Exact discrete minimization for TV+L0 image decomposition models
Loïc Denis, Florence Tupin, X. Rondeau

To cite this version:

HAL Id: ujm-00985427
https://hal-ujm.archives-ouvertes.fr/ujm-00985427
Submitted on 29 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
EXACT DISCRETE MINIMIZATION FOR TV+L0 IMAGE DECOMPOSITION MODELS

L. Denis
Observatoire de Lyon, CNRS CRAL, UCBL, ENS de Lyon, Université de Lyon, France

F. Tupin, X. Rondeau
Institut Telecom, Telecom ParisTech, CNRS LTCI, Paris, France

ABSTRACT
Penalized maximum likelihood denoising approaches seek a solution that fulfills a compromise between data fidelity and agreement with a prior model. Penalization terms are generally chosen to enforce smoothness of the solution and to reject noise. The design of a proper penalization term is a difficult task as it has to capture image variability. Image decomposition into two components of different nature, each given a different penalty, is a way to enrich the modeling. We consider the decomposition of an image into a component with bounded variations and a sparse component. The corresponding penalization is the sum of the total variation of the first component and the L0 pseudo-norm of the second component. The minimization problem is highly non-convex, but can still be globally minimized by a minimum s-t-cut computation on a graph. The decomposition model is applied to synthetic aperture radar image denoising.

Index Terms— denoising, discrete minimization, graph-cuts, synthetic aperture radar.

1. INTRODUCTION
Image denoising has been considered in the seminal work of Rudin, Osher and Fatemi (ROF) [1] as the decomposition of an image into a component with bounded variations (BV), and a noise component. The noise is removed by minimizing the total variation (TV) of the BV component, subject to a data fidelity constraint. The popularity of TV-based denoising comes from its ability to preserve sharp edges in the denoised image (i.e., in the BV component).

It has long been noted that ROF model leads to “cartoon-like” images. This can be better understood by considering ROF in a Bayesian framework. Data fidelity expresses the likelihood while total variation models the prior on the denoised image. Piecewise constant images have low total variation, and are therefore given a high prior probability. TV prior therefore biases the solution towards piecewise constant (i.e., “cartoon”) images. This is often considered a problem, sometimes referred to as a staircasing effect [2].

2. THE TV+L0 DECOMPOSITION MODEL
We describe the problem of the recovery of a scene $u$ from an observed image $v$ corrupted by noise. Images are considered sampled and quantized. The scene decomposition model is introduced in the following paragraph. Then, noise modeling is described. Finally, the TV+L0 minimization problem is stated.

2.1. Scene decomposition model
We consider scenes $u$ that can be decomposed as the sum of two terms: $u = u_{BV} + u_S$, where $u_{BV}$ is a component with low total variation (i.e., close to piecewise constant), and $u_S$ is a sparse component (i.e., with all pixels but a few equal to zero).

This work has been funded by DGA under contract REI 2008.34.0042.
The total variation of component $u_{BV}$ can be approximated\(^1\) by $TV(u_{BV}) = \sum_{(k,l)} \|w_{k,l}|u_{BV,k} - u_{BV,l}|$, where $(k, l)$ denotes a clique (i.e., pair of neighboring pixels), and $w_{k,l}$ is a weight. The sparsity of component $u_S$ is defined by the value of the $L0$ pseudo-norm, i.e., the number of non-zero pixels of $u_S$: $L0(u_S) = \sum_k 1 - \delta(u_{Sk})$.

By considering that the two components $u_{BV}$ and $u_S$ are statistically independent, the joint prior distribution $p(u_{BV}, u_S)$ can be modeled in the framework of Markov random fields by:

$$-\log p(u_{BV}, u_S) = -\log p(u_{BV}) - \log p(u_S) = \beta_{BV} TV(u_{BV}) + \beta_S L0(u_S). \tag{1}$$

2.2. Image formation model

We assume that the likelihood distribution that relates the scene $u$ to the noisy image $v$ is separable, i.e., can be written as a product of the likelihood at each pixel $k$:

$$p(v|u) = \prod_k p(v_k|u_k). \tag{2}$$

We therefore consider no mixing (such as that caused by blurring) or noise correlation in our image formation model.

If the noise is additive, the image $v$ can be written as $v = u_{BV} + u_S + n$, with $n$ a noise component. In the case of multiplicative noise, like speckle noise occurring in coherent imaging technique such as SAR, the image writes $v = (u_{BV} + u_S) \times n$. Note that for some noise distributions (e.g., Poisson noise, or Rice distribution), no such relation exists.

The likelihood distribution may depend on each term of the scene decomposition $p(v_k|u_{BV,k}, u_{Sk})$, as will be discussed in section 4.

2.3. Energy minimization problem

Starting from an observed noisy image $v$, the components $u_{BV}$ and $u_S$ can be estimated in the maximum a posteriori sense by solving the following minimization problem:

$$(u_{BV}, u_S) = \arg \min_{(u_{BV}, u_S)} -\log p(v|u_{BV}, u_S) - \log p(u_{BV}, u_S) = \arg \min_{(u_{BV}, u_S)} D(v, u_{BV}, u_S) + \beta_{BV} TV(u_{BV}) + \beta_S L0(u_S). \tag{3}$$

Given an input image $v$ with $N$ pixels, minimization problem (3) has $2N$ unknowns. The objective function is non-continuous and highly non-convex in variable $u_S$ (possibly also in variable $u_{BV}$ depending on $D(\cdot)$) and is not separable in variable $u_{BV}$ (the TV term induces a coupling between values at different sites). It is thus very challenging to solve.

We describe next section how the global minimum can still be found by exact discrete minimization via a combinatorial optimization technique based on a minimum s-t-cut computation.

3. EXACT DISCRETE MINIMIZATION BY GRAPH-CUTS

We show in the following paragraphs how to solve minimization problem (3) exactly among discrete images $u_{BV}$ and $u_S$. At each of their $N$ pixels, these images take values respectively in the discrete ordered sets $\{\alpha_1, \alpha_2, \cdots, \alpha_{m_{BV}}\}$ and $\{\gamma_1, \gamma_2, \cdots, \gamma_{m_S}\}$.

3.1. Problem reformulation

We begin by considering minimization problem (3) for $u_{BV}$ fixed. We write $u^*_S(u_{BV})$ the minimizer of one such restricted problem:

$$u^*_S(u_{BV}) = \arg \min_{u_S} D(v, u_{BV}, u_S) + \beta_S L0(u_S) \tag{4}$$

It comes from the separability of the data fidelity and $L0$ terms (see our hypotheses discussed in paragraph 2.2) that the computation of $u^*_S(u_{BV})$ requires to solve $N$ mono-dimensional minimization problems:

$$\begin{cases} u_{Sk}^*(u_{BV,k}) = u_{Sk}^\dagger & \text{if } -\log p(v_k|u_{BV,k}, u_{Sk}^\dagger) + \beta_S < -\log p(v_k|u_{BV,k}, u_{Sk} = 0), \\ u_{Sk}^*(u_{BV,k}) = 0 & \text{otherwise;} \\ \end{cases} \tag{5}$$

with $u_{Sk}^\dagger = \arg \min_{u_{Sk}} -\log p(v_k|u_{BV,k}, u_{Sk})$ either known in closed form, or numerically computed in $O(n_S)$ operations in the worst case.

We can now reformulate minimization problem (3) into an equivalent problem involving only $u_{BV}$:

$$\arg \min_{u_{BV}} D(v, u_{BV}, u^*_S(u_{BV})) + \beta_{BV} TV(u_{BV}) + \beta_S L0(u^*_S(u_{BV})) \tag{6}$$

The term (i) is of the form $\sum_k f_k(u_{BV,k})$ (i.e., it is separable), and (ii) is a convex term of the form $\sum_{(k,l)} g_{kl}(u_{BV,k}, u_{BV,l})$ (i.e., involving only first order cliques).

Solving problem (6) gives $u_{BV}^\star$. The pair $(u_{BV}^\star, u^*_S(u_{BV}^\star))$ obtained is then a global minimizer of problem (3). We describe in next paragraph how problem (6) can be stated equivalently as a minimum s-t-cut computation on a graph.

3.2. Graph representation

Minimization problems of the form of problem (6), with a non-convex but separable discrete potential $f_k$ and a con-
Fig. 1. Graph construction for discrete minimization.

(a) Pixel grid
(b) Ishikawa’s graph construction
(c) Arc orientations and weights

Fig. 2. Surface representation of Rice neg-log-likelihood (surface drawn for \(v_k = 1\)).

4. RESULTS

We illustrate in the following paragraphs the application of the discrete minimization algorithm to decompose a SAR image into homogeneous regions and strong scatterers with the TV+L0 model.

4.1. Model for SAR images

High and very high resolution SAR scenes are generally a mix of rather homogenous regions and point-like strong scatterers. These scatterers have strong contrast (amplitudes several orders of magnitude higher than that of the homogeneous background). Man-made structures (buildings, bridges, pylons, . . . ) produce such strong echoes.

The observed SAR amplitude image \(v\) is corrupted by speckle noise generated by the coherent summation of echoes coming from different scatterers. In a homogeneous region, noise follows Rayleigh distribution (Gamma distribution when considering intensity rather than amplitude). When a strong scatterer is present in a resolution cell, noise distribution is better described by a Rice distribution[10] (see Fig. 2):

\[
-\log p(v_k|u_{BV_k}, u_{Sk}) = \frac{v_k^2 + u_{Sk}^2}{2u_{BV_k}^2} + 2\log u_{BV_k}
\]

with \(I_0\) modified Bessel function of the first kind. Note that the amplitude \(u_{BV_k}\) must be strictly positive and that \(u_{Sk}\) is either equal to zero (absence of strong scatterer) or strictly positive. In the absence of a strong scatterer, equation (7) simplifies to the widely used Rayleigh model.
4.2. Energy minimization

Rice neg-log-likelihood defined in equation (7) is quasi-convex (but non-convex) in each variable. For a given value of $u_{BV}$, it reaches its unique minimum at $u_{S}^{*}(u_{BV})$ (red curve on Fig. 2). At each pixel $k$, and for each value $\{\alpha_1, \alpha_2, \ldots, \alpha_{m_{BV}}\}$, the corresponding value of $u_{S}^{*}(u_{BV_k})$ can be computed efficiently by bisection. The number of quantization levels $m_S$ of the sparse component can be chosen large to account for the high dynamic range of strong scatterers with very little overhead. The bounded variations component has very low dynamic range and choosing $m_{BV} \approx 100$ is often enough in practice. The positivity constraints $\forall k, u_{BV_k} > 0$ and $u_{S_k} \geq 0$ are straightforwardly enforced by considering discrete levels that satisfy the constraints. The tuning of parameters $\beta_{BV}$ and $\beta_S$ is done by hand by first setting a suitable $\beta_{BV}$ value with $\beta_S = 0$, and then choosing a value of $\beta_S$ that gives a satisfying number of strong scatterers.

4.3. Decomposition results

We illustrate the method on a real SAR image (Fig. 3.a: south of Toulouse city, France, ©CNES/DGA). Fig. 3 compares the results obtained when denoising with a total variation minimization approach and with the TV+L0 decomposition approach. The TV approach corresponds to setting $\beta_S$ to infinity. Strong scatterers are either suppressed or distorted (reduced amplitude, spreading) as they do not fit to the prior model (result of TV denoising shown Fig. 3.b). TV+L0 decomposition produces much smoother regions in the BV component (Fig. 3.c left) and sharp targets in the sparse component (Fig. 3.c right).

5. CONCLUSION

The proposed decomposition provides a better model to images such as in SAR imaging (both in terms of scene prior and noise model). The challenging highly non-convex problem can be exactly solved by computation of a min-cut. Though the number of quantization levels can be limited, the graph-based minimization algorithm is limited in its use by the cost of the memory representation of the graph: $N \times m_{BV}$ nodes and about $3N \times m_{BV}$ pairs of arcs.

6. REFERENCES


