A note on nice-levelable MRFs for SAR image denoising with contrast preservation

Champs de Markov bien nivelés pour la restauration d'Images ROS avec préservation du contraste

Jérôme Darbon
Marc Sigelle
Florence Tupin

2006D006

septembre 2006

Département Traitement du Signal et de l'Image
Groupe Multimédia

Dépôt légal : 2006 – 3ème trimestre
Imprimé à l'Ecole Nationale Supérieure des Télécommunications – Paris
ISSN 0751-1345 ENST D (Paris) (France 1983-9999)

Ecole Nationale Supérieure des Télécommunications
Groupe des Ecoles des Télécommunications - membre de ParisTech
46, rue Barrault - 75634 Paris Cedex 13 - Tél. + 33 (0)1 45 81 77 77 - www.enst.fr
Département TSI

A note on Nice-Levelable MRFs for SAR image denoising with contrast preservation

Jérôme Darbon[1,3] Marc Sigelle[2] Florence Tupin[2]

[1] present address: UCLA Mathematics Department, Box 901555, Los Angeles, CA 90095, USA

[2] Ecole Nationale Supérieure des Télécommunications TSI / CNRS LTCI UMR 5141
46 rue Barrault, F-75013 Paris, France

[3] EPITA Research and Development Laboratory (LRDE)
14-16 rue Voltaire 94276 CEDEX Le Kremlin-Bicêtre, France

E-mails: jerome@math.ucla.edu, marc.sigelle@enst.fr, florence.tupin@enst.fr

6th September 2006
N^o 2006D006
N^{o}ISSN: 0751-1345

A paper based on this technical report has been submitted to SPIE Electronic Imaging 2007 .

Abstract

It is well-known that Total Variation (TV) minimization with L^2 data fidelity terms (which corresponds to white Gaussian additive noise) yields a restored image which presents some loss of contrast. The same behavior occurs for TV models with non-convex data fidelity terms that represent speckle noise. In this note we propose a new approach to cope with the restoration of Synthetic Aperture Radar images while preserving the contrast.

Keywords: Image Restoration, Total Variation, Energy Minimization, Levelable Functions, Synthetic Aperture Radar.

Champs de Markov Bien Nivelés pour la Restauration d'Images ROS avec Préservation du Contraste

Jérôme Darbon[1,3] Marc Sigelle[2] Florence Tupin[2]

[1] adresse actuelle : UCLA Mathematics Department, Box 901555, Los Angeles, CA 90095, USA

[2] Ecole Nationale Supérieure des Télécommunications TSI / CNRS LTCI UMR 5141
46 rue Barrault, F-75013 Paris, France

[3] EPITA Research and Development Laboratory (LRDE)
14-16 rue Volraire 94276 CEDEX Le Kremlin-Bicêtre, France

E-mails : jerome@math.ucla.edu, marc.sigelle@enst.fr, florence.tupin@enst.fr

6 Septembre 2006
N^o 2006D006
N^{o}ISSN : 0751-1345
Un papier utilisant les résultats de ce rapport a été soumis au congrès SPIE Electronic Imaging 2007
.

Résumé

Il est bien connu que la minimisation de la Variation Totale (TV) avec une attache aux données de type L^2 (qui correspond à un bruit blanc Gaussian additif) aboutit à une image restaurée présentant une perte de contraste. Ce même comportement se produit avec une attache aux données non-convexe comme dans le cas du bruit de speckle (multiplicatif). Dans cette note nous proposons une nouvelle approche pour restaurer les images de Radar à Ouverture Synthétique (ROS) tout en préservant le contraste.

Mots-clés : Restauration d'Image, Variation Totale, Minimisation d'Energie, Fonctions Nivelées, Radar à Ouverture Synthétique.

1 Introduction

It is now well-known that when an image containing a bright object within a dark background and overall Gaussian noise corruption is restored using Total Variation (TV) regularization [21], a significant loss of grey level contrast between recovered object and background can happenn [18, 23]. We recently showed that TV is the paradigm of those regularization energy functionals which can be minimized level-by-level, which we called levelable functions [9]. We present in this paper the first application of this formalism to the denoising of Synthetic Aperture Radar (SAR) images where this loss of constrast effect can be preponderant.

Few works have addressed this loss of contrast issue under a Gaussian noise corruption assumption. In [20], Osher *et al.* propose an iterative regularization method which replaces the Total Variation prior by a generalized Bregman distance. The method amounts to minimizing a sequence of variational problems where each of them refine at each step a degraded image. This approach yields very good results compared to many other classical models. This scheme has been extended to a time-continuous nonlinear inverse scale space in [7, 5]. A proof of convergence can be found in [15] and rates of convergence in [6]. Such an approach has been succesfully extended to cartoon extraction from aerial images in [3], image denoising using wavelets in [26] and blind deconvolution by Marquina in [17].

In this paper we focus on the use of variational methods or Markov Random Fields (MRFs) [25] that make use of TV as priors. Note that many other approaches that do fit into this framework are available to perform SAR image denoising, such as [1, 2, 10, 14, 16, 19].

The contributions of this paper are the following. We propose a new framework based on MRF with levelable priors [9] for restoration of images corrupted by Gaussian or Speckle noise. A theoretical study is conducted and describes the behavior of filters defined by Levelable-MRFs. Some preliminary experiments suggest that this new approach performs very well. The remainder of this paper is organized as follows. We introduce our notation and briefly present Levelable and Nice-Levelable Markov Random Fields in Section 2. Section 3 is devoted to the study of the shape of the restored objects using nicelevelable MRFs. Section 4 describes the loss of contrast that occurs when performing the restoration using a Total Variation prior [21] (it is shown in [9] that the latter is levelable). In Section 5, we show how to prevent the result to have a loss of contrast using levelable priors. In Section 6 some very promising results are presented for synthetic images that are corrupted by white Gaussian additive noise and speckle noise. Finally, we draw some conclusions in Section 7.

2 Levelable and Nice-Levelable Markov Random Fields

In this section we briefly present Markov Random Fields with levelable priors. We refer the reader to [9] for further details.

Assume an image defined on a discrete grid S of cardinal N. Grey levels take values in the discrete set $[0, L-1]$, and we denote by $u_s \in [0, L-1]$ the label value of the image u at the site $s \in S$. We assume that the grid is endowed with a neighborhood system and we denote by $s \sim t$ the neighboring relationship between s and t and by (s,t) the second order clique. In this paper, only pairwise interactions are considered. We consider the decomposition of an image into its level sets using the decomposition principle [13, 12]. In other words we consider all thresholded images u^λ where $u_s^\lambda = \mathbb{1}_{u_s \leq \lambda}$ $\forall s \in S$. The original image u can reconstructed via the formula $u_s = \min\{\lambda, u_s^\lambda = 1\}\ \forall s \in S$.

A function is said *levelable* if and only if it can be rewritten as a sum on level sets, of functions of its variable level-sets. Since in this paper we only cope with MRFs with pairwise interactions, we only give the form a levelable function for functions of one and two variables. A function of one variable, $f : [0, L-1] \mapsto \mathbb{R}$, is always levelable since we have:

$$\forall u_s \in [0, L-1] \;\; f(u_s) = \sum_{\lambda=0}^{L-1} \left(f(\lambda+1) - f(\lambda)\right)(1 - u_s^\lambda) + f(0) \;\; .$$

From a Markovian point of view, data fidelity terms are functions of one variable. Next we consider functions of two variables that will corresponds to our priors. In it shown in [9] that a levelable symmetric function g of two variables, $g : [0, L-1]^2 \mapsto \mathbb{R}$, necessarily takes the following form

$$g(x, y) = F(\max(x, y)) - G(\min(x, y)) \;\; ,$$

where F and G are some functions that map $[0, L-1]$ to $\mathbb{R}$. Besides, if we also assume that $\forall y \in [0, L-1] \;\; g(\cdot, y)$ attains a minimum at y, then g takes the following form:

$$\begin{aligned} g(x, y) &= |S(x) - S(y)| + D(x) + D(y) \;\; , \\ &= \sum_{\lambda=0}^{L-2} R(\lambda)|\mathbb{1}_{\lambda<x} - \mathbb{1}_{\lambda<y}| + D(x) + D(y) \;\; , \end{aligned}$$

where $R(\lambda) = S(\lambda+1) - S(\lambda)$ is a *nonnegative* function on $[0, L-2]$ and where D is some mapping from $[0, L-1]$ to $\mathbb{R}$. Note that the non-negativeness of R implies that S is a non-decreasing function.

In the sequel we always assign $D \equiv 0$ and we say in this case that $g(x, y) = |S(x) - S(y)|$, with S *non-decreasing* is a *nice-levelable* function. It is thus immediate that the Total Variation [21], which corresponds to $g(x, y) = |x - y|$ and $S(x) = x$, is a nice-levelable function.

A levelable (resp. nice-levelable) Markov Random Field is a Markov random field whose pairwise interaction terms are levelable (resp. nice-levelable) functions. Although a (nice) levelable Markovian energy is generally not convex, a global minimizer can be computed by mapping the problem to a binary submodular function minimization (for which efficient algorithms are available). We refer the reader to [9] for details on this minimization. We also refer the reader to the work of Zalesky in [24] for levelable MRFs which involve higher order interaction terms. We are now ready to study the minimizers of a levelable Markovian energy.

3 A Theorem for the Shape of Restored Objects

In this section, we assume that we observe an image v corrupted by some noise and the restored version of v is referred to as u. We consider any *nice-levelable* posterior restoration energy (see previous section) so that it takes the following form:

$$E(u|v) = \sum_s U(v_s|u_s) + \beta \sum_{(s,t)} |S(u_s) - S(u_t)| \;\; ,$$

where the fonction U measures the fidelity of the restored image u to the observed data v. For instance, in the additive Gaussian noise case we have $U(v_s|u_s) = \dfrac{(v_s - u_s)^2}{2\sigma^2}$. Recall that since we consider nice-levelable MRFs, the function S is *non-decreasing* on $[0, L-1]$ and that for the TV prior we have $S(\lambda) = \lambda \quad \forall \lambda \in [0, L-1]$. In the reminder of this paper we use the notation $E(x|y)$ so signify that E is a function of x while y is a parameter fixed to some value; formally we have $E(\cdot|y) = E(\cdot, y)$. We may use the notation $\phi_s(x) = U(v_s|x)$ which is implicit in v_s.

We now generalize the results of [8, 22, 23] and show that under reasonable conditions only the constrast of object changes and not its shape. Levelable regularization functions are thus needed to prevent the loss of contrast. Let us now consider a cartoon object $\mathcal{O}$ with perimeter $\mathcal{L}(\mathcal{O})$, area $\mathcal{S}(\mathcal{O})$ and original luminance A, lying in a background of original luminance B. The whole image is corrupted by some (not necessarily gaussian) noise. We consider the posterior MRF energy when first regularizing by TV. In the following we shall make use of a *continuous* analysis, most often concerning grey levels and sometimes concerning the topology. Also, in the following the cardinal of some set E will be noted $|E| = Card(E)$ if no confusion ensures. For instance we shall often write $\mathcal{S}(\mathcal{O}) \approx |\mathcal{O}|$ in the discrete lattice framework.

The next theorem gives some sufficient conditions on the observed object and the Markovian energy so that the *shape* of the object in the result is preserved.

Theorem 1 *If the following conditions are met:*

- ***Assumption 1*** $\forall s \in S, \forall v_s \in \mathbb{R}$, *the attachment to data energy term*

$$\phi_s(\mu) = U(v_s \mid u_s = \mu) \text{ is minimal for } \mu = v_s \quad .$$

- ***Assumption 2*** *Moreover,* $\forall s \in S, \forall v_s \in \mathbb{R}$, $\phi_s(\mu)$ *is a* quasi-convex *function of parameter* μ *(see Appendix A Definition 1).*

- ***Assumption 3*** *The original (resp. restored) image are piecewise constant and verify:*

background: brilliance B *(resp.* b*)*	-	*object: brilliance* A *(resp.* a*)*
$B \leq b$	$\leq$	$a \leq A$

Then:
i) If the object $\mathcal{O}$ *to be restored is* convex, *then consider the class of all homothecies with power* $\lambda \gtrless 1$ *whose center is* interior *to the object. If both object and background sizes are statistically "significant", then either the object disappears, or the MRF posterior energy is minimal for* $\lambda = 1$ *i.e., the shape of the object is preserved through MRF restoration, whatever the* nice-levelable *regularization function employed.*

ii) Accordingly, the posterior energy of original shape position is minimal wrt. all candidate translations of original object $\mathcal{O}$ *(whatever its shape).*

Please, note that assumptions 1 and 2 apply for Nakagami, Gamma and Gaussian laws endowed with their usual parameter.

Proof: we proceed along the same line than [22, 23].

i) Minimization wrt. homothecies

Let us note $\mathcal{O}_\lambda = H_\lambda\ \mathcal{O}$ the candidate restoration object, supposed to be obtained from original object $\mathcal{O}$ by homothecy H_λ. We define now a methodology for computing the total attachment to data contribution to posterior energy

$$\mathcal{U} = \sum_{s \in S} U(v_s \mid u_s)$$

Of course it decomposes always as

$$\mathcal{U} = \sum_{s \in \mathcal{O}_\lambda} U(v_s \mid u_s = a) \ + \sum_{s \in S \setminus \mathcal{O}_\lambda} U(v_s \mid u_s = b) \tag{1}$$

Now, in each of these two terms, some observation random variables v_s are either emitted (we say "drawn", see below) by $u_s = A$ (if $s \in \mathcal{O}$), or by $u_s = B$ (if $s \in S \setminus \mathcal{O}$). Thus for homothecies it appears that two cases have to be investigated:

Case I) $\lambda \geq 1$: original object included in restored object $\mathcal{O} \subset \mathcal{O}_\lambda$

We split the first term of previous formula into two parts, yielding

$$\mathcal{U} \ = \underbrace{\sum_{s \in \mathcal{O}} U(v_s \mid u_s = a)}_{1)} \ + \underbrace{\sum_{s \in \mathcal{O}_\lambda \setminus \mathcal{O}} U(v_s \mid u_s = a)}_{2)} \ + \underbrace{\sum_{s \in S \setminus \mathcal{O}_\lambda} U(v_s \mid u_s = b)}_{3)}$$

since as a matter of fact (see Fig. 1 left part):

1) $\mathcal{O}$	cardinal: $\lvert\mathcal{O}\rvert$	is drawn from $P(\cdot \mid \mu = A)$ i.e., A-drawn (see Appendix A)	.
2) $\mathcal{O}_\lambda \setminus \mathcal{O}$	cardinal: $\lvert\mathcal{O}_\lambda\rvert - \lvert\mathcal{O}\rvert$	is drawn from $P(\cdot \mid \mu = B)$ i.e., B-drawn (")	.
3) $S \setminus \mathcal{O}_\lambda$	cardinal: $N - \lvert\mathcal{O}_\lambda\rvert$	is B-drawn	.

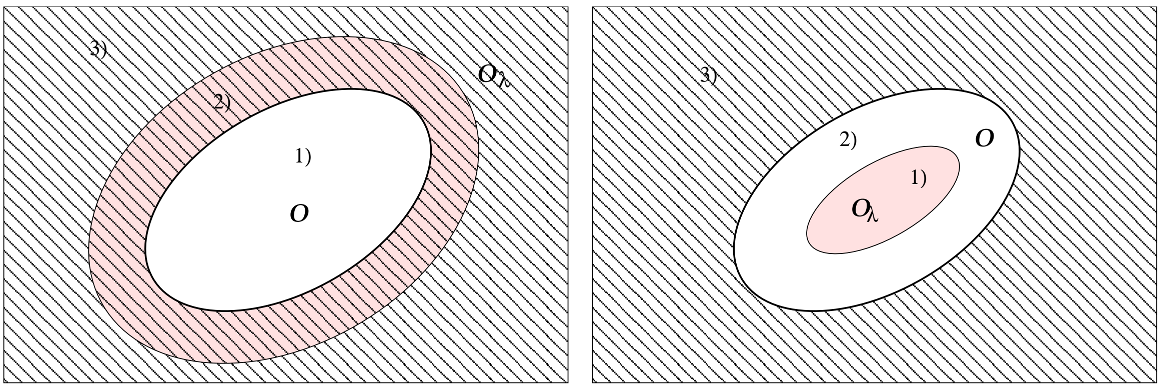


Figure 1: Homothecies. Left : case I) - Right : case II) .

Thanks to Proposition 3 of Appendix A the expression of $\mathcal{U}$ writes then approximately:

$$\mathcal{U} \approx |\mathcal{O}|\, U(v_s = A \mid \mu = a) \ + \ (|\mathcal{O}_\lambda| - |\mathcal{O}|)\, U(v_s = B \mid \mu = a) \ + \ (N - |\mathcal{O}_\lambda|)\, U(v_s = B \mid \mu = b)$$

so that the total posterior energy noted $E(\lambda)$ is approximately:

$$\begin{aligned} E(\lambda) &= E(u \mid v) \approx \\ &|\mathcal{O}|\, U(v_s = A \mid \mu = a) \;+\; (|\mathcal{O}_\lambda| - |\mathcal{O}|)\, U(v_s = B \mid \mu = a) \;+\; (N - |\mathcal{O}_\lambda|)\, U(v_s = B \mid \mu = b) \\ &+ \beta\, \mathcal{L}(\mathcal{O}_\lambda)\, (S(a) - S(b)) \quad (a \geq b) \ , \end{aligned}$$

Here we make a continuous approximation for topology by setting: $|\mathcal{O}_\lambda| \approx |\mathcal{O}|\ \lambda^2$ and $\mathcal{L}(\mathcal{O}_\lambda) \approx \lambda\, \mathcal{L}(\mathcal{O})$. Thus the "quadratic term in λ" in previous energy formula is

$$|\mathcal{O}_\lambda|\ [\, U(v_s = B \mid \mu = a) - U(v_s = B \mid \mu = b)\,]$$

It appears that this is a *positive* term since

$$U(v_s = B \mid \mu = a) - U(v_s = B \mid \mu = b) \geq 0 \ .$$

The latter inequality results indeed from the *quasi-convex* Hypothesis 2 and from the piecewise ordered Hypothesis 3. See Fig. 2 - left for an illustration.
Also, the linear term in λ which corresponds to regularization is almost by definition, positive and thus a non-decreasing function of λ.
Thus, as in Strong et al., the second-order polynomial (in λ) $E(\lambda)$ is *convex non-decreasing* for $\lambda \geq 1$. See right part of Fig. 3 .

Case II) $\lambda \leq 1$: restored object included in original object $\mathcal{O}_\lambda \subset \mathcal{O}$

Using the same approach we find that in this case the attachment to data contribution writes

$$\begin{aligned} \mathcal{U} &= \underbrace{\sum_{s \in \mathcal{O}_\lambda} U(v_s \mid u_s = a)}_{1)} + \underbrace{\sum_{s \in \mathcal{O} \setminus \mathcal{O}_\lambda} U(v_s \mid u_s = b)}_{2)} + \underbrace{\sum_{s \in S \setminus \mathcal{O}} U(v_s \mid u_s = b)}_{3)} \\ &\approx |\mathcal{O}_\lambda|\, U(v_s = A \mid \mu = a) + (|\mathcal{O}| - |\mathcal{O}_\lambda|)\, U(v_s = A \mid \mu = b) + (N - |\mathcal{O}|)\, U(v_s = B \mid \mu = b) \end{aligned}$$

since here (see Fig. 1 right part) :

1) $\mathcal{O}_\lambda$ is A-drawn, tested for $u_s = a$.
2) $\mathcal{O} \setminus \mathcal{O}_\lambda$ is A-drawn, tested for $u_s = b$.
3) $S \setminus \mathcal{O}$ is B-drawn, tested for $u_s = b$.

This quadratic term in λ (we make the same topological approximation as above) writes thus:

$$|\mathcal{O}_\lambda|\ [\, U(v_s = A \mid \mu = a) - U(v_s = A \mid \mu = b)\,] \leq 0$$

by invoking the same quasy-convexity and piecewise ordered hypotheses as above. See Fig. 2 - right for an illustration. Thus the second-order polynomial $E(\lambda) = E(u \mid v)$ is *concave* for $0 \leq \lambda \leq 1$. Two possibilities occur at this point:

- a) $E(\lambda = 0) < E(\lambda = 1)$: $E(\lambda)$ is minimal for $\lambda = 0$ i.e., the object disappears completely!
- b) $E(\lambda = 0) > E(\lambda = 1)$: $E(\lambda)$ is minimal for $\lambda = 1$ i.e., the shape of the object is recovered!

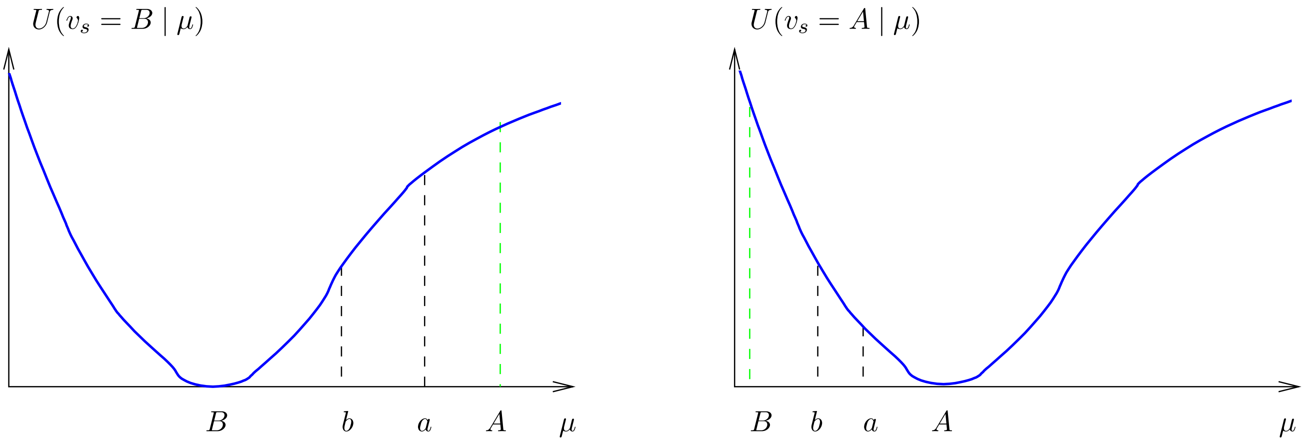


Figure 2: The quasi-convex behaviour of $U(v_s \mid \mu)$ and its consequence. Left : case I) - Right : case II) .

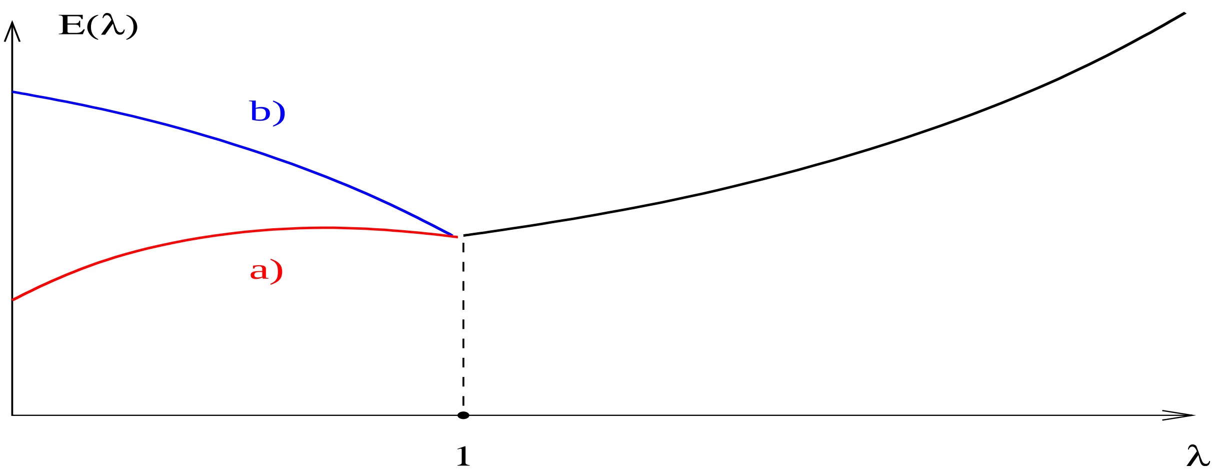


Figure 3: Sketch of the posterior energy $E(\lambda)$ as a function of the homothecy ratio λ .

We find the same "concave-convex" behaviour as Strong *et al.* as depicted in Fig. 3. This concludes the proof. □

Of course, some lack of accuracy of this development occurs around $\lambda = 0$ (resp. $\lambda = 1$), where $|\mathcal{O}_\lambda|$ (resp. $|\mathcal{O}| - |\mathcal{O}_\lambda|$) are "statistically small". Thus anything concerning the precise shape of the recovered object can happen around these ranges. Anyway we shall assume in the sequel that the theoretical conditions of Theorem 1 are met.

ii) Minimization wrt. translations

The same arguments as above apply to the set of translations of object $\mathcal{O}$. Let $\mathcal{O}_t$ be the translated candidate restored object and $D = \mathcal{O} \setminus (\mathcal{O} \cap \mathcal{O}_t)$, with $0 \leq |D| \leq |\mathcal{O}|$.
Now S can be decomposed in four subsets (see Fig. 4) :

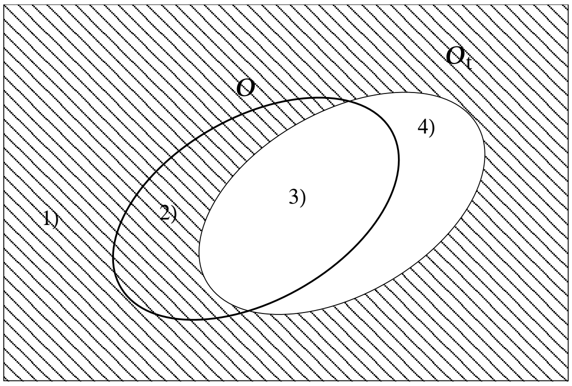


Figure 4: The case of translations.

1)	$S \setminus (\mathcal{O} \cup \mathcal{O}_t)$	cardinal:	$N - \|O\| - \|D\|$	B-drawn	tested for $u_s = b$
2)	$\mathcal{O} \setminus (\mathcal{O} \cap \mathcal{O}_t)$		$\|D\|$	A-drawn	tested for $u_s = b$
3)	$\mathcal{O} \cap \mathcal{O}_t$		$\|\mathcal{O}\| - \|D\|$	A-drawn	tested for $u_s = a$
4)	$\mathcal{O}_t \setminus (\mathcal{O} \cap \mathcal{O}_t)$		$\|D\|$	B-drawn	tested for $u_s = a$

Since the regularization component of posterior energy is translation invariant (!) we just cope with the attachment to data energy, and make use of previous statistical arguments:

$$\begin{aligned} \mathcal{U} \approx \quad & \underset{1)}{(N - |O| - |D|)\, U(v_s = B \mid \mu = b)} + \underset{2)}{|D| \quad U(v_s = A \mid \mu = b)} \\ & + \underset{3)}{(O| - |D|)\, U(v_s = A \mid \mu = a)} + \underset{4)}{|D| \quad U(v_s = B \mid \mu = a)} \end{aligned}$$

The linear component in $|D|$ of this expression is thus

$$|D| \; [\, U(v_s = A \mid \mu = b) - U(v_s = A \mid \mu = a) + U(v_s = B \mid \mu = a) - U(v_s = B \mid \mu = b) \,] \quad .$$

The sum of two first terms as well as that of the two last ones is *positive* from the quasi-convexity hypothesis and $B \leq b \leq a \leq A$.
The posterior energy is thus minimum for $|D| = 0$. □

Once again this statistical-based demonstration is no more valid for ”small” translations of the object. Anyway we shall assume that the restored object does not move at all.

We would also like to emphasize that this proof holds for *any* nice-levelable priors. In the remainder of this paper, we assume that conditions of theorem 1 are fulfilled and furthermore that the shape of the object is completely recovered.

4 A Theorem for the Brilliance of Restored Objects

We have cope with the shape of the restored object in the previous section and now we study its gray level value. The following theorem explains how there is necessarily a loss of contrast when one is using a suitable levelable prior to do the filtering.

Theorem 2 *If the requirements and results of Theorem 1 hold, namely:*

- **Hypotheses 1 and 2***: $\forall v_s$ attachment to data energy $\phi_s(\mu) = U(v_s|\mu)$ is a quasi-convex function of μ, (Appendix A Definition 1) and attains its minimum at $\mu = v_s$.*
- **Hypothesis** *3 (***mild***): the brilliance of piecewise constant restored image satisfies $b \leq a$.*
- **Theorem 1***: the shape and position of the object $\mathcal{O}$ are preserved.*

Then, regularizing with a nice-levelable pairwise energy implies that the brilliance of object decreases whereas that of background increases: $B \leq b \leq a \leq A$: i.e., **Hypothesis** *3 (***strong***) holds.*

Proof: this Theorem can be proved either in a continuous or even in a discrete grey-level framework. The total posterior energy, noted $E(a,b) = E(u \mid v)$ writes indeed:

$$E(a,b) = \sum_{s\in\mathcal{O}} U(v_s \mid \mu = a) \ + \sum_{s\in S\setminus\mathcal{O}} U(v_s \mid \mu = b) \ + \ \beta\,\mathcal{L}(\mathcal{O})\,(S(a) - S(b)) \ \ (a \geq b)$$

From the statistical hypothesis that both object and backgound sizes are large, this writes as:

$$E(a,b) \approx |\mathcal{O}|\,U(v_s = A \mid \mu = a) \ + \ (N - |\mathcal{O}|)\,U(v_s = B \mid \mu = b) \ + \ \beta\,\mathcal{L}(\mathcal{O})\,(S(a) - S(b))$$

Thus for fixed b the total energy term wrt. variable a writes:

$$E(a) \approx |\mathcal{O}|\,U(v_s = A \mid \mu = a) \ + \ \beta\,\mathcal{L}(\mathcal{O})\,S(a) \ \ (a \geq b) \tag{2}$$

From the quasi-convex hypothesis + the levelable hypothesis ($S(\cdot)$ is a *non-decreasing* function), this is a *non-decreasing* function of a for $a \geq A$. Thus the minimizer value a^* verifies $b \leq a^* \leq A$.
Conversely for a fixed the total energy term wrt. variable b writes as

$$E(b) \approx (N - |\mathcal{O}|)\,U(v_s = B \mid \mu = b) - \beta\,S(b) \ \ (b \leq a) \tag{3}$$

Using the same arguments as above this is a *non-increasing* function of b for $b \leq B$. Thus the minimizer value b^* verifies $a \geq b^* \geq B$. This concludes the proof. □

We propose in the next section a new approach to circumvent this loss of contrast obtained using a modified TV prior.

5 Why do we need levelable regularization energies?

Let us apply previous results to the usual Gaussian noise case. Previous equation 2 for the posterior energy of candidate restored object $\mathcal{O}$ with brilliance a writes then

$$E(a) = E(u \mid v) = \mathcal{S}(\mathcal{O})\, \frac{(A-a)^2}{2\sigma^2} + \beta\, \mathcal{L}(\mathcal{O})\, S(a) \quad (S(a) > 0)$$

Now, in the *continuous* grey level framework the following typical loss of brilliance is found by minimizing $E(a)$ wrt. a, i.e. by setting $\frac{\partial E}{\partial a} = 0$:

$$a^* - A = -\frac{\mathcal{L}(\mathcal{O})}{\mathcal{S}(\mathcal{O})}\, \sigma^2\, \beta \left(\frac{\partial S}{\partial a}\right)_{a^*} \quad (\left(\frac{\partial S}{\partial a}\right)_a = 1\ \forall a \text{ for } TV) \ .$$

The magnitude of this contrast loss will be lower if the "effective" regularization parameter at grey level A, namely $(\frac{\partial S}{\partial a})_{a=A}$ is low ! We are thus set between two contradictory objectives: regularization and contrast preservation. Thus we design and use an adapted levelable function with low (discrete) "slope" $R(\lambda) = S(\lambda+1) - S(\lambda)$ for each convenient grey level values $\lambda = A$ to be recovered!
This approach can be generalized to other types of noise as Gamma and Nakagami laws for instance. We present just an outline for this purpose: assume that minimizer $a^* \approx A$. Then

$$\left(\frac{\partial U(A \mid a)}{\partial a}\right)_{a^*} \approx \underbrace{\left(\frac{\partial U(A \mid a)}{\partial a}\right)_A}_{0} + (a^* - A) \left(\frac{\partial^2 U(A \mid a)}{\partial a^2}\right)_A$$

The minimizer value a^* is thus given by

$$a^* - A \approx \frac{\mathcal{L}(\mathcal{O})}{\mathcal{S}(\mathcal{O})}\, \beta \left(\frac{\partial S}{\partial a}\right)_A \Big/ \left(\frac{\partial^2 U(A \mid a)}{\partial a^2}\right)_A$$

It remains to show that indeed $a^* \approx A$, and also that a similar reasoning holds for background (which is more likely since its size is usually quite larger than that of the object itself).

Recall that in this section we have assumed that S is a differentiable function and that $U(A|\cdot)$ is twice differentiable. Since a levelable MRF is defined on a finite set of labels, the above consideration does not apply directly. However, although in this paper we have assumed that the set of label is the discrete set $\{0, L-1\}$, one can chose an arbitrary fine quantization of the continuous segment $[0, L-1]$, i.e., $\{0, \delta, \ldots, L-\delta\}$ with $\delta > 0$ and thus getting a fine approximation of the first and second derivatives using classical finite difference schemes.

6 Experiments

In this section we present some results on synthetic images corrupted by Gaussian additive noise or speckle noise.

6.1 $L^2 + TV$

First we investigated the validity of previous developments on the usual $L^2 + TV$ model: a circle was created with diameter $D = 40$, brilliance of background (resp. object) $\mu_1 = 60$ (resp. $\mu_2 = 80$) . Gaussian noise was then added with standard deviation $\sigma = 30$ i.e., similar to that of a Rayleigh distribution for these mean values. In our experiment the levelable function is prescribed as $R(\lambda) = S(\lambda + 1) - S(\lambda) = 0.01$ for both $\lambda = \lambda_1 = 59$ and $\lambda = \lambda_2 = 79$, whereas $R(\lambda) = 1 \ \ \forall \lambda \neq \lambda_1, \lambda_2$ as for TV ! Comparison of results with standard TV is shown on Figs. 5 and 6. We clearly see on Fig. 6 that the minimization using the latter nice-levelable function achieves both noise removal and contrast preservation. This is to compare to TV regularization which only successes in noise removal, as predicted by the theory.

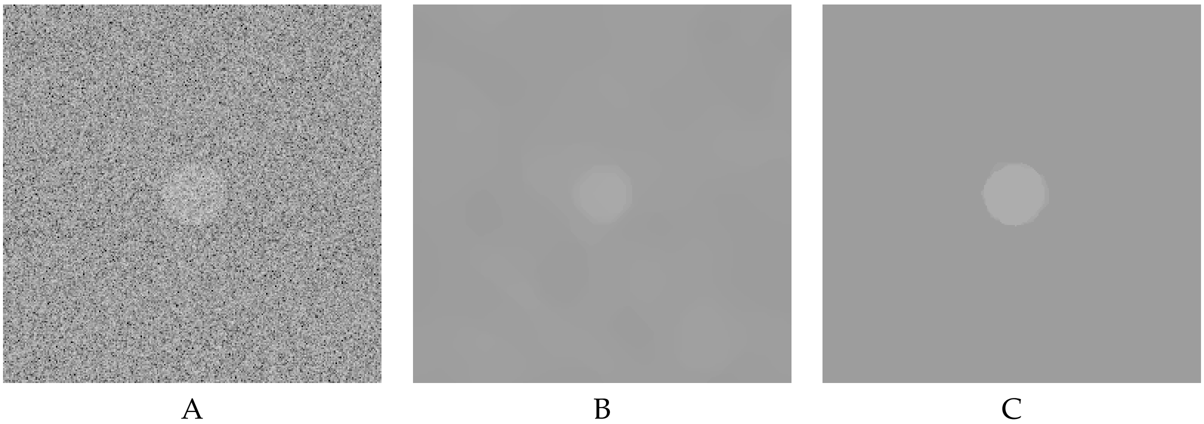


Figure 5: A: original noisy image (Gaussian noise) - B: result with TV - C: result with adapted levelable function.

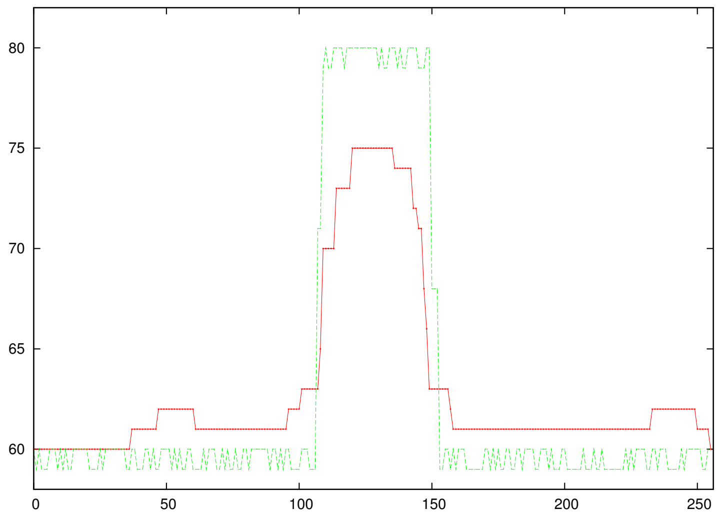


Figure 6: Slices of the noisy and restored images at vertical line $x = 128$. Red: TV regularization - Green: levelable regularization (L^2+ modified TV).

6.2 Rayleigh + TV model

We now generalize previous effect to M-look speckled SAR images following a Nakagami law [11]:

$$E(v_s|u_s) = M\ [\frac{v_s^2}{u_s^2} + 2\ \log u_s]$$

To this end we synthetize a mire with original grey levels 20, 40, 60 and 80 on which we superimpose a Nakagami law of parameter $M = 1$ (this is a Rayleigh distribution). This noisy image is depicted on figure 7-A. The restored images using TV and adapted levelable regularization are respectively presented in Figure 7-B and -C. Note that for visualization purposes, we have applied a change of contrast on these images. Although both results are visually very similar, the effect of adaptive levelable regularization is clearly seen on figure 8: contrast is better preserved, while still removing noise.

The size of the images in these experiments in 256×256. Although the minimization method described in [9] may require a huge amount of memory (almost 3 Gigabytes for the images in this report) it takes only about 20 seconds a Pentium 4 3GHz to perform the optimization. Recall that the obtained minimizer are *exact* although the functional is not convex. These time results are much lower than the ones presented in [9] (using exactly the same algorithm) which restore images corrupted by impulsive noise using TV as a prior.We conjecture that this behavior is due to the fact the the functionals we minimize in this paper are somehow more ”convex” than the one used in [9]. This behavior is currently under investigation. Approximate energy minimization approaches for these problems which require much less memory will be presented in a forthcoming paper.

7 Conclusion

In this paper we first presented a statistical-based extension of [22, 23] concerning the shape conservation and loss of contrast for piecewice-constant restored images with Total Variation and general noise such as speckle. We then showed how a judicious use of levelable regularization functions i.e., decomposable on level sets [9] can overcome this loss of contrast effect, and applied this formalism to the denoising of Synthetic Aperture Radar (SAR) images while preserving the reflectivity of each region of interest. Preliminary results are very promising. A main issue is how to estimate automatically the levelable functions. This point will be addressed in a forthcoming paper.

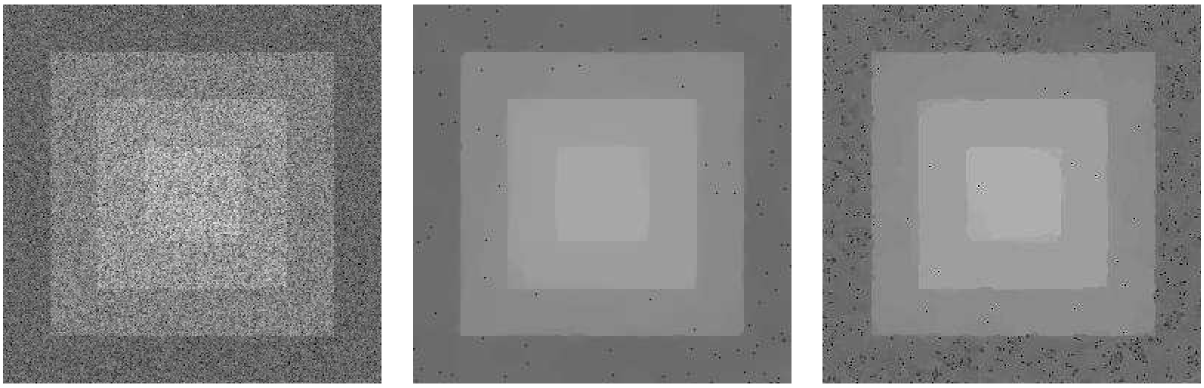
Appendix A: recall on sufficient statistics and exponential families

In this Appendix we sketch our definitions and notations and recall the main properties of sufficient statistics [4] in the case of exponential families, which is well adapted to the MRF approach.

Definitions and notations

Definition 1 *A* quasi-convex *function of several variables is a function whose level-sets are convex.*

In one-dimension, this means that the function is non-increasing, attains its minimal value and then increases.



A B C

Figure 7: A: original noisy image (Rayleigh speckle noise)- B: result with TV - C: result with adapted levelable function.

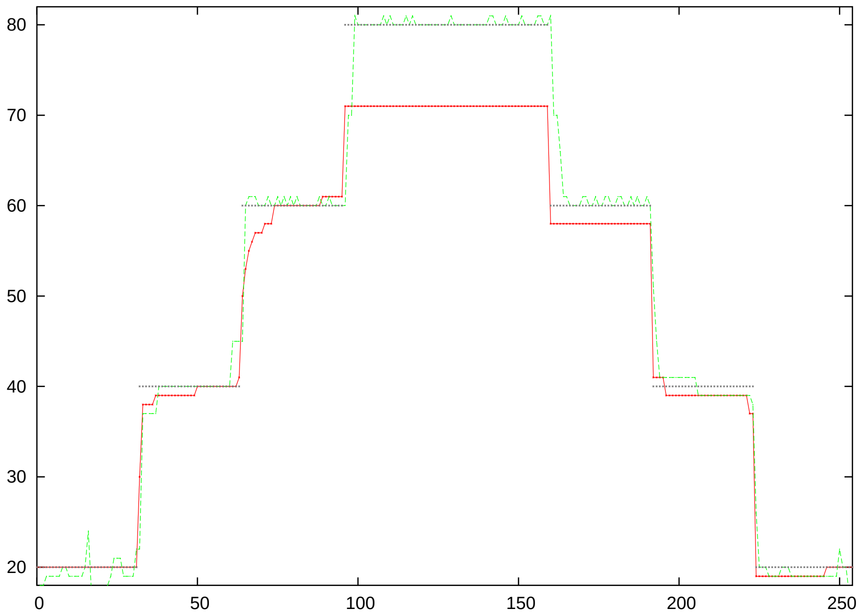


Figure 8: Slices of the noisy and restored images at horizontal line $y = 128$. Grey-Blue: non-noisy image - Red: TV - Green: levelable regularization (Rayleigh + modified TV).

Definition 2 *Expectation of random variable X under parametric distribution $P_a(\cdot)$ is noted :* $\mathbb{E}_a\left[X\right]$

Definition 3 *A subset $E \subset S$ is said to be A-drawn if $(V_1, \ldots V_s)_{s\in E}$ is drawn according to the conditional law*

$$P_A(V1, \ldots V_s) = P(V1, \ldots V_s \mid \mu = A) = P(V1, \ldots V_s \mid u_1 = \ldots u_s = A) \tag{4}$$

Sufficient statistics and exponential families for MRF observation pdf's

We address a μ-drawn subset $E \subset S$ with cardinal $|E| = Card(E)$ and write:

$$\begin{aligned} P(V_1 = v_1 \ldots V_s = v_s \mid \mu) &= h(v_1 \ldots v_s)\ \exp - [\ \chi(\mu)\ T(v_1 \ldots v_s) + \Gamma(\mu)\ ] \\ &\propto \ \exp -U(v_1 \ldots v_s \mid u_1 = \ldots = u_s = \mu) \end{aligned} \tag{5}$$

with

$$U(v_1 \ldots v_s \mid u_1 = \ldots = u_s = \mu) = \chi(\mu)\ T(v_1 \ldots v_s) + \Gamma(\mu)\ .$$

Here $T(v_1 \ldots v_s)$ is the sufficient statistics associated to $P_\mu(V1, \ldots V_s)$. In the following we shall *always* assume that the random variables $V_1 \ldots V_s$ are i.i.d for sake of simplicity, which corresponds to conditional independance of observations in the MRF framework. Thus:

$$\begin{aligned} T(V_1 \ldots V_s) &= \sum_{s\in E} T(V_s) \text{ and} \\ U(v_1 \ldots v_s \mid u_1 = \ldots = u_s = \mu) &= \sum_{s\in E} U(v_s \mid u_s = \mu) && (6) \\ &= \chi(\mu)\ \left(\sum_{s\in E} T(V_s)\right) + \Gamma(\mu) && (7) \\ \text{with } \Gamma(\mu) &= |E|\ \gamma(\mu)\ . \end{aligned}$$

For instance in the Nakagami law, one has:

$$\begin{aligned} T(v_1 \ldots v_s) = \sum_{s\in E} v_s^2, \qquad & h(v1 \ldots v_s) \propto \prod_{s\in E} v_s^{2M-1} \\ \chi(\mu) = \frac{M}{\mu^2}, \qquad & \Gamma(\mu) = |E|\ \gamma(\mu) \text{ with } \gamma(\mu) = 2M\ \log(\mu)\ . \end{aligned}$$

Also, definition of the MRF observation energy in last equation (7) is coherent, since from a MRF point of view v_s is observed and fixed.
Parameter μ in (5) can be defined up to a monotone function change. In the sequel, we shall make the fundamental assumption that parameter μ has the following precise, physical meaning:

Hypothesis 4 *For any μ-drawn subset $E \subset S$ and $\forall A \in \mathbb{R}$ fixed, the likelihood*

$$L_E(\mu) = P(V_1 = \ldots = V_s = A \mid \mu) = P(V_s = A \mid \mu)^{|E|}$$

is maximal at $\mu = A$. In a MRF context, this means that the attachment to data energy $\sum_{s\in E} U(v_s = A \mid u_s = \mu)$ *is minimal for $\mu = A$.*

This is in fact Hypothesis 1 of this paper.
We are now equipped to state the two following Propositions:

Proposition 1 *$\forall A \in \mathbb{R}$ and for any A-drawn subset$E \subset S$:*

$$\mathbb{E}_A\left[T(V_1 \ldots V_s)\right] = T(V_1 = \ldots V_s = A)$$

Proof: from the ML Hypothesis 4 one has

$$\left(\frac{\partial \displaystyle\sum_{s \in E} U(V_s = A \mid \mu)}{\partial \mu}\right)_{\mu=A} = \left(\frac{\partial \chi(\mu)}{\partial \mu}\right)_{\mu=A} T(V_1 = \ldots V_s = A) + \left(\frac{\partial \Gamma(\mu)}{\partial \mu}\right)_{\mu=A} = 0 \quad (8)$$

On the other hand, a classical result of Probability theory establishes that for any parametric pdf $P_\mu(\cdot)$:

$$\mathbb{E}_\mu\left[\frac{\partial \log P_\mu(V_1 \ldots V_s)}{\partial \mu}\right] = 0 \quad \forall \mu \in \mathbb{R}$$

In our case this writes as:

$$\mathbb{E}_\mu\left[\frac{\partial \displaystyle\sum_{s \in E} U(V_s \mid \mu)}{\partial \mu}\right] = \frac{\partial \chi(\mu)}{\partial \mu}\, \mathbb{E}_\mu\left[T(V_1 \ldots V_s)\right] + \frac{\partial \Gamma(\mu)}{\partial \mu} = 0 \quad \forall \mu \in \mathbb{R} \quad (9)$$

Now, setting $\mu = A$ in this formula and identifying with previous equation (8) establishes the result, provided that $\chi(\mu)$ is invertible at $\mu = A$ i.e. , $\left(\frac{\partial \chi(\mu)}{\partial \mu}\right)_{\mu=A} \neq 0$. □

Proposition 2 *Let $E \subset S$ be A-drawn, and $V_1 \ldots V_s$ i.i.d.*

$$Then \lim_{|E| \to +\infty} \frac{T(V_1 \ldots V_s)}{|E|} = \frac{\mathbb{E}_A\left[T(V_1 \ldots V_s)\right]}{|E|} = T(A)$$

Hence the name "sufficient statistics": for instance, estimator of parameter $\mu = A$ for the Nakagami law is given by $\hat{A}^2 = \left(\displaystyle\sum_{s \in E} V_s^2\right) / |E|$.

Proof: this relies immediately from the (weak) law of large numbers for i.i.d. random variables V_s and from Proposition 1. □

The next result, of significant physical interpretation, follows at once:

Proposition 3 *Let $E \subset S$ be A-drawn, and $V_1 \ldots V_s$ i.i.d. Then:*

$$\forall \mu \in \mathbb{R}, \lim_{|E| \to +\infty} \frac{\displaystyle\sum_{s \in E} U(v_s \mid u_s = \mu)}{|E|} = U(v_s = A \mid u_s = \mu)$$

Proof: indeed one has from (7):

$$\begin{aligned} Q &= \frac{\displaystyle\sum_{s \in E} U(v_s \mid u_s = \mu)}{|E|} = \chi(\mu) \frac{\displaystyle\sum_{s \in E} T(v_1 \ldots v_s)}{|E|} + \frac{\Gamma(\mu)}{|E|} \\ \lim_{|E| \to +\infty} Q &= \chi(\mu)\, T(A) + \gamma(\mu) = U(v_s = A \mid u_s = \mu) \end{aligned}$$

Acknowledgements

The authors would like to thanks Boris Zalesky for fruitful discussions. Jérôme Darbon's research is supported by the Office of Naval Research through grant ONR N00014-06-1-0345, the National Institute of Health through grant NIH U54-RR021813 and the National Science Foundation through grant NSF DMS-0610079.

References

[1] A. Achim, E. Kuruoglu, and J. Zerubia. SAR image filtering based on the heavy-tailed Rayleigh model. *IEEE Transactions on Image Processing*, 15(9):2686–2693, September 2006.

[2] H. H. Arsenault and M. Levesque. Combined homomorphic and local statistics processing for restoration of images degraded by signal-dependent noise. *Applied Optics*, 23(6):845–850, 1984.

[3] B. Berkels, M. Burgers, M. Droske, O. Nemitz, and M. Rumpf. Cartoon extraction based on anisotropic image classification. Technical report, UCLA, 2006.

[4] A. Borovkov. *Mathematical Statistics*. Gordon and Breach, 1998.

[5] M. Burger, G. Gilboa, S. Osher, and J. Xu. Nonlinear inverse scale space methods. *Communications in Mathematical Sciences*, 5(1):175–208, 2006.

[6] M. Burger and S. Osher. Convergence rates of convex variational regularization. *Inverse Problem*, 20:1411–1421, 2004.

[7] M. Burger, S. Osher, J. Xu, and G. Gilboa. Nonlinear inverse scale space methods for image restoration. In Springer-Verlag, editor, *Variational, Geometric, and Level Set Methods in Computer Vision (VLSM)*, volume 3752 of *Lecture Notes in Computer Science*, pages 25–36, 2005.

[8] T.F. Chan, S. Esedoglu, and M. Nikolova. Algorithms for Finding Global Minimizers of Image Segmentation and Denoising Models. *SIAM Journal on Applied Mathematics*, 66(5):1632–1648, 2006.

[9] J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation part II: Levelable functions, convex priors and non-convex cases. *Journal of Mathematical Imaging and Vision*, 2006.

[10] V.S. Frost, J. A. Stiles, K.S. Shanmugan, and J. C. Holtzman. A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-4(2):157–166, February 1982.

[11] J.W. Goodman. Some fundamental properties of speckle. *Journal Optical Society of America*, 66(11):1145–1150, 1976.

[12] F. Guichard and J.-M. Morel. *Image Iterative Smoothing and PDE s.* downloadable manuscript: please write email to fguichard@poseidon-tech.com, 2000.

[13] F. Guichard and J.-M. Morel. Mathematical morphology "almost everywhere". In *Proceedings of ISMM*, pages 293–303. Csiro Publishing, April 2002.

[14] J. S. Lee. Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2(2):165–168, February 1980.

[15] J. Lie and J.M. Nordbotten. Inverse scale space for nonlinear regularization. *Journal of Mathematical Imaging and Vision*, to appear., 2006.

[16] A. Lopes, R. Touzi, and E. Nezry. Adaptative speckle filters and scene heterogeneity. *IEEE Transactions on Geoscience and Remote Sensing*, 28(6):992–1000, 1990.

[17] A. Marquina. Inverse scale space methods for blind deconvolution. Technical Report 36, UCLA, 2006.

[18] Y. Meyer. *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*. American Mathematical Society, 2001.

[19] J-M. Nicolas, F. Tupin, and H. Maître. Smoothing speckled SAR images by using maximum homogeneous region filters: an improved approach. In *IGARSS*, volume 3, Sydney, Australia, July 2001.

[20] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An Iterative Regularization Method for Total Variation Based Image Restoration. *SIAM journal on Multiscale modeling and Application*, 4:460–489, 2005.

[21] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D.*, 60:259–268, 1992.

[22] D. Strong and T.F. Chan. Exact Solutions to Total Variation Regularization Problems. Technical report, CAM, UCLA, October 1996.

[23] D. Strong and T.F. Chan. Edge-Preserving and Scale-Dependent Properties of Total Variation Regularization. *Inverse Problems*, 9:S165–S187, November 2000.

[24] B.A. Zalesky `http://www.citebase.org/cgi-bin/citations?id=oai:arXiv.org:math/0304041`. Efficient Determination of Gibbs Estimators with Submodular Energy Functions, 2003.

[25] G. Winkler. *Image Analysis, Random Fields and Dynamic Monte Carlo Methods. A Mathematical Introduction*. Applications of mathematics. Springer-Verlag, 3rd edition, 2006.

[26] J. Xu and S. Osher. Iterative regularization and nonlinear inverse scale space applied to wavelet based denoising. *IEEE Transactions on Image Processing*, to appear., 2006.