INVERSE SCALE SPACE METHODS FOR BLIND DECONVOLUTION
ANTONIO MARQUINA *

Abstract. In this paper we propose a blind deconvolution algorithm based on the total variation regularization formulated as a nonlinear inverse scale space method that allows an efficient recovery of edges and textures of blurry and noisy images. The proposed explicit scheme evolve in time the signal and the estimated kernel until a stopping criterion is satisfied. Numerical results indicate that our scheme is robust and converges very fast to the recovered image for images convolved with an experimental point spread function (guided star) and contaminated with Gaussian white noise.

Key words. Total variation restoration, blind deconvolution, Gaussian blur, denoising, inverse scale space methods.

1. Introduction. Given a blurry and noisy image $f : \Omega \rightarrow R$, where $\Omega$ is a bounded open subset in $R^2$, the model of degradation we assume is

$$f = k \ast u + n$$

where $k(x, y)$ is the convolution kernel (Gaussian kernel), $u$ is the original image, $n$ is Gaussian white noise with zero mean and the blurring operator is defined through the convolution

$$(k \ast u)(x, y) = \int_{\Omega} u(s, r)k(x-s, y-r)dsdr.$$

Our goal is to recover the unknown $u$ and/or $k$ from $f$.

Given the knowledge of the kernel $k$, the standard way to solve the problem when noise is not present is to solve the constrained minimization problem:

$$u = \arg \min_u \left\{ F(u) := \frac{\lambda}{2} ||k \ast u - f||_2^2 + \frac{1}{2} \int_{\Omega} |\nabla u|^2 \right\}$$

where $\lambda > 0$ is a scale parameter regularizing the high frequencies of the solution $u$.

This is the Tikhonov regularization of a constrained minimization problem with a quadratic penalization. The Euler-Lagrange (EL) equation associated to (1.1) is the following linear elliptic equation

$$\lambda k \ast (f - k \ast u) + \Delta u = 0$$

with homogeneous Neumann boundary conditions that can be solved by using the fast cosine transform (see [10] for details).

When noise is present the above model procedure is not useful since high frequencies get amplified. The most successful approach to solve this problem is the TV

*The author is with the Departamento de Matematica Aplicada Universidad de Valencia, 46100-Burjassot, Spain (e-mail: marquina@uv.es). This work was supported by DGICYT project MTM2005-07708, Ayudas Estancias 2006 Generalitat Valenciana, DMS-0312222, ACI-0321917, DMI-0327077 and NIH-U54RR021813.
regularization problem proposed by Rudin, Osher and Fatemi ([2]) (we will call the ROF model).

They solve the constrained minimization problem

$$u = \arg\min_u \left\{ F(u) := \frac{\lambda}{2} \|k * u - f\|^2_2 + \|u\|_{TV} \right\},$$

(1.3)

where

$$\|u\|_{TV} = \int |\nabla u|$$

is the total variation norm of $u$. The main advantage of this approach is that sharp edges are much better recovered than the linear model (1.2) does.

If $k$ is the delta function, $\delta * u = u$, (1.3) becomes a pure denoising model which effectively recovers the image with edges from noisy data.

For general convolution kernels $k$, the restoration of the image $u$ is usually a numerically ill-posed problem ([4]). The associated EL equation to (1.3) is

$$\lambda \tilde{k} * (f - k * u) + \nabla \cdot \frac{\nabla u}{|\nabla u|} = 0$$

(1.4)

where $\tilde{k}$ is the adjoint of $k$ with homogeneous Neumann boundary conditions.

Equation (1.4) can be solved by using gradient-descent projection method proposed in [2], (see also [9]).

$$u_t = \lambda \tilde{k} * (f - k * u) + \nabla \cdot \frac{\nabla u}{|\nabla u|}$$

(1.5)

using as initial data the original noisy and blurry signal $f$. The flow defined by (1.5) is the so-called TV-flow.

An iterative refinement applied to the ROF model was introduced in [12] to recover finer scales. This procedure is related to the so-called Bregman distance [11] and reads as follows:

- Set $u_0 = 0; \ p_0 = 0$
- Given $u_{n-1}$ and $p_{n-1} \in \partial J(u_{n-1}), n \geq 1$
  1. Compute $u_n = \arg\min_u Q_n(u)$ with

$$Q_n(u) := J(u) - J(u_{n-1}) - \langle u - u_{n-1}, p_{n-1} \rangle + \frac{\lambda}{2} \|k * u - f\|^2_{L_2}$$

  2. Update the dual variable

$$p_n := p_{n-1} + \lambda \tilde{k} * (f - k * u_{n-1}) \in J(u_n)$$

- $n \rightarrow n + 1$
where $p_n$ is an element of 
\[ \partial J(u) = \{ p : E \to R : \langle v - u, p \rangle \leq J(v) - J(u), \forall v \in E \} \]
the subgradient of $J$. It is possible to show that this algorithm is well-defined and the following is satisfied

\[ p_n = \lambda \sum_{j=1}^{n} \hat{k} * (f - k * u_j) \] (1.6)

The minimizer $u_n$ is unique when $\text{ker} [K] = \{0\}$.

This could not be the case when $k$ is a compact convolution operator so we have some degrees of freedom to choose different sequences [12].

Scale space methods (SSM) are those smoothing small scale features faster than large scale ones. The following is an example of SSM since small scale features smooth faster.

\[ u_t = \nabla \cdot \frac{\nabla u}{|\nabla u|}, \] (1.7)

is an example of SSM since small scale features smooth faster.

Inverse scale space methods (ISSM) have been introduced in [11] and are based on a different idea. Instead of starting with the noisy image and smoothing it when the method evolves in time, ISSM start with the image $u = 0$ and approach the noisy image $f$ as time increases with large scale features converging faster than small ones. An optimal stopping time criterion might give a satisfactory restored image. ISSM methods have been constructed in [11] for quadratic penalization (regularization) functionals. Recently, a new class of nonlinear ISSM has been successfully constructed and tested for the total variation functional in [13] as the limit of the iterated refinement procedure ([12]).

Considering $\lambda$ as a time scale parameter running from the coarser scale $\lambda = 0$ until $\lambda \to \infty$ Burger et al. in [13] state the nonlinear ISSM as the differential equation in $p(t)$

\[ \frac{\partial p}{\partial t} = \hat{k} * (f - k * u(t)), \quad p(t) \in \partial J(u(t)) \] (1.8)

with initial condition $u(0) = p(0) = 0, \quad (p(0) \in \partial J(0))$

The differential equation (1.8) can be integrated in time if we can compute $u(t)$ as function of $p$ in every time step under the condition $p \in \partial J(u)$. This can be done explicitly in 1D but in multidimensions the equation is too involved. To overcome this problem a relaxed ISSM was introduced in [13]. The relaxed ISSM consists in solving the system:

\[ u_t = -p(u) + \lambda \left[ \hat{k} * (f - k * u) + v \right] \] (1.9)

\[ v_t = \alpha \hat{k} * (f - k * u) \]

where $u|_{t=0} = v|_{t=0} = 0$, and $\lambda > 0$ and $\alpha > 0$ are constants.

Fourier analysis of this system applied to the quadratic regularization functional shows that in this case the solution $u$ is the linear combination of two components, one is a Gaussian convolution decaying quickly in time and the other component corresponds to the inverse scale space solution when $\alpha \leq \frac{1}{4}$.
In this paper we are interested in recovering $u$ and the kernel $k$ without having a priori knowledge of $k$ and $u$. This is what we call the blind deconvolution problem.

You and Kaveh ([7]) and later Chan and Wong ([1]) proposed models for blind deconvolution based on a joint minimization problem with respect to $u$ and $k$ as follows:

$$(u,k) = \arg \min_{u,k} \left\{ G(u,k) := \frac{1}{2} \|f - k * u\|_2^2 + \alpha_1 J(u) + \alpha_2 J(k) \right\}$$

with either $J(u) = \frac{1}{2} \int |\nabla u|^2$ in [7] or $J(u) = \int |\nabla u|$ in [1].

This functional is not jointly convex and the solution of these variational problems are not unique. In [14] a new variational problem was proposed to improve the previous results where a discussion on the non-uniqueness of the solution is included.

Here we propose a relaxed inverse scale procedure to obtain a more efficient algorithm to approximate a solution to the blind deconvolution problem.

In section II we have introduced the idea of “progressive deconvolution” to justify the convenience of evolving the estimated kernel to perform deconvolution in a more efficient way. In section III we present our new model and we include preliminary numerical tests in section IV to show the algorithm converges fast to the restored image. We draw our conclusions in section V.

2. A simple model: progressive deconvolution. We consider the ROF deconvolution model (1.3) given the kernel $k$ and a blurry and noisy image $f = k * u + n$, where $n$ is Gaussian white noise.

There are two types of kernels $k$ we can deal with: reducible kernels and irreducible kernels.

1. Reducible kernels There is $N > 0$ such that

$$k = k_1 * k_2 * k_3 * \cdots * k_N$$

and the support of $k_j$ is much smaller than the one of $k$.

Example: Gaussian kernels. These kernels are infinitely reducible since

$$k_3(x,y) = \frac{1}{4\pi\delta} e^{-(x^2 + y^2)/4\delta}$$

$$k = k_3 * k_3 * \cdots * k_3$$

for arbitrary $N > 0$, and the support of $k_N$ is small for $N$ large.

2. Irreducible kernels There is no $k_1 \neq k$ such that $k_1 * k_2 = k$.

Example: Motion blur. In one dimension, $k = \chi_{[-a,a]}$ will $\int k = 1$.

Deconvolution for irreducible kernels using the ROF-model appears to be better conditioned than the one for reducible ones, (like Gaussian blur with large support plus noise).

Then, the idea is to design a procedure for reducible kernels, we call ”Progressive Deconvolution”, that is more accurate and better conditioned than the direct application of the ROF-model using the original kernel.
We consider a reducible kernel $k$, with $N$ factors ($N \gg 1$):

$$k = k_N * k_{N-1} * \cdots * k_2 * k_1$$

Then, if we fix a convenient Lagrange multiplier $\lambda$, (depending on $N$), we solve the following sequence of deconvolution problems:

$$u_1 = \arg \min_u \left\{ \int |\nabla u| + \frac{\lambda}{2} ||k_1 * u - f||^2 \right\},$$

$$u_2 = \arg \min_u \left\{ \int |\nabla u| + \frac{\lambda}{2} ||k_2 * u - u_1||^2 \right\},$$

$$\vdots$$

$$u_N = \arg \min_u \left\{ \int |\nabla u| + \frac{\lambda}{2} ||k_N * u - u_{N-1}||^2 \right\}$$

Then we have the following sequence of decompositions, ([5],[6]):

$$f = k_1 * u_1 + v_1$$

$$u_1 = k_2 * u_2 + v_2$$

$$u_2 = k_3 * u_3 + v_3$$

(2.1)

$$\vdots$$

$$u_{N-1} = k_N * u_N + v_N$$

then

$$f = k_1 * k_2 * k_3 * \cdots * k_N * u_N + v = k * u + v$$

where

$$v = k_1 * \cdots * k_{N-1} * v_N + k_1 * \cdots * k_{N-2} * v_{N-1} + \cdots + k_1 * v_2 + v_1$$

and $u = u_N$.

The advantage of this approach is that the algorithm is accurate and efficient for each of the problems.

The "blind deconvolution problem" can be addressed from this point of view:

We start with a kernel of small support: $k_0$. We may choose either:

- A regularization of the delta function consistent with the resolution of the original signal
- A "model kernel" (also with small support) that fits with experimental data

Then, we apply the progressive deconvolution (as in (2.1)) taking

$$k_1 = k_0; \; k_2 = k_0; \; \cdots; \; k_N = k_0$$
until "certain N" for which we have recovered the scales we want.

The blind deconvolution of Gaussian blur starting with a convenient regularization of the delta function appears to be effective.

The algorithm based on the ROF-model consists in solving a sequence of Euler-Lagrange equations:

\[-\nabla \cdot \frac{\nabla u}{|\nabla u|} + \lambda \hat{k}_1 * (k_1 * u - f) = 0\]

whose solution \(u_1\) is the input data for

\[-\nabla \cdot \frac{\nabla u}{|\nabla u|} + \lambda \hat{k}_2 * (k_2 * u - u_1) = 0\]

and so on. The last Euler-Lagrange equation to be solved is

\[-\nabla \cdot \frac{\nabla u}{|\nabla u|} + \lambda \hat{k}_N * (k_N * u - u_{N-1}) = 0\]

whose solution \(u = u_N\) is the solution to the deconvolution problem with kernel

\[k = k_1 \ast k_2 \ast k_3 \cdots \ast k_N\]

The Lagrange multiplier \(\lambda > 0\) used for all partial deconvolutions performed is constant and it is chosen according to the SNR ratio of the original signal \(f\).

To analyze the behavior of the progressive deconvolution in terms of the Fourier transform we formulate the basic variational problem using the quadratic penalization as regularizing functional and we obtain the following sequence of Euler-Lagrange equations:

\[-\Delta u_1 + \lambda \hat{k}_1 * (k_1 * u_1 - f) = 0\]
\[-\Delta u_2 + \lambda \hat{k}_2 * (k_2 * u_2 - u_1) = 0\]
\[
\vdots
\]
\[-\Delta u_N + \lambda \hat{k}_N * (k_N * u_N - u_{N-1}) = 0\]

then, using the Fourier transform we have the following set of algebraic equations for every frequency \((k, l)\):

\[\hat{u}_j = \frac{\lambda \hat{u}_{j-1} \hat{k}_j}{1/\epsilon^2 + \lambda k_j^2} + \lambda k_j^2\] (2.2)

then

\[\hat{u}_N = \frac{\lambda^N \hat{k}_1 \cdots \hat{k}_N \hat{f}}{\prod_{j=1}^{N} \left[1/\epsilon^2 + \lambda k_j^2\right]} = \frac{\lambda^N \hat{k} \hat{f}}{\prod_{j=1}^{N} \left[1/\epsilon^2 + \lambda k_j^2\right]}\] (2.3)
If we solve the whole problem with appropriate $\lambda_0 > 0$

\[-\Delta u + \lambda_0 \hat{k} * (k * u - f) = 0\]

using the Fourier transform we have

\[\hat{u}_j = \frac{\lambda_0 \hat{k} \hat{f}}{\frac{1}{k^2 + \tau^2} + \lambda_0 k^2}\]  \hspace{1cm} (2.4)

If we match the frequencies of $u_N$ and $\hat{u}$, we have that

\[\hat{u} \approx \hat{u}_N\]

and

\[\prod_{j=1}^{N} \frac{\lambda^N}{\frac{1}{k_j^2 + \tau^2} + \lambda^2 k_j^2} \approx \frac{\lambda_0}{\frac{1}{k^2 + \tau^2} + \lambda_0 k^2}\]  \hspace{1cm} (2.5)

The continuous version of the progressive deconvolution of Gaussian blur by means of a nonlinear inverse scale space method can be written as follows

\[
\begin{align*}
\nu_t &= \lambda[(\hat{k} * (f - k * u)) + v] + \nabla \cdot \frac{\nabla u}{|\nabla u|} \\
\varepsilon_t &= \alpha \hat{k} * (f - k * u) \\
k_t &= \frac{1}{\mu} \Delta k
\end{align*}
\]

where $\lambda > 0$, $\mu > 0$, $\alpha = \lambda/4$ with initial data $u|_{t=0} = 0$, $v|_{t=0} = 0$ and $k|_{t=0} = \delta$.

This is a very simple model to "de-Gaussify" without a priori knowledge of the kernel.

3. The new model. One step forward is to reformulate the constrained blind deconvolution model proposed in ([14]), as an nonlinear inverse scale space that evolves the signal and the kernel starting from the zero signal and the $\delta$ kernel.

We consider the following general variational problem with respect to $k$ and $u$

\[
(u, k) = \arg \min_{u,k} \{ F(u, k) := H(u, k, f) + J_1(u) + J_2(k) \} .
\]

Then we consider the following relaxed inverse scale space model for blind deconvolution

\[
\begin{align*}
\nu_t &= -p(u) + \lambda [\partial_u H(u, k, f) + v] \\
\varepsilon_t &= -\alpha \partial_u H(u, k, f), \hspace{1cm} p \in \partial J_1(u) \\
k_t &= -q(k) + \mu [\partial_k H(u, k, f) + w] \\
w_t &= -\beta \partial_k H(u, k, f), \hspace{1cm} q \in \partial J_2(u)
\end{align*}
\]
where $u|_{t=0} = v|_{t=0} = 0$, and $k|_{t=0} = w|_{t=0} = \delta$

In our case we will consider $J_1(u) = \int |\nabla u|$, $J_2(k) = \int |\nabla k|$ or $\frac{1}{2} \int |\nabla k|^2$ and

$$H(u, k, f) = \frac{1}{2} ||f - k * u||^2_2 + \frac{\alpha_1}{2} \left( \int u - w \right)^2 + \frac{\alpha_2}{2} \left( \int |u| - w \right)^2 + \frac{\beta_1}{2} \left( \int k - 1 \right)^2 + \frac{\beta_2}{2} \left( \int |k| - 1 \right)^2,$$

The constraints $\int k = 1, \int u = w = \int f$ and $u, k \geq 0$ are approximately satisfied for $\alpha_1, \alpha_2, \beta_1, \beta_2$ positive and large, (see [14]).

Then the model will be
\[
\begin{align*}
    u_t &= -p(u) + \lambda \left[ \hat{k} \ast (f - k \ast u) + v + \alpha_1(w - \int u) \right] + \alpha_2(w - \int |u| \operatorname{sgn}(u)) \\
    v_t &= \alpha \left[ \hat{k} \ast (f - k \ast u) + \alpha_1(w - \int u) \right] + \alpha_2(w - \int |u| \operatorname{sgn}(u)) \\
    k_t &= -q(k) + \mu \left[ \hat{u} \ast (f - u \ast k) + w + \beta_1(1 - \int k) \right] + \beta_2(w - \int |k| \operatorname{sgn}(k)) \\
    w_t &= \beta \left[ \hat{u} \ast (f - u \ast k) + w + \beta_1(1 - \int k) \right] + \beta_2(w - \int |k| \operatorname{sgn}(k))
\end{align*}
\] (3.5) (3.6) (3.7) (3.8)

where \( p \in \partial_u J_1(u) \), \( q \in \partial_k J_2(k) \) and \( \operatorname{sgn} \) is the sign function.

For the linear case \((J_1(u) = \frac{1}{2} \int |\nabla u|^2 \) and \( J_2(k) = \frac{1}{2} \int |\nabla k|^2 \) and disregarding the \( \alpha_1, \alpha_2, \beta_1, \beta_2 \) terms, we can perform similar Fourier analysis as the one in [13] to conclude that either ”freezing” \( u \) or \( k \) the equations converge to steady state under the conditions

\[
\begin{align*}
    \alpha &\leq \frac{\lambda}{4} |\hat{k}|^2 \\
    \beta &\leq \frac{\mu}{4} |\hat{u}|^2
\end{align*}
\]

where \( \hat{k} \) and \( \hat{u} \) are respectively Fourier transforms of \( k \) and \( u \).

If \( u_s \) and \( k_s \) are the steady state solution then \( \hat{k}_s \ast f = \hat{k}_s \ast k_s \ast u_s \), \( v_s = \frac{p(u_s)}{\lambda} \) and \( w_s = \frac{q(k_s)}{\mu} \).

The inverse scale space method described by the equations (3.5), (3.6), (3.7) and (3.8) for a given blurry and noisy signal \( f \) approaches a steady state solution \( k_s, u_s \) when time increases that satisfy \( \hat{k}_s \ast f = \hat{k}_s \ast k_s \ast u_s \) which might not be desirable since \( u_s \) is expected to be noisy and the constraint

\[
||f - k(t) \ast u(t)||_{L^2}^2 = \sigma^2
\]

is satisfied exactly. The stopping time criterion we will use in our numerical tests consists in choosing the solution at the minimum time for which the \( L^2 \)-norm of the residual is the standard deviation of the noise.

**4. Numerical results.** In this section we present preliminary numerical results we have obtained using the blind deconvolution model based on ISSM. We will use \( J_1(u) = \int |\nabla u| \) and \( J_2(k) = \frac{1}{2} \int |\nabla k|^2 \).

If \( f \) is the noisy/blurred image and \( k \) is the convolution kernel (estimated or given) then the difference \( n = f - k \ast u \) is Gaussian white noise of zero mean, \( \sigma^2 \approx ||f - k \ast u||_{L^2}^2 \) is the estimated variance of the noise. If \( f \) is the mean of \( f \), we compute the signal
to noise ratio by means of

$$SNR := \frac{||f - \tilde{f}||_{L^2}}{||n - \tilde{n}||_{L^2}}$$

for Gaussian white noise $n$, with mean $\tilde{n}$.

All numerical solutions of the blind deconvolution model were obtained by discretizing the evolution equations to first order in time using central differences for the second order terms and fast cosine transform for the computation of the convolutions to ensure homogeneous Neumann boundary conditions, ([3], [10]).

We use the regularization of $J_1(u) = \int |\nabla u|$ by using the perturbation $J_1(u) := \int \sqrt{||\nabla u||^2 + \epsilon^2}$ where $\epsilon$ is a small positive number. In our calculations we use $\epsilon \approx 10^{-4}$.

We will consider two-dimensional images of $256 \times 256$ resolution points and kernels of $129 \times 129$ resolution points in all numerical experiments presented here.
4.1. Example 1: Pure deblurring problem. We consider the blurred image $f$ in Figure 2.1 (top right) obtained by convolving the original image in Figure 2.1 (top left), with the experimental point spread function of Figure 2.1 bottom left. The point spread function used here is slight asymmetric and different in shape from a Gaussian kernel.

We run our model for blind deconvolution represented by (3.1), (3.2), (3.3) and (3.4), and we obtain the results for times $t = 25, 50, 75$ and $100$ using $\lambda = 1$, $\alpha = 0.25$, $\mu = 2$, $\beta = 0.25$ and $\Delta t = 0.1$ displayed in Figure 3.1. We used the original kernel to deconvolve $f$ by means of the ROF model and the ISSM based on the ROF energy and we get comparable results but we recover more textures with our blind deconvolution model. This means that our model is better conditioned that the one using the original kernel since the large support of the kernel makes the evolution of the first two equations less efficient and accurate.

Fig. 4.1. Deblurring and denoising problem (Blind): $t = 25$ (top left); $t = 50$ (top right); $t = 75$ (bottom left); $t = 100$ (bottom right).
4.2. Example 2: Deblurring and denoising problem. We consider a blurry and noisy image obtained (Figure 2.1 bottom right) obtained by adding Gaussian white noise with standard deviation $\sigma = 2.78$ to the blurry image appearing in Figure 2.1 top right. We run our model using $\lambda = 1$, $\alpha = 0.25$, $\mu = 2$, $\beta = 0.5$ and $\Delta t = 0.1$ for $t = 25, 50, 75$ and $t = 100$. We display the numerical results in Figure 4.1. We observe that the signal is recovered very efficiently in spite the support of the kernel is large. We display in Figure 4.2 the $L^2$-norm of the residual $f - k * u$, versus time for both experiments and we observe a good numerical convergence until $\|f - k * u\|_{L^2} = \sigma$. At later times, textures and noise are back.

5. Conclusion. We present a model for blind deconvolution based on total variation regularization formulated as a nonlinear inverse scale space method that evolves the signal and the kernel from the zero signal and a delta function as initial kernel. The nonlinear inverse scale space formulation appears to be more appropriate for deconvolution of signals since large scale features converge faster than small ones. We implement a simple time dependent explicit algorithm that is robust and converges fast to the restored image by using a stopping criterion based on the $L^2$-norm of the residual. Our preliminary numerical tests with blurred and noisy images with Gaussian blur and Gaussian white noise indicates the algorithm works efficiently and accurate.

REFERENCES