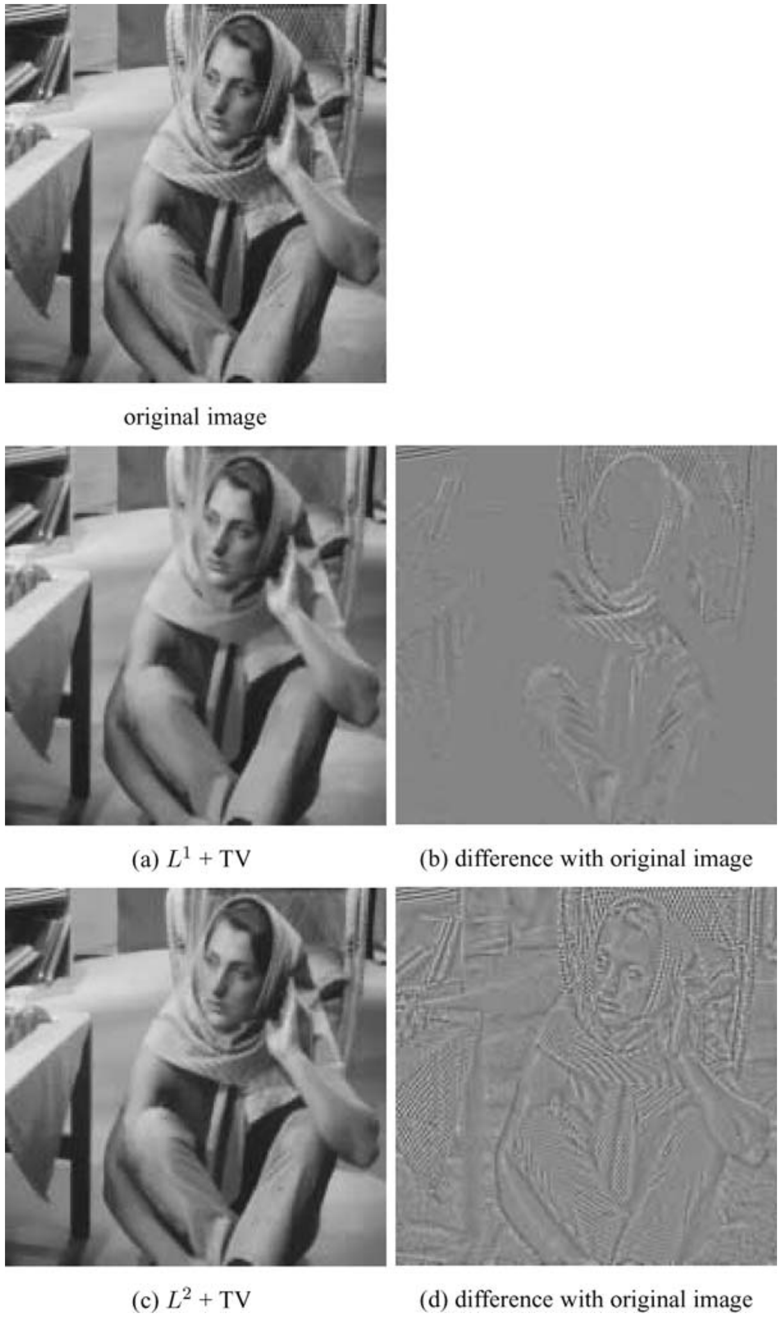


Figure 8. Minimizers of TV with L^1 and L^2 fidelity for image *Barbara*. Differences are shown (texture). The zero is set to the grey level 128 for both difference images.

original images. This behavior is described in [12] by Chan *et al.* and is mainly due to the case of the L^2-norm as data fidelity. These authors suggest to replace the L^2-norm by the L^1-norm in order to preserve the contrast. This is justifed by our result in section 3 which shows that minimization of the model $L^1 + TV$ yields a morphological filter. Figures 6. and 7. depict some results for the L^1+ TV model on respective *woman* (232×522) and *Montpellier* (512×512) aerial images. The more the regularization coefficient β is high, the more images are simplified. For the image *woman* the details of the face disappear while the background tends to become homogeneous. Moreover such a filtering drastically reduce the number of level lines. The same comments apply for the aerial image. Clearly, on both results, the image is simplified while the contrast is maintained. This is due to the morphological behavior of the filter.

As noted by Meyer in [32], the model $L^2 + TV$ can be used for image decomposition. The latter consists in decomposing an image into two components: the first one contains the geometry of the image while the second one contains the texture information. In [46], Yin *et al* study the model $L^1 + TV$ for such decompositions. Note that a new minimization algorithm for $L^1 + TV$, based on iterative thresholding is presented in [4]. Figure 8. depicts results for image *Barbara* using both L^1 and L^2 fidelity terms and TV regularization. The coefficient β is chosen to yield the best visual result. Clearly, the model $L^1 + TV$ outperforms the other one. All the textures are well captured into the texture component for the L^1-based model while some of them are missed by the L^2-based one. Note that, for the L^2-based decomposition, some contours appears in the texture although they should be in the geometric component. Other norms instead of the total variation are considered in [3, 32, 43] and references therein.

Time results (on a 3GHz Pentium IV) for the divide-and-conquer-based algorithm and for the sequential-based algorithm are given on Table 1. for L^2 and L^1 data fidelity. Clearly the divide-and-conquer based algorithm along with dichotomy outperforms the sequential algorithm.

6. Conclusion

In this paper we have presented an algorithm which computes an exact solution for the minimization of the total variation under a convex constraint. The method relies on the decomposition of the problem into binary ones using the level sets of an image. Compared to the state of the art, our algorithm is quite fast and provides a global minimizer in any dimensions. First, the algorithms described by Ishikawa in [30] and Pollak *et al.* in [38] which perform exact optimizations lack of one of these properties. Then, our algorithm presents some improvements compared to those proposed by Zalesky in [47] and Chambolle in [10]. Last, the main difference between our algorithm and the one proposed by Hochbaum [29] is that she makes use of a parametric-based approach [24] while we rely on a divide-and-conquer scheme. Besides, we have shown that minimization of the model $L^1 + TV$ yields a morphological filter.

Several future works are under investigation. First of all, comparison to other existing algorithms performing exact energy minimization (and in particular the one of Hochbaum [29]) has to be made. Besides the calculus of the complexity of our algorithm remains to be done. An extension of the model $L^1 + TV$ for vectorial mathematical morphology is presented in [15].

In the next part of this paper, we extend the proposed approach of energy decomposition on the level sets to a more general class of energies. We show that the case of the total variation is indeed a particular case.

Table 1. Time results (in seconds on a Pentium4 3GHz) with L^1 and L^2 data fidelity term for different weighted term β. Time for the divide-and-conquer and for the sequential (in parentheses) approaches are presented.

L^2 fidelity	Image	$\beta = 23.5$	$\beta = 44.5$
	Girl (256 × 256)	0.79 (12.40)	0.97 (13.24)
	Aerial (512 × 512)	2.88 (47.66)	3.40 (51.75)
L^1 fidelity	Image	$\beta = 2.7$	$\beta = 4.7$
	Girl (256 × 256)	0.73 (12.34)	0.85 (13.22)
	Aerial (512 × 512)	3.53 (56.64)	5.03 (71.86)

Acknowledgements

The authors would like to greatly thank Vincent Arsigny (INRIA Sophia ASCLEPIOS Project), Jean-Francois Aujol (CMLA ENS Cachan) and Antonin Chambolle (CMAP Ecole Polytechnique) for very fruitful discussions and bright insights.

Preliminary versions of this paper appeared in the Proceedings of the *International Workshop on Combinatorial Image Analysis 2004* [16] and in the proceedings of the *Iberian Conference on Pattern Recognition and Image Analysis 2005* [17].

References

1. S. Alliney, "An Algorithm for the Minimization of Mixed l^1 and l^2 norms with Application to Bayesian Estimation," *IEEE Transactions on Signal Processing*, Vol. 42, No. 3, pp. 618–627, 1994.
2. S. Alliney, "A Property of the Minimum Vectors of a Regularizing Functional defined by Means of the Absolute Norm," *IEEE Transactions on Signal Processing*, Vol. 45, No. 4, pp. 913–917, 1997.
3. J.-F. Aujol, G. Aubert, L. Blanc-Feraud, and A. Chambolle, "Image decomposition into a bounded variation component and oscillating component," *Journal of Mathematical Imaging and Vision*, Vol. 22, No. 1, pp. 71–88, 2005.

4. J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-Texture Image Decomposition—Modeling, Algorithms, and Parameter Selection," Technical Report 10, UCLA, 2005.
5. Y. Boykov and V. Kolmogorov, "Computing Geodesic and Minimal Surfaces via Graph Cuts," In: *International Conference on Computer Vision*, Vol. 1, pp. 26–33, 2003.
6. Y. Boykov and V. Kolmogorov, An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 9, pp. 1124–1137, 2004.
7. Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 11, pp. 1222–1239, 2001.
8. O. Catoni and I. Gaudron, "Détection de contours par seuillage adaptatif et restauration stochastique d'images binaires," In: André Gagalowicz (ed.), *Second Annual Conference on Computer Graphics in Paris, Pixim 89 ACM SIGGRAPH FRANCE Hermes*, pp. 341–355, 1989.
9. A. Chambolle, "An Algorithm for Total Variation Minimization and Applications," *Journal of Mathematical Imaging and Vision*, Vol. 20, pp. 89–97, 2004.
10. A. Chambolle, "Total Variation minimization and a class of binary MRF models," In Springer-Verlag, (ed.), *5th International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition(EMMCVPR)*, Vol. LNCS 3757, pp. 136–152, 2005.
11. T. Chan, S. Esedoglu, and M. Nikolova, "Algorithms for Finding Global Minimizers of Image Segmentation and Denoising Models," Technical Report 54, UCLA, 2004.
12. T. Chan and S. Esedoğlu, "Aspect of Total Variation Regularized L^1 Function Approximation," *SIAM Journal of Applied Mathematics*, Vol. 65, No. 5, pp. 1817–1837, 2005.
13. T. Chan, A. Marquina, and P. Mulet, "High-Order Total Variation-Based Image Restoration," *SIAM J. on Scientific Computing*, Vol. 22, No. 2, pp. 503–506, 2000.
14. T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to Algorithms*. The MIT Press, 2001.
15. J. Darbon and S. Peyronnet, "A Vectorial Self-Dual Morphological Filter based on Total Variation Minimization," In: Springer-Verlag, (ed.), *International Symposium on Visual Computing (ISVC)*, Vol. LNCS 3804, pp. 388–395, 2005.
16. J. Darbon and M. Sigelle, "Exact Optimization of Discrete Constrained Total Variation Minimization Problems," In: Springer-Verlag, (ed.), *Tenth International Workshop on Combinatorial Image Analysis (IWCIA 2004)*, Vol. LNCS 3322, pp. 540–549, 2004.
17. J. Darbon and M. Sigelle, "A Fast and Exact Algorithm for Total Variation Minimization," In: Springer-Verlag, (ed.), *2nd Iberian Conference on Pattern Recognition and Image Analysis (IbPria)*, Vol. LNCS 3522, pp. 351–359, 2005.
18. J. Darbon, "Total Variation Minimization with L^1 Data Fidelity as a Contrast Invariant Filter," In *Proceedings of the 4th IEEE International Symposium on Image and Signal Processing and Analysis (ISPA 2005)*, Zagreb, Croatia, 2005.
19. P. Djurić, Y. Huang, and T. Ghirmai, "Perfect Sampling : A Review and Applications to Signal Processing," *IEEE Transactions on Signal Processing*, Vol. 50, No. 2, pp. 345–256, 2002.
20. R. L. Dobrushin, "The Problem of Uniqueness of a Gibbsian Random Field and the Problem of Phase Transition," *Functional Analysis and Applications*, Vol. 2, pp. 302–312, 1968.
21. D. Dobson and C. Vogel, "Recovery of Blocky Images from Noisy and Blurred Data," *SIAM Journal on Applied Mathematics*, Vol. 56, No. 4, pp. 1181–1199, 1996.
22. S. Durand, F. Malgouyres, and B. Rougé, "Image Deblurring, Spectrum Interpolation and Application to Satellite Imaging," *SIAM Journal on Applied Mathematics*, Vol. 56, No. 4, pp. 1181–1199, 1996.
23. L. Evans and R. Gariepy, *Measure Theory and Fine Properties of Functions*. CRC Press, 1992.
24. G. Gallo, M. Grigoriadis, and R. Tarjan, "A fast parametric maximum flow algorithm and applications," *Siam Journal on Computing*, Vol. 18, No. 1, pp. 30–55, 1989.
25. S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Pattern Analysis and Machine Intelligence*, Vol. 6, No. 6, pp. 721–741, 1984.
26. H-O. Georgii, *Gibbs Measures and Phase Transitions*, de Gruyter - Studies in Mathematics, Vol. 9. 1988.
27. D. Greig, B. Porteous, and A. Seheult, "Exact maximum a posteriori estimation for binary images," *Journal of the Royal Statistics Society*, Vol. 51, No. 2, pp. 271–279, 1989.
28. F. Guichard and J. Morel, "Mathematical Morphology, Almost Everywhere", In: *Proceedings of ISMM*, pp. 293–303. (2002).
29. D. Hochbaum, "An efficient algorithm for image segmentation, Markov Random Fields and related problems," *Journal of the ACM*, Vol. 48, No. 2, pp. 686–701, 2001.
30. H. Ishikawa, "Exact optimization for Markov random fields with convex priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 10, pp. 1333–1336, 2003.
31. V. Kolmogorov and R. Zabih, "What Energy can be Minimized via Graph Cuts?" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 2, pp. 147–159, 2004.
32. Y. Meyer. *Oscillating patterns in image processing and nonlinear evolution equations*, Vol. 22 of *University Lecture Series*, American Mathematical Society, Providence, RI, 2001, The fifteenth Dean Jacqueline B. Lewis memorial lectures.
33. H. Nguyen, M. Worring, and R. van den Boomgaard, "Watersnakes: Energy-Driven Watershed Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 3, pp. 330–342, 2003.
34. M. Nikolova, "Minimizers of Cost-Functions Involving Nonsmooth Data-Fidelity Terms," *SIAM J. Num. Anal.*, Vol. 40, No. 3, pp. 965–994, 2002.
35. M. Nikolova, "A Variational Approach to Remove Outliers and Impulse Noise," *Journal of Mathematical Imaging and Vision*, Vol. 20, pp. 99–120, 2004.
36. S. Osher, A. Solé, and L. Vese, "Image Decomposition and Restoration Using Total Variation Minimization and the $\mathbf{H}^{-1}$ *Norm*," *J. Mult. Model. and Simul.*, Vol. 1, No. 3, 2003.
37. E. Pechersky, A. Maruani, and M. Sigelle, "On Gibbs Fields in Image Processing," *Markov Processes and Related Fields*, Vol. 1, No. 3, pp. 419–442, 1995.
38. I. Pollak, A. Willsky, and Y. Huang, "Nonlinear Evolution Equations as Fast and Exact Solvers of Estimation Problems," *IEEE Transactions on Signal Processing*, Vol. 53, No. 2, pp. 484–498, 2005.
39. J.G. Propp and D.B. Wilson, "Exact sampling with coupled Markov chains and statistical mechanics," *Random Structures and Algorithms*, Vol. 9, No. 1, pp. 223–252, 1996.
40. L. Rudin, S. Osher, and E. Fatemi, "Nonlinear Total Variation Based Noise Removal Algorithms," *Physica D.*, Vol. 60, pp. 259–268, 1992.

41. K. Sauer and C. Bouman, "Bayesian Estimation Of Transmission Tomograms Using Segmentation Based Optimization," *IEEE Transactions on Nuclear Science*, Vol. 39, No. 4, pp. 1144–1152, 1992.
42. M. Sigelle, Champs de Markov en Traitement d'Images et Modèles de la Physique Statistique : Application à la Relaxation d'Images de Classification http://www.tsi.enst.fr/sigelle/tsi-these.html. PhD thesis, ENST, 1993.
43. L. Vese and S. Osher, "Image Denoising and Decomposition with Total Variation and Oscillatory Functions," *Journal of Mathematical Imaging and Vision*, Vol. 20, No. 1–2, pp. 7–18, 2004.
44. C. Vogel and M. Oman, "Iterative Method for Total Variation Denoising," *SIAM J. Sci. Comput.*, Vol. 17, pp. 227–238, 1996.
45. G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*," Applications of mathematics. Springer-Verlag, 2003.
46. W. Yin, D. Goldfarb, and S. Osher, "Total Variation Based Image Cartoon-Texture Decomposition," In *3nd IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision*, Beijing, China, 2005.
47. B. Zalesky, "Network Flow Optimization for Restoration of Images," *Journal of Applied Mathematics*, Vol. 2, No. 4, pp. 199–218, 2002.

Jérôme Darbon was born in Chenôve, France in 1978. From 1998 to 2001, he studied computer science at Ecole Pour l'Informatique et les Techniques Avancées (EPITA), France. He received the M.Sc. degree in applied mathematics from E.N.S. de Cachan, France, in 2001. In 2005, he received the Ph.D. degree from Ecole Nationale des Télécommunications (ENST), Paris, France. He is currently a postdoc at the Department of Mathematics, University of California, Los Angeles, hosted by Prof. T.F. Chan. His main research interests include fast algorithms for exact energy minimization and mathematical morphology.

Marc Sigelle was born in Paris on 18th March 1954. He received an engeneer diploma from Ecole Polytechnique Paris 1975 and from Ecole Nationale Supérieure des Télécommunications Paris in 1977. In 1993 he received a PhD from Ecole Nationale Supérieure des Télécommunications. He worked first at Centre National d'Etudes des Télécommunications in Physics and Computer algorithms. Since 1989 he has been working at Ecole Nationale Supérieure des Télécommunications in image and more recently in speech processing. His main subjects of interests are restoration and segmentation of signals and images with Markov Random Fields (MRF's), hyperparameter estimation methods and relationships with Statistical Physics. His interests concerned first reconstruction in angiographic medical imaging and processing of remote sensed satellital and synthetic aperture radar images, then speech and character recognition using MRF's and bayesian networks. His most recent interests concern a MRF approach to image restoration with Total Variation and its extensions.