A Finite Markov Random Field approach to fast edge-preserving image recovery

Michele Ceccarelli *

Research Center on Software Technologies – RCOST, University of Sannio, Via Traiano, 82100 Benevento, Italy

Received 22 April 2004; received in revised form 29 March 2006; accepted 16 May 2006

Abstract

We investigate the properties of edge-preserving smoothing in the context of Finite Markov Random Fields (FMRF). Our main result follows from the definition of discontinuity adaptive potential for FMRF which imposes to penalize linearly image gradients. This is in agreement with the Total Variation based regularization approach to image recovery and analysis. We also report a fast computational algorithm exploiting the finiteness of the field, it uses integer arithmetic and a gradient descent updating procedure. Numerical results on real images and comparisons with anisotropic diffusion and half-quadratic regularization are reported.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Markov random fields; Image denoising; Edge-preserving potentials

1. Introduction

The Bayesian framework is particularly suited for solving computer vision problems as it can embed in a unique model the data consistency constraints, observation model and a priori assumptions. The underlying probabilistic model is the Markov Random Field [9], and it has been successfully applied to several inverse imaging problems such as deconvolution, denoising, interpolation, segmentation, depth estimation, shape from shading and shape from texture. The ill-posed nature of these inverse imaging problems is typically treated by recurring to Gibbs priors encompassing both the uncertainty about the solution and the desirable characteristics it should have. The generic and most popular assumption regards the smoothness of the solution [9,21]. It tends to prefer solutions characterized by local coherence and homogeneity. However, it can lead, in many situations, to over smoothed solutions due to the imposition of the constraint everywhere in the image. Indeed, classical image restoration approaches are essentially based on the least squares criteria, which are basically linear and tend to smooth out edges in the output image. Therefore, the application of the smoothness constraint which preserve discontinuities has been one of the most active research areas in the computer vision community [7,15,17–19,23]. In particular, the concept of discontinuity adaptive prior (or edge-preserving regularization) [15] is becoming even more adopted also due to the availability of fast and accurate algorithms [7,25].

Here we show that the concept of discontinuity adaptive prior can be introduced even in the context of Finite Markov Random Fields (FMRF) where the underlying space of the solution is assumed to be finite. In particular, we classify a potential function as being edge preserving if it treats in the same way all the monotone functions in a given interval. This definition avoids to introduce the behavior of the potential function at the infinity and therefore it is more suited for FMRF. We show that in order to be edge preserving, a potential function should weight linearly the image gradient, in agreement with the recent approaches based on the Total Variation norm [5,19]. We also show how to develop a fast computational algorithm for exploiting the finiteness of the field, using integer arithmetic.
We focus on the problem of image denoising. In computer vision noise may refer to any entity, in images, data or intermediate results, that is not interesting for the purposes of the main computation [24]. For example, in image acquisition the effect of noise is that image values are not those expected, as these are corrupted during the various stages of image acquisition. As a consequence, the pixel values of two images of the same scene taken by the same camera and in the same light conditions are never exactly the same. Such fluctuations will introduce errors in the successive analysis steps. In this paper we restrict our attention with the assumption that the noise is random and additive, as this is a common assumption for natural images. Other types of noise, such as multiplicative speckle noise, which are typical of particular acquisition processing such as tomography and remote sensing require different techniques which are not covered here [26]. Impulsive noise also usually occurs during image acquisition, impulsive noise alters random pixels, making their values very different from the true values and very often from those of neighboring pixels too. Rank order filters and mathematical morphology [8] are often applied for efficiently approach this problem. Indeed, the assumption of additive random noise is widely accepted for natural images and it is yet the subject of many studies [19,25]. Here we will deal with uniform and Gaussian additive noise.

As far as the application fields where image recovery plays a fundamental role, let us mention various modalities and diagnostic fields such as functional nuclear medicine [3], ultrasound [26], remote sensing [1] and it is generally considered as a first preprocessing step in every image analysis workflow.

The paper is organized as follows, the next Section reports a brief introduction to the theory of Random Markov Fields applied to image reconstruction. Whereas in Section 3, we present the definition of edge-preserving potential for Finite Markov Random Fields and its properties. The last section is devoted to the experimental evaluation of the algorithm and its comparison to well-known denoising methods.

2. The MRF approach

Here we consider the problem of restoring an image corrupted by noise. Let \( I_{ij}^0 \), \( i = 1, \ldots, M \) and \( j = 1, \ldots, N \) an observed image and \( I_{ij} \) the “true” image, then our model is

\[
I_{ij}^0 = I_{ij} + n_{ij},
\]

where \( n_{ij} \) denotes the noise. This problem can be solved in the context of Bayesian paradigm. The goal is to estimate the image \( \hat{I} \) with the maximum a posteriori probability given \( I_0 \)

\[
\hat{I} = \arg \max \ p(I|I_0).
\]

According to the Bayes rule, the posterior \( p(I|I_0) \) is proportional to the product between the likelihood \( P(I|I_0) \), which depends on the noise distribution, and the prior \( P(I) \), which encompasses our a priori information about the true image. Usually, the likelihood distribution is assumed to be Gaussian:

\[
P(I_0|I) = e^{-|I-I_0|/2\sigma^2}/Z_\sigma.
\]

where \( Z_\sigma \) is a normalization factor which does not depend on \( I \). In addition, according to the Markov Random Field (MRF) model, the prior has a Gibbs distribution [16]

\[
P(I) = \frac{1}{Z} e^{-\mathcal{A}(I)}
\]

where \( \mathcal{A}(I) \) is the prior energy functional, it measures the “quality of the image” \( I \), in the sense that smaller values of \( \mathcal{A}(I) \) correspond to “better” images. \( \mathcal{A}(I) \) is the sum of local contribution form each image pixel. When there is no particular knowledge about the kind of images and the specific domain, the most natural assumption about \( I \) is its smoothness, therefore, \( \mathcal{A}(I) \) should be aimed at measuring the irregularities of the solution \( I \), such irregularities being naturally depend on the derivative magnitudes of \( I \).

Problem (2), by taking the logarithm and changing the sign, corresponds to the solution of

\[
\arg \min_I \{||I - I_0||^2 + \lambda \mathcal{A}(I)\}
\]

where \( \lambda \) is the so-called regularization parameter that controls the weight given to the minimization of the prior term with respect to the residual of the likelihood. Alternatively, if the noise variance is known or estimated, one can be interested in solving the corresponding constrained minimization problem

\[
\arg \min_I \mathcal{A}(I) \quad \text{subject to } ||I - I_0||^2 \leq \sigma
\]

Then the problem is to select the “best” image among those matching the constraints imposed by the statistics of the noise, and the problem (3) can be seen as the Lagrangian formulation of it with the parameter \( \lambda \) imposing the constraint.

Classical prior energy functionals are essentially based on the \( \| \|_2 \) norm of the gradient, which has the advantage of producing a set of linear equations to be satisfied by the solution. The main drawback in their use is that these functionals do not allow discontinuities in the solution, i.e. the edges are not well restored. Recently, people is even more interested in edge-preserving methods which produce much better results both from the perceptive point of view and in terms of signal-to-noise ratio. The price to be paid for these advantages is the solution of, sometimes complex, non-linear differential equation arising from the minimum condition of problem (4). In general, the prior energy has the form

\[
\mathcal{A}(I) = \sum_{i,j} \phi(|D_x I|) + \sum_{i,j} \phi(|D_y I|)
\]

where \( \phi \) is the potential function, \( D_x \) and \( D_y \) are the discretized derivative operators in the \( x \) and \( y \) directions:

\[
(D_x I)_{ij} = (I_{ij} - I_{i,j-1})/\delta_x \quad (D_y I)_{ij} = (I_{i,j} - I_{i,j-1})/\delta_y
\]
In order to be a suitable potential function, \( \phi \) should satisfy the following general assumptions:

(i) \( \phi(t) \geq 0 \), for any \( t \);
(ii) \( \phi(t) = \phi(-t) \);
(iii) \( \phi(t) \) is increasing for \( t \geq 0 \) and decreasing for \( t \leq 0 \).

In addition to these assumptions, a potential function \( \phi \) is considered edge preserving or discontinuity adaptive if it further satisfies [7, 15]

(iv) \( \lim_{t \to \infty} \frac{\phi(t)}{t} = 0 \);
(v) \( 0 \leq \lim_{t \to 0} \frac{\phi(t)}{t} < \infty \).

These conditions are quite natural in the context of images belonging to a continuous framework. However, in practice digital images have values over some finite set, such as \( \{0, \ldots , 255\} \). In such case the underlying image model is called Finite Markov Random Field (FMRF) representing the fact that \( I_{ij} \) can take only a finite set of values. In this context, the concept of infinity, of course, does not make sense, and condition (iv) just represents an ideal behavior. Therefore, successfully edge-preserving recovery algorithms should necessarily rely on some additional scale parameter representing thresholds which select candidate edges of the basis of gradient values which are above this threshold. In particular, the study reported in [15] classifies discontinuity adaptive potential functions in terms of the band, which is the interval where \( \phi'(x) > 0 \), outside this interval the penalty term does not depend on \( x \), it can be either zero (no smoothing) or constant as for example the so-called line process potential function [4].

A number of edge-preserving potential functions have been proposed in literature such as those reported in Table 1. These potential functions are typically parametrized by a parameter \( \gamma \) which allow to shrink or expand the band of the potential thus allowing a king of smooth threshold for the transition between uniform areas and candidate edges. This parameter being chosen as function of the image scale and the amount of edges one wants to consider inside the image.

Here we want to consider an alternative derivation of the energy potential function which does not depend on the specific values attained by each pixel and therefore it is suitable for FMRF. In particular, we will show that by taking into account the finite nature of the field, among the regularization functions reported in Table 1 the most suitable for FMRF is the simplest one \( |\phi(x) - x| \), which has recently received much attention, and is known as the Total Variation regularizer [19]. Our study is an improvement and extension of the model in [5], here we specialize this model for FMRF obtaining a very fast and efficient algorithm. We arrive at the model by considering that the condition (iv) is not suitable for Finite Markov Fields, and that a suitable adaptation not assuming that the domain of the field is unbounded.

### 3. The edge-preserving property for FMRF

In order to introduce the concept of discontinuity adaptive potential for FMRF let us consider the simple one-dimensional example plotted in Fig. 1 reporting two functions \( I^1 \) and \( I^2 \), the first containing an evident step discontinuity, the second being a smooth transition from a value to another. Let us consider a discretization of \( I^1 \) and \( I^2 \) as two sequences \( \{I^1_i\}_{i=0, \ldots , N} \) and \( \{I^2_i\}_{i=0, \ldots , N} \) with the same discretization step, then

\[
\mathcal{R}(I^1) = \sum_{i=1}^{N} \phi(I^1_i - I^1_{i-1}); \quad \mathcal{R}(I^2) = \sum_{i=1}^{N} \phi(I^2_i - I^2_{i-1}).
\]

If we want \( \phi \) to be an edge-preserving potential then \( I^1 \) should not be penalized more than \( I^2 \), in the sense that the solution of the problem (4) should not be biased toward \( I^2 \), this means that

\[
\mathcal{R}(I^1) \leq \mathcal{R}(I^2).
\]

This equation guarantees that sharp edges are preserved because the likelihood of solution \( I^1 \) is at least as much as that of \( I^2 \). In other words, this model does not prefer the smooth behavior of the second solution with respect to the sharp discontinuity of the first. However, if, on the contrary, \( \mathcal{R} \) is biased toward \( I^1 \) when applied to the image \( I^2 \), the solution of (4) will introduce artificial step discontinuities. This could be seen as the lack of the causality principle in the smoothing process aimed at solving problem (4), or equivalently that the smoothing behavior induced by such a potential can introduce artificial features during the regularization process. This last event is particularly disastrous in image recovery processes where the aim is to automatic analyze image contents. Therefore we also must have

\[
\mathcal{R}(I^1) \geq \mathcal{R}(I^2).
\]

From these last two inequalities we derive our definition:

**Definition 1.** A potential function \( \phi \) satisfying the general conditions (i)-(iii) is said FRMF edge preserving if given two monotonically increasing (decreasing) sequences \( \{I^1_i\}_{i=0, \ldots , N} \) and \( \{I^2_i\}_{i=0, \ldots , N} \) such that \( I^1_0 = I^2_0 \) and \( I^1_N = I^2_N \) it satisfies

\[
\sum_{i=1}^{N} \phi(I^1_i - I^1_{i-1}) = \sum_{i=1}^{N} \phi(I^2_i - I^2_{i-1}).
\]

#### Table 1

<table>
<thead>
<tr>
<th>( \phi(x) )</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{1}{2} (1 + r^2) )</td>
<td>[11]</td>
</tr>
<tr>
<td>( \log(1 + r^2) )</td>
<td>[13]</td>
</tr>
<tr>
<td>( \log \cosh \gamma r )</td>
<td>[12]</td>
</tr>
<tr>
<td>(</td>
<td>x</td>
</tr>
<tr>
<td>( \sqrt{1 + r^2} - 1 )</td>
<td>[7]</td>
</tr>
<tr>
<td>( \frac{1}{2} \phi(x) = \frac{1}{(1 + x^2)} )</td>
<td>[18]</td>
</tr>
<tr>
<td>( \min(x, y) )</td>
<td>[4]</td>
</tr>
</tbody>
</table>

...
Eq. (5) is suited for characterizing edge-preserving potentials for FMRF. Actually it does not make use of the behavior of \( \phi'(x) \) infinity, rather, it is based on the weight it gives to similar sequences which eventually contain abrupt changes representing edges. Neither it requires the choice, in terms of appropriate thresholds, of what a discontinuity is.

We try now to characterize the properties a function should satisfy in order to be a FMRF edge-preserving potential. The first consequence of our definition is that among the potential functions listed in Table 1 the Total variation norm [19] is FMRF edge preserving.

**Proposition 1.** A FMRF edge-preserving potential \( \phi(x) \) is a linear function of \( x \) for \( x \geq 0 \).

This result is, of course, not surprising. For example, several functions of Table 1 have a linear behavior at infinity such as \([12,19]\) and \([7]\). In addition to the edge-preserving property these three functions have the nice property of convexity, which is of help in the solution of (4). Indeed, one can show that convex edge-preserving potentials which are not bounded must necessarily satisfy

\[
0 < \lim_{i \to \infty} \frac{\phi(x)}{x} < \infty
\]  
(6)

\[i.e., convex edge-preserving potentials have a linear behavior at infinity. This can be easily shown by considering the rate of convergence of \( \phi(x) \) at infinity, it is defined as a real \( x \), such that

\[
0 < \lim_{i \to \infty} \frac{\phi(x)}{x^\alpha} < \infty
\]  
(7)

and proving that necessarily \( \alpha = 1 \). If \( \alpha > 1 \) then \( \phi(x) \) is convex but cannot be edge preserving. Indeed, if (7) is true, then also \( \lim_{i \to \infty} \frac{\phi(x)}{x^\alpha} \) is finite, and therefore \( \phi(x) \) cannot be edge preserving as condition (iv) cannot hold. If, on the contrary, \( \alpha < 1 \), the potential can be edge preserving but no more convex. Hence \( \alpha = 1 \).

Our derivation, however, states that in order to have an edge-preserving potential, this linear behavior should be always satisfied, clarifying what implicitly stated by condition (iv). In addition, as a consequence of property (ii) we have that

\[\phi(x) = |l(x)|\]

where \( l(x) \) is a linear function.

Given an FMRF edge-preserving potential we want to show how we can solve problem (4) with a fast and efficient algorithm. Let us call \( \mathcal{G} = \{g_0, \ldots, g_{L-1}\} \) the finite set where the image pixels take values, i.e. \( I_{i,j} \in \mathcal{G} \) and consider the minimum difference between two image values:

\[A = \min_{k \neq i} |g_k - g_i|\]

Here we develop a simple iterative algorithm, inspired by gradient descent, aimed at the minimization of the discrete functional \( \mathcal{R} \) with the given constraint by following the iterative scheme

\[
I_{i,j}^{n+1} = I_{i,j}^n + \Lambda \cdot \text{sign}(I_{i,j}^n - I_{i,j}^n) \\
+ \text{sign}(I_{i,j}^n - I_{i,j}^n) - \text{sign}(I_{i,j}^n - I_{i,j}^n))
\]  
(8)

this scheme is iterated while \( \|I^{n+1} - I^n\| \leq \sigma \) is true. Since \( \mathcal{R}(I) \geq 0 \), the following proposition states the convergence of the scheme.

**Proposition 2.** The sequence of potentials \( \mathcal{R}(I^n) \) generated by scheme (8) decreases monotonically.

The next proposition states the causality property, which is fundamental for every iterative smoothing process. Roughly speaking, the causality principle states that each feature at a coarse scale must have a cause at a finer scale. This means that the smoothing process does not introduce spurious features. Formally, it can be shown that every causal smoothing process must be governed by, or be the discretized version of, a parabolic partial differential equation obeying a maximum principle [2].

![Fig. 1. The function \( I^1 \) contains an abrupt change, whereas \( I^2 \) is a smooth transition from a value to another.](image-url)
Proposition 3. The scheme (8) satisfies
\[
\min \{P_{ij}^n, P_{i-1,j}^n, P_{i+1,j}^n, P_{i,j-1}^n, P_{i,j+1}^n\} \leq P_{ij}^{n+1} \\
\leq \max \{P_{ij}^n, P_{i-1,j}^n, P_{i+1,j}^n, P_{i,j-1}^n, P_{i,j+1}^n\}
\]

Note that the first sign operator in the scheme (8) was necessary just to ensure from one point that the numerical scheme was causal, and to use just integer arithmetic, bearing in mind that the set $\mathcal{Z}$ is finite. Other choices are of course possible, if we want to perform floating point operations, it is possible to choose time steps which guarantee the causality [5].

4. Experiments and comparisons

In this section we will present some experimental results about the application of the developed algorithm to synthetic and real grayscale. In particular the algorithm reads as:

**Digital Picture Recovery Algorithm**

//Input: a discrete image $I_{ij}$, $i = 1, \ldots, N$
//M and $j = 1, \ldots, N$
//Output: the recovered image

1. Estimate $\tilde{\sigma}$
2. $I_{ij} := I_{ij}^0$
3. $n := 0$
4. while $(\sum_{ij} (I_{ij}^n - I_{ij}^0)^2 \leq MN \cdot \sigma)$
5. Apply (8) to each image pixel
6. $n := n + 1$
7. end while
8. output $F$

Note that there are several methods to perform step 1. Our implementation adopts a variant of the method proposed in [14]. In particular,

$$\tilde{\sigma} = \frac{1}{36} \text{Variance} \begin{bmatrix} 1 & -2 & 1 \\ 2 & -4 & 2 \\ 1 & -2 & 1 \end{bmatrix}$$

where $\otimes$ represents the convolution operator. Note that the algorithm does not require any parameter. In our implementation we have the choice to implement filter (8) in a recursive manner, i.e., the updating is performed in place, this kind of updating produces a significant speed-up of the convergence while maintaining the causality properties of the method. In any case the experiments presented below are based on batch updating. Since we also report computing time, the adopted computing platform is significant, all the experiments were performed on a 600 MHz Pentium II Linux Personal Computer.

The first experiment we present is aimed at the evaluation of the edge-preserving ability of the algorithm even at higher scales (where for iterative causal smoothing filters the scale is intended as the number of iterations). Fig. 3 presents the result of the application of Eq. (8) to the image of Fig. 2. It is a 256 x 256 grayscale image, and the algorithm requires about 0.035 s per iteration. As we can appreciate from the Fig. 3, sharp edges remain stable at higher scales as abrupt changes are preserved by the smoothing process.

Before entering into the quantitative assessment of the algorithm, we want to briefly introduce other well-known edge-preserving recovery algorithms. In particular the half-quadratic regularization method by the ARTUR algorithm and anisotropic diffusion.

4.1. Half-quadratic regularization

The idea of half-quadratic regularization [7,10] consists into transforming the energy $\mathcal{R}(I)$ into its dual form

$$\mathcal{R}^*(I) = \sum_{ij} \phi^+(\vert D_i I_{ij}, (b_{ij})_x \vert) + \phi^-(\vert D_j I_{ij}, (b_{ij})_y \vert)$$

where $b$ are auxiliary variables is such a way that $\mathcal{R}^*(I)$ is quadratic in $I$ when the $b$ are fixed. In [7], the authors present a very efficient iterative algorithm for solving problem (3) which is actually considered the fastest edge-preserving image recovery algorithm. It is based on an alternate minimization

$$I^1 = 0$$

repeat

$$\begin{array}{l}
(b_{ij}^{n+1})_x = \frac{\phi^+(\vert D_i I_{ij} \vert)}{2D_i I_{ij}} \\
(b_{ij}^{n+1})_y = \frac{\phi^-(\vert D_j I_{ij} \vert)}{2D_j I_{ij}}
\end{array}$$

solve the linear system $(I - 2D_i I_{ij}^{n+1}) I_{ij}^{n+1} = \hat{I}$ for $I^{n+1}$

where $I^{n+1} = -D_i \text{diag}([b_{ij}^{n+1}]) D_j - D_j \text{diag}([b_{ij}^{n+1}]) D_i$

$n := n + 1$

until convergence
be efficiently used even for denoising. Despite its efficiency in terms of speed of convergence it requires the choice of several important parameters such as the regularization parameter $\alpha$ and the stopping criteria both for the outer loop and for the adopted iterative method for solving the inner system.

4.2. Anisotropic diffusion

The basic idea of anisotropic diffusion is to evolve from the original image $I_0$ a family of increasingly smooth images $I_t$ derived from the solution of the following partial differential equation [18]:

Fig. 3. The Lenna image after various number of iterations, and the corresponding edges detected with non-maximal suppression. From the first to the last row the number of iterations are respectively 20, 60, 100 and 140.
\[
\frac{\partial I}{\partial t} = \text{div}[c(|\nabla I|)\nabla I]
\]  
(10)

In order for such evolution to be well posed and edge preserving the diffusion coefficient \(c()\) should satisfy several properties \([5,22]\), as being in the form \(c(x) = \frac{\phi(x)}{\sigma}\), where \(\phi\) satisfies the conditions of being a potential function reported in the previous section. In our experiments we use the constrained Total Variation diffusion corresponding to \(c(x) = \frac{1}{2}\) \([19]\), which reads as

\[
\frac{\partial I}{\partial t} = \nabla \cdot \left( \frac{\nabla I}{|\nabla I|} \right) - \mu(I - I_0)
\]

(11)

where \(\mu\) is a Lagrange multiplier chosen in such a way that the constraint \(|I - I_0| \leq \sigma\) is satisfied.

4.3. Gray scale images

In order to evaluate the behaviour of the algorithm, and to compare it with other edge-preserving denoising, we artificially add to the original image some amount of noise and then measure the quality of the reconstruction as function of the iteration. It is well known that the quantitative measures of image reconstruction may often fail with respect to perceptually plausible measures. For example the mean squared error measure tends to compress small errors and to overweight large errors. In this paper we adopt as a quantitative measure of the reconstruction the so-called Mean Error (ME) defined as

\[
\text{ME}(I, I^0) = \frac{1}{MN} \sum_{i,j} |I_{i,j} - I^0_{i,j}|
\]

(12)

The first experiment adopts uniformly distributed additive noise. In particular we add uniform noise at an amount of 8.5 dB, 11 dB and 14.5 dB of Signal to Noise Ratio.

![Fig. 4. The lena image, corrupted by uniform noise at 8.5 dB of SNR (a), and its reconstruction by the proposed method (b), the anisotropic diffusion (c) and half-quadratic regularization (d) (\(\sigma = 0.075\)).](image-url)

![Fig. 5. The ME measure as function of the iteration for the reported algorithms. For the half-quadratic regularization the measure is computed at each iteration of the iterative algorithm adopted to solve the inner linear system, which in our case is a conjugate gradient algorithm with a multigrid preconditioner. This figure refers to uniform noise at 14.5 dB of SNR.](image-url)
The corrupted image and the corresponding reconstructions are reported in Fig. 4. As it can be seen from the images, there is no significant difference between the reconstructions, at least from the perceptual level. In order to quantitatively appreciate the behavior of the algorithm we report the ME as function of the iteration number. As Fig. 5–7 show, the proposed algorithm compares well in terms of quality of reconstruction with the other algorithms reported. For what concerns the parameters adopted for the generation of such figure let us mention that our algorithm does not need any free parameters, whereas for the case of anisotropic diffusion we choose the time step as 0.5, the maximum number of iteration is 80 and the noise variance the same that estimated by (9). Whereas, for the half-quadratic reconstruction we adopted the regularization parameters reported in the figure, and fixed the maximum number of outer iterations to four and the maximum number of inner iteration of the preconditioned conjugate gradient algorithm to seven. The above figures also show that the ARTUR algorithm has a very fast convergence rate however one

![Fig. 6. The ME measure as function of the iteration for the reported algorithms. This figure refers the uniform noise at 11 dB of SNR.](image)

![Fig. 7. The ME measure as function of the iteration for the reported algorithms. This figure refers to uniform noise at 8.5 dB of SNR.](image)

![Fig. 8. Computing times (in percentages) of each iteration of the considered methods.](image)

![Fig. 9. The ME measure as function of the iteration for the reported algorithms. This figure refers to Gaussian noise at 11 dB of SNR.](image)

![Fig. 10. The ME measure as function of the iteration for the reported algorithms. This figure refers to Gaussian noise at 9 dB of SNR.](image)
should consider that the price in terms of computation is much higher with respect to the proposed algorithm. As an example, Fig. 8 reports the computing times of each iteration of the reported algorithm. The histogram is in percentages with respect to the times of the inner iteration of the ARTUR algorithm. Specifically, for the reported 256 × 256 grayscale image the computing times over the adopted platform are 0.027, 0.076 and 0.562 s, respectively for the proposed algorithm, an isotropic diffusion and half-quadratic regularization. If, in addition, we consider that the right regularization parameter must be typically chosen in an experimental trial and error fashion, the

Fig. 11. Reconstruction of an edge-rich noisy image. The original image is presented in figure (a) and its noisy version is in (b). The reconstruction by the anisotropic diffusion and by the proposed method is reported in figures (c) and (d), respectively. On the last row we report the Mean Error measured as function of the iteration for both methods.
advantage of the proposed method is even more evident. Finally, Figs. 11 and 12 report similar results for other well-known test images.

For what concerns Gaussian noise, however, things are a little bit different, Figs. 9 and 10 report the reconstruction result of the same image corrupted by Gaussian noise.

Fig. 12. Reconstruction of another noisy image. The original image is presented in figure (a) and its noisy version is in (b). The reconstruction by the anisotropic diffusion and by the proposed method is reported in figures (c) and (d), respectively. On the last row we report the Mean Error measured as function of the iteration for both methods.
Here, we can see that, the anisotropic diffusion gets a slightly better behavior, this is mainly due to floating point representation of the solution which accounts for small variations in the data.

4.4. Color images

Recently, there has been several works on the restoration of color images for non-flat features such as CB and HSV models [6,20], where the typical channel-by-channel approach is compared to the so-called vectorial approach. Despite its simplicity, the former can get good results with a small computational effort, see for example the results in [6]. Here we show how the simplest extension to RGB color images of the proposed approach works on a real image as reported in Figs. 13 and 14. The main advantage of the proposed method consists into the fact that our algorithm maintains an integer valued representation of the images and uses inter arithmetic, these features can be of particular importance for applications using large color image such as for example digital film restoration where space for image representation and computing times are a matter.

5. Conclusions

We have reported an image recovery algorithm which is based on the Finite Markov Random Field model. We have investigated the properties of edge-preserving potential functions for FMRF and clarified that the linear behavior of potential functions is fundamental for convex edge preserving priors. The resulting algorithm is fast and efficient, does not require any choice of free parameters and has been compared with other well-known denoising methods both in terms of quality of reconstruction and comput-
Both \( f^1 \) and \( f^2 \) are increasing, therefore \( f^1_0 - f^1_{-1} \geq 0 \) and \( f^2_0 - f^2_{-1} \geq 0 \), moreover \( f^1_0 = f^2_0 \) and \( f^1_{-1} = f^2_{-1} \), then let \( c = b - a \), if \( f \) is a FMRF edge preserving then

\[
\phi(a + c) = \phi(b) = \sum_i \phi(f^2_i - f^1_{i-1}) = \sum_i \phi(f^1_i - f^1_{i-1})
\]

\[
= \phi(a) + \phi(b - a) = \phi(a) + \phi(c)
\]

and this is true for any \( a \geq 0 \) and \( c \geq 0 \). \( \square \)

**Proof of Proposition 2.** We can show that \( \mathcal{R}(I^{n+1}) \subseteq \mathcal{R}(I^n) \) in the unidimensional case, since the 2D case is just a tensor product. Let us consider a sign updating of the \( i \)-th point

\[
I^{n+1}_i = I^n_i + A \cdot \text{sign}[(\text{sign}(I^n_{i+1} - I^n_i) - \text{sign}(I^n_i - I^n_{i-1}))]
\]

(13)

The contribution of \( I^{n+1} \) to \( \mathcal{R}(I^{n+1}) \) depends on the value of \( \vert I^n_{i+1} - I^n_i \vert + \vert I^n_i - I^n_{i-1} \vert \), the proof is based of the possible configurations of \( I^n_{i+1}, I^n_i \) and \( I^n_{i-1} \), if \( I^n_{i+1} \leq I^n_i \leq I^n_{i-1} \) or \( I^n_{i-1} \leq I^n_i \leq I^n_{i+1} \), then since by construction the filter leaves unaltered monotone functions, we have \( I^{n+1}_i = I^n_i \) and the proposition is true. Two cases remain, the first when \( I^n_{i-1} \leq I^n_i \leq I^n_{i+1} \) and \( I^n_i \leq I^n_{i-1} \), in this case from (13)

\[
I^{n+1}_i = I^n_i + A
\]

and therefore \( \vert I^{n+1}_{i+1} - I^{n+1}_i \vert = \vert I^n_{i+1} - I^n_i \vert - A \leq \vert I^n_{i+1} - I^n_i \vert \) and \( \vert I^{n+1}_{i-1} - I^{n+1}_i \vert = \vert I^n_{i-1} - I^n_i \vert \leq \vert I^n_{i-1} - I^n_i \vert \), therefore also in this case the proposition is true. The last case is when \( I^n_i \leq I^n_{i+1} \) and \( I^n_i \leq I^n_{i-1} \), the proof in this case is similar to the previous. \( \square \)

**Proof of Proposition 3.** Let \( m = \min \{I^n_{i,j}, I^n_{i-1,j}, I^n_{i+1,j}, I^n_{i,j-1}, I^n_{i,j+1}, I^n_{i-1,j-1}, I^n_{i-1,j+1}, I^n_{i+1,j-1}, I^n_{i+1,j+1}\} \) if \( I^n_{i,j} = m \) then it is easy to verify that

\[
\text{sign}(I^n_{i,j} - I^n_{i-1,j}) - \text{sign}(I^n_{i,j} - I^n_{i,j-1}) + \text{sign}(I^n_{i,j+1} - I^n_{i,j}) - \text{sign}(I^n_{i,j} - I^n_{i,j+1}) \leq 4
\]

and, in this case, from (8)

\[
I^{n+1}_{i,j} = I^n_{i,j} + A \geq m
\]

In the other cases, from the definition of \( A \), we have

\[
I^n_{i,j} \geq m + A
\]

and therefore

\[
I^{n+1}_{i,j} \geq I^n_{i,j} - A \geq m
\]

where the first inequality comes from (8). The proof that \( I^{n+1}_{i,j} \leq M \) is analogous. \( \square \)

**References**


Fig. 14. A real RGB image, its noisy version and the recovered image.


