

# Semi-Blind Image Restoration via Mumford-Shah Regularization

L. Bar      N. Sochen\*      N. Kiryati

School of Electrical Engineering

\* Dept. of Applied Mathematics

Tel Aviv University

Tel Aviv, 69978, Israel

## Abstract

Image restoration and segmentation are both classical problems, that are known to be difficult and have attracted major research efforts. This paper shows that the two problems are tightly coupled and can be successfully solved together. Mutual support of image restoration and segmentation processes within a joint variational framework is theoretically motivated, and validated by successful experimental results. The proposed variational method integrates semi-blind image deconvolution (parametric blur-kernel), and Mumford-Shah segmentation. The functional is formulated using the  $\Gamma$ -convergence approximation and is iteratively optimized via the alternate minimization method. While the major novelty of this work is in the unified treatment of the semi-blind restoration and segmentation problems, the important special case of known blur is also considered and promising results are obtained.

## 1 Introduction

Image analysis systems usually operate on blurred and noisy images. The standard model  $g = h * f + n$  is applicable to a large variety of image degradation processes that are encountered in practice. Here  $h$  represents an (often unknown) space-invariant blur kernel (point spread function),  $n$  is the noise and  $f$  is an ideal version of the observed image  $g$ .

The two following problems are at the basis of successful image analysis. (1) Can we estimate the blur kernel  $h$  and recover  $f$ ? (2) Can we segment  $g$  in agreement with the structure of  $f$ ? Blind restoration and image segmentation are both classical problems, that

are known to be difficult and have attracted major research efforts, see e.g. [3, 12, 13, 20].

Had the correct segmentation of the image been known, blind image restoration would have been facilitated. Clearly, the blur kernel could have then been estimated based on the smoothed profiles of the known edges. Furthermore, denoising could have been applied to the segments without over-smoothing the edges. Conversely, had adequate blind image restoration been accomplished, successful segmentation would have been much easier to achieve. Blind restoration and image segmentation are therefore tightly coupled tasks: the solution of either problem would become fairly straightforward given that of the other.

This paper presents an integrated framework for simultaneous semi-blind restoration and image segmentation. As will be seen, strong arguments exist in favor of constraining the recovered blur kernel to parameterized function classes, see e.g. [6]. Our approach is presented in the context of the fundamentally important model of isotropic Gaussian blur, parameterized by its (unknown) width. Preliminary results were presented in [4].

## 2 Fundamentals

### 2.1 Segmentation

The difficulty of image segmentation is well known. Successful segmentation requires top-down flow of models, concepts and a priori knowledge in addition to the image data itself. In their segmentation method, Mumford and Shah [14] introduced top-down information via the preference for piecewise-smooth segments separated by well-behaved contours. Formally, they proposed to minimize a functional that includes a fidelity term, a piecewise-smoothness

term, and an edge integration term:

$$\mathcal{F}(f, K) = \frac{1}{2} \int_{\Omega} (f - g)^2 dA + \beta \int_{\Omega \setminus K} |\nabla f|^2 dA + \alpha \int_K d\sigma \quad (1)$$

Here  $K$  denotes the edge set and  $\int_K d\sigma$  is the total edge length. The coefficients  $\alpha$  and  $\beta$  are positive regularization constants. The primary difficulty in the minimization process is the presence of the unknown discontinuity set  $K$  in the integration domains.

The  $\Gamma$ -convergence framework approximates an irregular functional  $\mathcal{F}(f, K)$  by a sequence  $\mathcal{F}_{\epsilon}(f)$  of regular functionals such that

$$\lim_{\epsilon \rightarrow 0} \mathcal{F}_{\epsilon}(f) = \mathcal{F}(f, K)$$

and the minimizers of  $\mathcal{F}_{\epsilon}$  approximate the minimizer of  $\mathcal{F}$ . Ambrosio and Tortorelli [1] applied this approximation to the Mumford-Shah functional, and represented the edge set by a characteristic function  $(1 - \chi_K)$  which is approximated by an auxiliary function  $v$ , i.e.,  $v(x) \approx 0$  if  $x \in K$  and  $v(x) \approx 1$  otherwise. The functional thus takes the form

$$\mathcal{F}_{\epsilon}(f, v) = \frac{1}{2} \int_{\Omega} (f - g)^2 dA + \beta \int_{\Omega} v^2 |\nabla f|^2 dA + \alpha \int_{\Omega} \left( \epsilon |\nabla v|^2 + \frac{1}{4\epsilon} (v - 1)^2 \right) dA. \quad (2)$$

See also Richardson and Mitter [16]. Discretization of the Mumford-Shah functional and its  $\Gamma$ -convergence approximation is considered in [7]. Additional perspectives on variational segmentation can be found in Vese and Chan [23] and in Samson *et al* [19].

Simultaneous segmentation and restoration of a blurred and noisy image has recently been presented in [11]. A variant of the Mumford-Shah functional was approached from a curve evolution perspective. In that work, the discontinuity set is limited to an isolated

closed curve in the image and the blurring kernel  $h$  is assumed to be a priori known.

## 2.2 Restoration

Restoration of a blurred and noisy image is difficult even if the blur kernel  $h$  is known.

Formally, finding  $f$  that minimizes

$$\|h * f - g\|_{L_2(\Omega)}^2 \quad (3)$$

is an ill-posed inverse problem: small perturbations in the data may produce unbounded variations in the solution. In Tikhonov regularization [22], a smoothing term  $\int_{\Omega} |\nabla f|^2 dA$  is added to the fidelity functional (3). In image restoration, Tikhonov regularization leads to over-smoothing and loss of important edge information. For better edge preservation, the Total Variation approach [17, 18] replaces  $L_2$  smoothing by  $L_1$  smoothing. The functional to be minimized is thus

$$\mathcal{F}(f, h) = \frac{1}{2} \|h * f - g\|_{L_2(\Omega)}^2 + \beta \int_{\Omega} |\nabla f| dA . \quad (4)$$

This nonlinear optimization problem can be approached via the half-quadratic minimization technique [2, 9, 10]. An efficient alternative approach, based on the lagged diffusivity fixed point scheme and conjugate gradients iterations, was suggested by Vogel and Oman [24].

Image restoration becomes even more difficult if the blur kernel  $h$  is not known in advance. In addition to being ill-posed with respect to the image, the blind restoration problem is ill-posed in the kernel as well. To illustrate one aspect of this additional ambiguity, suppose

that  $h$  represents isotropic Gaussian blur, with variance  $\sigma^2 = 2t$ :

$$h_t = \frac{1}{4\pi t} e^{-\frac{x^2+y^2}{4t}} .$$

The convolution of two Gaussian kernels is a Gaussian kernel, the variance of which is the sum of the two originating variances:

$$h_{t_1} * h_{t_2} = h_{t_1+t_2} . \tag{5}$$

Assume that the true  $t$  of the blur kernel is  $t = T$ , so  $g = h_T * f$  . The fidelity term (3) is obviously minimized by  $f$  and  $h_T$ . However, according to equation (5),  $g$  can also be expressed as

$$g = h_{t_1} * h_{t_2} * f \quad \forall (t_1 + t_2) = T .$$

Therefore, an alternative hypothesis, that the original image was  $h_{t_2} * f$  and the blur kernel was  $h_{t_1}$ , minimizes the fidelity term just as well. This exemplifies a fundamental ambiguity in the division of the apparent blur between the recovered image and the blur kernel, i.e., that the scene itself might be blurred. For meaningful image restoration, this hypothesis must be rejected and the largest possible blur should be associated with the blur kernel. It can be achieved by adding a kernel-smoothness term to the functional.

Blind image restoration with joint recovery of the image and the kernel, and regularization of both, was presented by You and Kaveh [26], followed by Chan and Wong [8]. Chan and Wong suggested to minimize a functional consisting of a fidelity term and total variation ( $L_1$

norm) regularization for both the image and the kernel:

$$\mathcal{F}(f, h) = \frac{1}{2} \|h * f - g\|_{L_2(\Omega)}^2 + \alpha_1 \int_{\Omega} |\nabla f| dA + \alpha_2 \int_{\Omega} |\nabla h| dA . \quad (6)$$

In their iterative algorithm, the recovered image  $f$  is initialized as the observed image  $g$ , and the blur kernel  $h$  as the impulse function. The parameters  $\alpha_1$  and  $\alpha_2$  control the  $L_1$ -smoothness of the recovered image and the kernel respectively. As  $\alpha_1$  is increased, the recovered image becomes smoother, at the expense of lower fidelity and a narrow kernel. Increasing the parameter  $\alpha_2$  would lead to a wider kernel, at the expense of lower fidelity and ringing in the recovered image. Much can be learned about the blind image restoration problem by studying the characteristics and performance of this algorithm. Consider the images shown in Fig. 1. An original image (upper left) is degraded by isotropic Gaussian blur with  $\sigma = 2.1$  (top-right). Applying the algorithm of [8] (with  $\alpha_1 = 10^{-4}$  and  $\alpha_2 = 10^{-4}$ ) yields a recovered image (bottom-left) and an estimated kernel (bottom-right). It can be seen that the identification of the kernel is inadequate, and that the image restoration is sensitive to the kernel recovery error.

To obtain deeper understanding of these phenomena, we plugged the original image  $f$  and the degraded image  $g$  into the functional (6), and carried out minimization only with respect to  $h$ . The outcome was similar to the kernel shown in Fig. 1 (bottom-right). This demonstrates an excessive dependence of the recovered kernel on the image characteristics. At the source of this problem is the aspiration for general kernel recovery: the algorithm of [8] imposes only mild constraints on the shape of the reconstructed kernel. This allows the distribution of edge directions in the image to have an influence on the shape of the recovered kernel, via the trade-off between the fidelity and kernel smoothness terms. In the

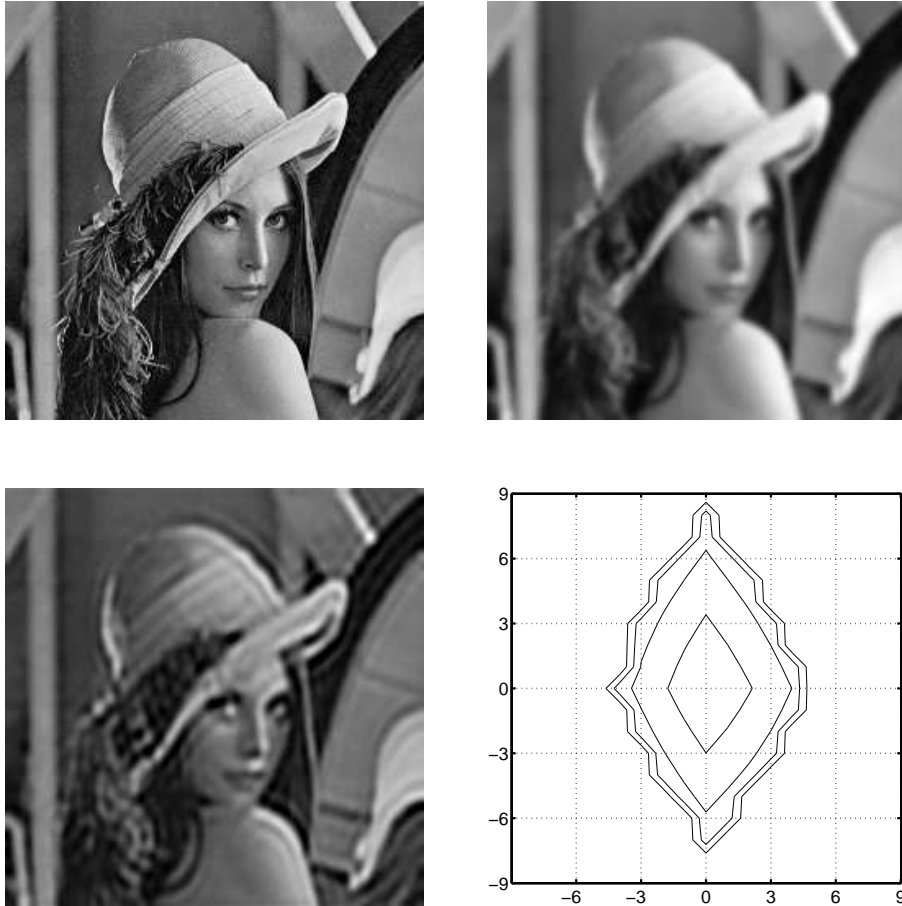


Figure 1: Blind image restoration using the method of [8]. *Top-left:* Original. *Top-right:* Blurred using an isotropic Gaussian kernel ( $\sigma = 2.1$ ). *Bottom-left:* Recovered image. *Bottom-right:* Reconstructed kernel.

example of Fig. 1, there are more vertical edges than horizontal ones. Since the functional (6) trades between the smoothness of the estimated kernel and the smoothness of the recovered image, it has an incentive to overestimate the vertical scale of the kernel. That is because the resulting penalty in image smoothness (ringing across the few horizontal edges) is relatively small. For additional insight see Fig. 2.

Facing the ill-posedness of blind restoration with a general kernel, two approaches can be taken. One is to add relevant data; the other is to constrain the solution. Recent studies have adopted one of these two approaches, or both. In [21], the blind restoration problem is

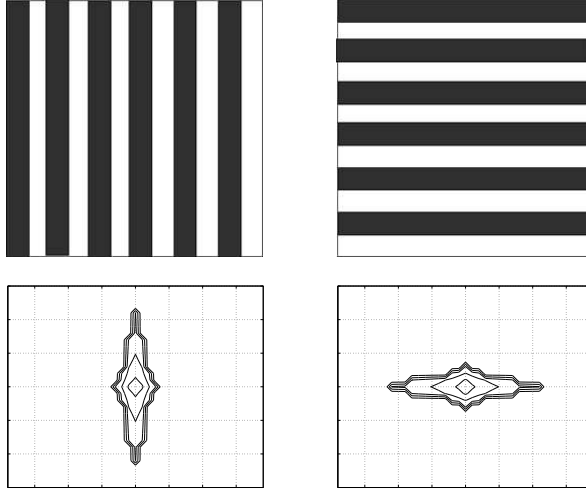


Figure 2: Experimental demonstration of the dependence of the recovered kernel on the image characteristics in [8]. Each of the two synthetic bar images (top-row) was smoothed using an isotropic Gaussian kernel, and forwarded to the blind restoration algorithm of [8] (equation 6). The respective recovered kernels are shown in the bottom row.

considered within a multichannel framework, where *several* input images can be available.

In many practical situations, the blurring kernel can be modeled by the physics/optics of the imaging device and the set-up. The blurring kernel can then be constrained and described as a member in a class of parametric functions. This constraint was exploited in the direct blind deconvolution algorithm of [6]. In [15], additional relevant data was introduced via learning of similar images and the blur kernel was assumed to be Gaussian.

By modifying the framework of [8], we can demonstrate the gains that can be obtained by reducing the blind image restoration problem to semi-blind restoration, in which the kernel is assumed to belong to a class of parametric functions. Fig. 3 is obtained by forcing the algorithm to restrict the estimated kernel to isotropic Gaussians of unknown width. The parameters in this example were  $\alpha_1 = 10^{-4}$  and  $\alpha_2 = 2$ , and the estimated  $\sigma$  was 1.94. The significant improvement with respect to Fig. 1 (bottom left) is clear.





Figure 3: Semi-blind restoration obtained by restricting the algorithm of [8] to isotropic Gaussian kernels of unknown width. Compare with Fig. 1 (bottom-left).

### 3 Coupling semi-blind restoration with segmentation

The observation, discussed in the introduction, that blind restoration and image segmentation can be mutually supporting, is the fundamental motivation for this work. We present an algorithm, based on functional minimization, that iteratively alternates between segmentation, parametric blur identification and restoration. We note that large classes of practical imaging problems are compatible with constrained blur models. In particular, Carasso [6] described the association of the Gaussian case with diverse applications such as undersea imaging, nuclear medicine, computed tomography scanners, and ultrasonic imaging in non-destructive testing.

In this work we integrate Mumford-Shah segmentation with semi-blind deconvolution of isotropic Gaussian blur. This is accomplished by extending the Mumford-Shah functional and applying the  $\Gamma$ -convergence approximation as described in [1]. The observed image  $g$  is modeled as  $g = h_\sigma * f + n$  where  $h_\sigma$  is an isotropic Gaussian kernel parameterized by its

width  $\sigma$ , and  $n$  is white Gaussian noise. The objective functional used is

$$\mathcal{F}_\epsilon(f, v, \sigma) = \frac{1}{2} \int_{\Omega} (h_\sigma * f - g)^2 dA + G_\epsilon(f, v) + \gamma \int_{\Omega} |\nabla h_\sigma|^2 dA \quad (7)$$

where  $G_\epsilon(f, v)$  is the image regularization term, defined as

$$G_\epsilon(f, v) = \beta \int_{\Omega} v^2 |\nabla f|^2 dA + \alpha \int_{\Omega} \left( \epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right) dA.$$

In the sequel it is assumed that the image domain  $\Omega$  is a rectangle in  $\mathbb{R}^2$  and that image intensities are normalized to the range  $[0, 1]$ .

The functional depends on the functions  $f$  (ideal image) and  $v$  (edge integration map), and on the width parameter  $\sigma$  of the blur kernel  $h_\sigma$ . The fidelity term accounts for the blur process  $h_\sigma * f$ . Braides [5] presents a proof for the  $\Gamma$ -convergence of  $G_\epsilon$  to the regularization terms in the Mumford-Shah functional (1).

The last term in Eq. 7 stands for the regularization of the kernel, necessary to resolve the fundamental ambiguity in the division of the apparent blur between the recovered image and the blur kernel. This means that we prefer to reject the hypothesis that the blur originates from  $f$ , and assume that it is due to the convolution with the blur kernel. From the range of possible kernels, we thus select a wide one. This preference is represented by the kernel smoothness term: the width of the Gaussian corresponds to its smoothness, measured by the  $L_2$  norm of its gradient. The convenient  $L_2$  norm was preferred because the blur kernels are restricted to Gaussians, that are smooth functions. Note that the kernel regularization term in [8] had to use the  $L_1$  norm, because they admitted piecewise smooth kernels.

Minimization of the objective functional (7) with respect to  $f$  and  $v$  is carried out us-

ing the Euler-Lagrange (E-L) equations (8) and (9), with Neumann boundary conditions. The E-L equations are *linear* partial differential equations; their derivation is presented in Appendix A. Minimization with respect to the scalar parameter  $\sigma$  is determined by differentiation of the objective functional (Eq. 10).

$$\frac{\delta \mathcal{F}_\epsilon}{\delta v} = 2\beta v |\nabla f|^2 + \alpha \cdot \frac{v-1}{2\epsilon} - 2\epsilon \alpha \nabla^2 v = 0 \quad (8)$$

$$\frac{\delta \mathcal{F}_\epsilon}{\delta f} = (h_\sigma * f - g) * h_\sigma(-x, -y) - 2\beta \text{Div}(v^2 \nabla f) = 0 \quad (9)$$

$$\frac{\partial \mathcal{F}_\epsilon}{\partial \sigma} = \int_\Omega \left[ (h_\sigma * f - g) \left( \frac{\partial h_\sigma}{\partial \sigma} * f \right) + \gamma \frac{\partial}{\partial \sigma} |\nabla h_\sigma|^2 \right] dA = 0 \quad (10)$$

where in Eq. 10,

$$\frac{\partial h_\sigma}{\partial \sigma} = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \cdot \left( \frac{x^2+y^2}{\sigma^3} - \frac{2}{\sigma} \right), \quad (11)$$

and

$$\frac{\partial}{\partial \sigma} |\nabla h_\sigma|^2 = \frac{1}{2\pi^2\sigma^4} e^{-\frac{x^2+y^2}{\sigma^2}} \cdot \left( \frac{x^2+y^2}{\sigma^7} - \frac{4}{\sigma^5} \right). \quad (12)$$

Studying the objective functional (7), it can be seen that it is strictly convex and lower bounded with respect to the functions  $f$  and  $v$  if the other one and the blur kernel  $h_\sigma$  are fixed. The convexity property also holds for general blur kernels  $h(x, y)$ . However, because of the non-linear dependence of  $h_\sigma$  on  $\sigma$ , the functional is not convex with respect to  $\sigma$ . Nevertheless, no convergence problems have been encountered in our experiments. Note that since the Gaussian is an even function, the search for  $\sigma$  can be limited to positive values. Following [8], the alternate minimization (AM) approach is applied: in each step of the iterative procedure we minimize with respect to one function and keep the other two fixed. This leads to a local minimum. The following algorithm is obtained:

Initialization:  $f = g, \quad \sigma = \varepsilon_1, \quad v = 1, \quad \sigma_{prev} \gg 1$

while  $(|\sigma_{prev} - \sigma| > \varepsilon_2)$  repeat

1. Solve equation (8) for  $v$
2. Solve equation (9) for  $f$
3. Set  $\sigma_{prev} = \sigma$ , and solve equation (10) for  $\sigma$ .

Here  $\varepsilon_1$  and  $\varepsilon_2$  are small positive constants.

## 4 Discrete solution

The essence of the proposed algorithm is iterative solution of the linear partial differential equations (8,9) and equation (10). These equations are discretized and solved numerically.

The discretization scheme used was CCFD (cell-centered finite difference) [21, 24]. Consider a function  $u(x, y)$  defined above the unit square. In CCFD, the domain is divided into  $M \times M$  identical  $h \times h$  cells ( $h = 1/M$ ), and  $u(x, y)$  is approximated by a function  $U(x, y)$  that is constant within each cell. Formally, the cell centers are  $(x_i, y_j)$ , where

$$x_i = (i - 1/2)h, \quad i = 1, \dots, M \quad (13)$$

$$y_j = (j - 1/2)h, \quad j = 1, \dots, M \quad (14)$$

and  $U(x_i, y_j) = u(x_i, y_j)$ . The forward and backward finite difference approximations of the derivatives  $\partial u(x, y)/\partial x$  and  $\partial u(x, y)/\partial y$  are respectively denoted by  $\Delta_{\pm}^x u_{ij} = \pm(u_{i\pm 1, j} - u_{ij})$  and  $\Delta_{\pm}^y u_{ij} = \pm(u_{i, j\pm 1} - u_{ij})$ .

Step 1 of the algorithm calls for the solution of equation (8). Following discretization, we obtain a linear system of equations

$$2\beta v_{ij} [(\Delta_+^x f_{ij})^2 + (\Delta_+^y f_{ij})^2] + \alpha \cdot \frac{v_{ij} - 1}{2\epsilon} - 2\alpha\epsilon (\Delta_-^x \Delta_+^x v_{ij} + \Delta_-^y \Delta_+^y v_{ij}) = 0. \quad (15)$$

With column-stack ordering of  $\{v_{ij}\}$ , this system is of the form  $Pv = q$ , where the matrix  $P$  is symmetric and sparse. It is solved using the Minimal Residual algorithm [25].

Let  $H_\sigma$  denote the operator of convolution with a Gaussian  $h_\sigma$ , i.e.,  $H_\sigma f = h_\sigma(x, y) * f$ .  $H_\sigma^*$  is the adjoint operator,  $H_\sigma^* f = h_\sigma(-x, -y) * f$ . Using the notation of [24], let  $L(v)$  denote the differential operator

$$L(v)f = -\text{Div}(v^2 \nabla f). \quad (16)$$

Now, equation (9) can be expressed as

$$H_\sigma^*(H_\sigma f - g) + 2\beta L(v)f = 0 \quad (17)$$

Let  $A(v)f = H_\sigma H_\sigma^* f + 2\beta L(v)f$ . Rearranging equation (17) we obtain

$$A(v)f = H_\sigma^* g \quad (18)$$

$f$  is iteratively determined. To obtain  $f^{n+1}$ , a correction term  $d^n$  is added to the current value  $f^n$ :  $f^{n+1} = f^n + d^n$ .  $d^n$  is estimated by

$$A(v)d^n = H_\sigma^* g - A(v)f^n \quad (19)$$

In Appendix B it is shown that the operator  $A(v)$  is self-adjoint and positive definite. In the

discrete domain, equation (19) corresponds to a linear system of equations with a symmetric and positive definite matrix  $A(v)$ . Consequently, equation (19) is solved for  $d^n$  via the Conjugate Gradients method. In step 3, equation (10) takes the form

$$\sum_{i,j \in \Omega} (h_\sigma * f - g) \left( \frac{\partial h_\sigma}{\partial \sigma} * f \right)_{i,j} + \gamma \sum_{i,j \in \Omega} \left( \frac{\partial}{\partial \sigma} |\nabla h_\sigma|^2 \right)_{i,j} = 0, \quad (20)$$

where  $\partial h_\sigma / \partial \sigma$  and  $\partial(|\nabla h_\sigma|^2) / \partial \sigma$  are given in Eqs. (11) and (12) respectively. The equation was solved for  $\sigma$  using the bisection method. The discrete support of the Gaussian was limited to about  $20\sigma \times 20\sigma$ , which in our experiments was much smaller than the image size. The numerical integral of  $h_\sigma$  was normalized to 1.

## 5 Special case: known blur kernel

If the blur kernel is known, the restriction to Gaussian kernels is no longer necessary. In this case, the kernel-regularization term in the objective functional (7) can be omitted. Consequently, the algorithm can be simplified by skipping step 3 and replacing the stopping criterion by a simple convergence measure.

The resulting algorithm for coupled segmentation and image restoration is fast, robust and stable. Unlike [11], the discontinuity set is not restricted to isolated closed contours. Its performance is exemplified in Fig. 4. The top-left image in Fig. 4 is a blurred version of an original  $170 \times 170$  *Notes* image (not shown). The blur kernel corresponds to horizontal 9-pixel motion. The top-right image is the reconstruction obtained using total variation regularization [17,24], with a regularization parameter of  $3 \cdot 10^{-7}$ . The bottom-left image is the outcome of the proposed method ( $\beta = 10^{-4}, \alpha = 10^{-8}, \epsilon = 10^{-3}$ ), with a known blur kernel. The bottom-right image shows the associated edge map  $v$  determined by the algorithm.



Figure 4: The case of a known (9-pixel horizontal motion) blur kernel. *Top-left*: Corrupted image. *Top-right*: Restoration using the TV method [17,24]. *Bottom-left*: Restoration using the suggested method. *Bottom-right*: Edge map produced by the suggested method.

Acceptable restoration is obtained with both methods. Nevertheless, the suggested method yields a sharper result, and is almost free of “ghosts” (white replications of notes) that can be seen in the top-right image (e.g., between the C notes in the right part of the top staff). Computing time was about 2 minutes in interpreted MATLAB on a 2GHz PC.

## 6 Results: semi-blind restoration

Consider the example shown in Fig. 5. The top-left image was obtained by blurring the original  $200 \times 200$  *coin* image (not shown) with a Gaussian kernel with  $\sigma = 2.1$ . Restoration



Figure 5: Semi-blind restoration. *Top-left*: Blurred ( $\sigma = 2.1$ ) image. *Top-right*: Restoration using the method of [8]. *Bottom-left*: Restoration using a modified version of [8], in which the kernel was restricted to the family of isotropic Gaussians. *Bottom-right*: Restoration using the method suggested in this paper.

using [8] ( $\alpha_1 = 10^{-4}$ ,  $\alpha_2 = 10^{-5}$ ) is shown top-right. The bottom-left image is the outcome of a modified version of the method of [8]. The kernel was restricted to the family of isotropic Gaussians, and kernel smoothness was measured using the  $L_2$  norm. The parameters were tuned for best performance:  $\alpha_1 = 10^{-4}$  and  $\alpha_2 = 20$ . The resulting estimate of  $\sigma$  was 1.61. The gains brought by the restriction of the recovered kernel to a class of parametric kernels are clear. The image recovered using the method suggested in this paper shown bottom-right. In this case the parameters were  $\beta = 10^{-4}$ ,  $\alpha = 10^{-8}$ ,  $\epsilon = 10^{-3}$ ,  $\gamma = 20$  and the resulting estimate of  $\sigma$  was 2.05. Note that the bottom-right image, corresponding to



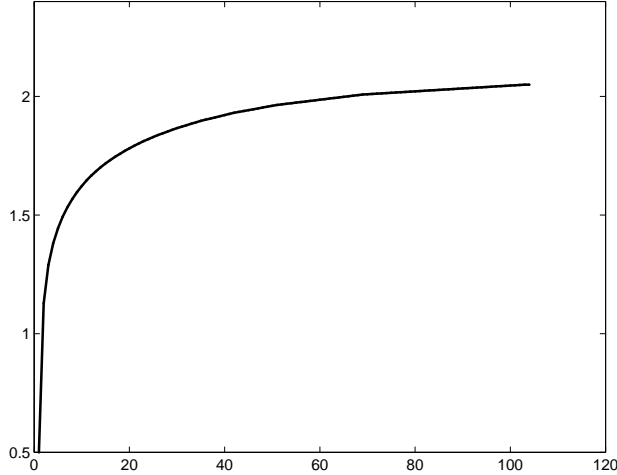


Figure 6: The convergence of the estimated width  $\sigma$  of the blur kernel as a function of the iteration number in the semi-blind recovery of the *noiseless coin* image (Fig. 5).

our method, is sharper than the bottom-left image. This highlights the advantage of the Mumford-Shah image regularizer over the  $L_1$  norm. The convergence process is illustrated in Fig. 6.

In Fig. 7, the blurred input image (top-left) is similar to that of Fig. 5, but with slight white Gaussian noise added (SNR=30dB). Shown top-right is the image recovered using the method of [8] ( $\alpha_1 = 10^{-4}, \alpha_2 = 10^{-4}$ ). The bottom-left image is the outcome of the modified version of the method of [8]. The kernel was restricted to the family of isotropic Gaussians and kernel smoothness was measured using the  $L_2$  norm. Here  $\alpha_1 = 0.01, \alpha_2 = 60$  and the resulting estimate of  $\sigma$  was 1.33. The gains brought by the restriction of the recovered kernel to a class of parametric kernels are again clear. The image recovered using the method suggested in this paper is shown bottom-right ( $\beta = 0.01, \alpha = 0.001, \epsilon = 0.1, \gamma = 60$ ). The resulting estimate of  $\sigma$  was 1.77.

Note that in the presence of noise, the image and kernel smoothness parameters are increased in order to compensate for the noise, and the resulting  $\sigma$  is decreased. Nevertheless, with or without noise, and even though  $\gamma = \alpha_2$ , the suggested method yields better (higher)



Figure 7: Semi-blind restoration with some additive noise (SNR=30dB). *Top-left*: Blurred ( $\sigma = 2.1$ ) and noisy image. *Top-right*: Restoration using the method of [8]. *Bottom-left*: Restoration using the modified version of [8], in which the kernel was restricted to the family of isotropic Gaussians. *Bottom-right*: Restoration using the suggested method.

estimates of  $\sigma$ , and consequently sharper recovered images.

In all the experiments described in the sequel, the parameters used in the algorithm were  $\beta = 10^{-4}$ ,  $\alpha = 10^{-8}$ ,  $\epsilon = 10^{-3}$  and  $\gamma = 40$ . The initial value of  $\sigma$  was 0.5 and the convergence tolerance was  $\varepsilon_2 = 10^{-2}$ .

Additional examples are shown in Figs. 8 and 9. The top, middle and bottom rows in both figures correspond to the blurred, recovered and edge images respectively. The images were blurred using a Gaussian kernel with  $\sigma = 2.1$  for the *Lena*, *Cameraman* and *Einstein* images, and  $\sigma = 2.6$  for the *Sails* image. Computing time in these examples was about 12 minutes in interpreted MATLAB on a 2GHz PC. The unknown width parameters of the blur



Figure 8: Semi-blind restoration. *Top row:* Blurred images. *Middle row:* Restoration using the suggested method. *Bottom row:* Edge maps produced by the suggested method.

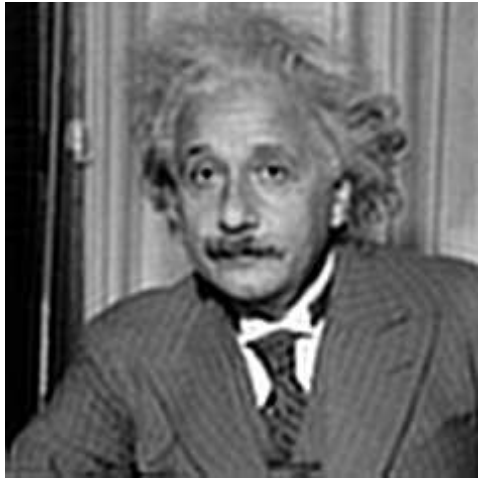
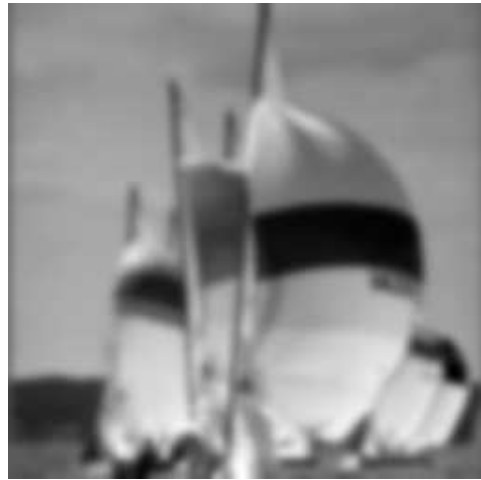
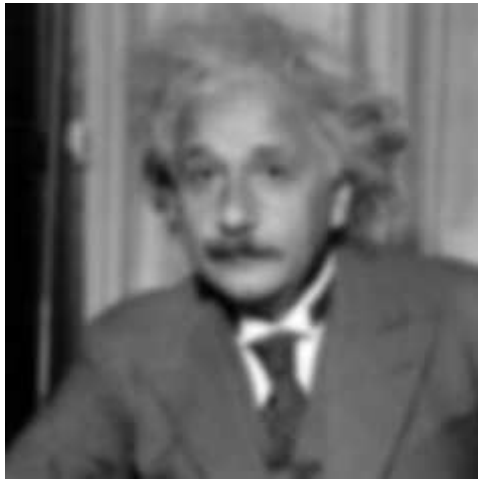


Figure 9: Semi-blind restoration. *Top row:* Blurred images. *Middle row:* Restoration using the suggested method. *Bottom row:* Edge maps produced by the suggested method.

kernels were estimated to be  $\sigma = 2.01, 1.95, 2.27, 2.57$ , and are in pleasing agreement with their true value.

## 7 Discussion

Inverse problems in image analysis are difficult and often ill-posed. This means that searching for the solution in the largest possible space is not always the best strategy. A-priori knowledge should be used, wherever possible, to limit the search and constrain the solution. In the context of pure blind restoration, Carasso [6] analyzes previous approaches and presents convincing arguments in favor of restricting the class of blurs.

Along these lines, in this paper the blur kernels are constrained to the class of isotropic Gaussians parameterized by their width. This is a sound approximation of physical kernels encountered in diverse contexts [6]. The advantages brought by this restriction are well-demonstrated in the experimental results that we provide.

In the suggested variational framework, the restored image is regularized using the Mumford-Shah terms ( $G_\epsilon(f, v)$ ). This is a better model than the  $L_1$  norm for the piecewise smoothness of common images, and leads to better experimental results. The experimental results also demonstrate the robustness of the algorithm with respect to the parameters. The images tested in Figs. 8 and 9 are quite different, yet they were all successfully restored using the same parameter set.

The functional (7) is not generally convex. Nevertheless, the algorithm gracefully converged to a visually appealing restoration in all our experiments. The risk of convergence to local minima is common and inherent to most variational image processing algorithms. Only in special cases can convergence to the global minimum be proved. This is an interesting

and challenging topic for future research.

## Acknowledgments

This research was supported by MUSCLE: Multimedia Understanding through Semantics, Computation and Learning, a European Network of Excellence funded by the EC 6th Framework IST Programme. It was also supported by the Israel Academy of Sciences. Leah Bar was supported by the Weinstein Center for Signal Processing Research at Tel Aviv University.

## Appendices

### A

In this appendix we derive the Euler-Lagrange equations (8),(9) that correspond to the objective functional (7):

$$\begin{aligned} \mathcal{F}_\epsilon(f, v, \sigma) &= \frac{1}{2} \int_{\Omega} (h_\sigma * f - g)^2 dA + \beta \int_{\Omega} v^2 |\nabla f|^2 dA + \\ &+ \alpha \int_{\Omega} \left( \epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right) dA + \gamma \int_{\Omega} |\nabla h_\sigma|^2 dA. \end{aligned}$$

In the following derivation, let  $f(x, y)$  be extended to the whole  $\mathbb{R}^2$  with the Neumann boundary condition, meaning that the function values are uniform along the normal of the image boundary. In addition, the integration element  $dA$  denotes  $dx dy$ .

### Variation with respect to $f$

The variation of  $\mathcal{F}_\epsilon$  with respect to  $f$  can be expressed as

$$\frac{\delta \mathcal{F}_\epsilon}{\delta f} = \frac{\partial}{\partial \lambda} \mathcal{F}_\epsilon(f + \lambda \eta) \Big|_{\lambda=0} = \frac{\partial}{\partial \lambda} [\mathcal{F}_\epsilon^I + \mathcal{F}_\epsilon^{II}] \Big|_{\lambda=0} = 0$$

where  $\eta \in C^1(\Omega)$  with

$$\mathcal{F}_\epsilon^I = \frac{1}{2} \int_{\Omega} [h_\sigma * (f + \lambda \eta) - g]^2 dA$$

and

$$\mathcal{F}_\epsilon^{II} = \beta \int_{\Omega} v^2 [(f_x + \lambda \eta_x)^2 + (f_y + \lambda \eta_y)^2] dA.$$

For the first part,

$$\frac{\partial \mathcal{F}_\epsilon^I}{\partial \lambda} \Big|_{\lambda=0} = \int_{\Omega} (h_\sigma * f - g)(h_\sigma * \eta) dA.$$

By substituting the expression for  $h_\sigma * \eta$  and exchanging variables we obtain

$$\frac{\partial \mathcal{F}_\epsilon^I}{\partial \lambda} \Big|_{\lambda=0} = \int_{\Omega} (h_\sigma * f - g) * h_\sigma(-x, -y) \eta(x, y) dA. \quad (21)$$

For the second part,

$$\frac{\partial \mathcal{F}_\epsilon^{II}}{\partial \lambda} \Big|_{\lambda=0} = 2\beta \int_{\Omega} v^2 (f_x \eta_x + f_y \eta_y) dA. \quad (22)$$

Integrating by parts yields

$$\begin{aligned} \frac{1}{2\beta} \frac{\partial \mathcal{F}_\epsilon^{II}}{\partial \lambda} \Big|_{\lambda=0} &= \int_{\Omega} \frac{\partial}{\partial x} (v^2 \eta f_x) dA + \int_{\Omega} \frac{\partial}{\partial y} (v^2 \eta f_y) dA - \\ &\quad - \int_{\Omega} \left[ \frac{\partial}{\partial x} (v^2 f_x) + \frac{\partial}{\partial y} (v^2 f_y) \right] \eta dA. \end{aligned}$$

By applying the divergence theorem to the vector field  $v^2 \eta \nabla f$ , it is seen that the first two terms are equal to the closed contour integral

$$\int_{\partial\Omega} v^2 \eta \nabla f \cdot dn \quad (23)$$

where  $n$  is the vector normal to the contour. Adopting the Neumann boundary condition

$$\frac{\partial f}{\partial n} = 0 \quad (x, y) \in \partial\Omega ,$$

integral (23) vanishes, hence equation (22) takes the form

$$\left. \frac{\partial \mathcal{F}_\epsilon^{II}}{\partial \lambda} \right|_{\lambda=0} = -2\beta \int_{\Omega} \left[ \frac{\partial}{\partial x} (v^2 f_x) + \frac{\partial}{\partial y} (v^2 f_y) \right] \eta \, dA . \quad (24)$$

Adding equations (21),(24) and using the fundamental lemma of calculus of variations we obtain Eq. (9):

$$\frac{\delta \mathcal{F}_\epsilon}{\delta f} = (h_\sigma * f - g) * h_\sigma(-x, -y) - 2\beta \text{Div}(v^2 \nabla f) = 0 .$$

### Variation with respect to $v$

The variation of  $\mathcal{F}_\epsilon$  with respect to  $v$  can be expressed as

$$\frac{\delta \mathcal{F}_\epsilon}{\delta v} = \left. \frac{\partial}{\partial \lambda} \mathcal{F}_\epsilon(v + \lambda\psi) \right|_{\lambda=0} = \left. \frac{\partial}{\partial \lambda} [\mathcal{F}_\epsilon^I + \mathcal{F}_\epsilon^{II}] \right|_{\lambda=0} = 0$$

where

$$\mathcal{F}_\epsilon^I = \beta \int_{\Omega} (v + \lambda\psi)^2 |\nabla f|^2 \, dA + \frac{\alpha}{4\epsilon} \int_{\Omega} (v + \lambda\psi - 1)^2 \, dA$$



and

$$\begin{aligned}
\mathcal{F}_\epsilon^{II} &= \alpha\epsilon \int_{\Omega} [(v_x + \lambda\psi_x)^2 + (v_y + \lambda\psi_y)^2] dA . \\
\frac{\partial \mathcal{F}_\epsilon^I}{\partial \lambda} \Big|_{\lambda=0} &= \int_{\Omega} \left[ 2\beta v |\nabla f|^2 + \frac{\alpha}{2\epsilon}(v-1) \right] \psi dA \\
\frac{\partial \mathcal{F}_\epsilon^{II}}{\partial \lambda} \Big|_{\lambda=0} &= 2\alpha\epsilon \int_{\Omega} [v_x\psi_x + v_y\psi_y] dA
\end{aligned} \tag{25}$$

As with Eq. (22), we integrate by parts, apply the divergence theorem and adopt the Neumann boundary condition

$$\frac{\partial v}{\partial n} = 0 \quad (x, y) \in \partial\Omega .$$

we obtain

$$\frac{\partial \mathcal{F}_\epsilon^{II}}{\partial \lambda} \Big|_{\lambda=0} = -2\alpha\epsilon \int_{\Omega} (v_{xx} + v_{yy})\psi dA. \tag{26}$$

Adding equations (25),(26) and using the fundamental lemma of calculus of variations we arrive at Eq. (8):

$$\frac{\delta \mathcal{F}_\epsilon}{\delta v} = \left( 2\beta |\nabla f|^2 + \frac{\alpha}{2\epsilon} - 2\alpha\epsilon \nabla^2 \right) v = \frac{\alpha}{2\epsilon}$$

## B

**Theorem.** *The operator  $A(v)$  defined as*

$$A(v) = H_\sigma H_\sigma^* - 2\beta \text{Div}(v^2 \nabla) \tag{27}$$

*is self adjoint and positive definite.*

*Proof.* Let  $A(v) = A^I + A^{II}$ , where

$$A^I u = H_\sigma H_\sigma^* u$$

and

$$A^{II} u = -2\beta \text{Div}(v^2 \nabla u) .$$

Let  $H$  be a convolution operator, i.e.,  $Hf(x) = h(x) * f(x)$  where  $x \in \mathbb{R}^2$ . Its adjoint operator  $H^*$  is defined by  $\langle g, Hf \rangle = \langle H^*g, f \rangle$ , where  $\langle \cdot, \cdot \rangle$  denotes inner product.

Here,

$$\langle H^*g, f \rangle = \langle g, Hf \rangle = \int_{\mathbb{R}^2} g \cdot (h * f) dx = \int_{\mathbb{R}^2} [h(-x) * g(x)] \cdot f(x) dx$$

hence  $H^*g(x) = h(-x) * g(x)$ . Now

$$A^{I*} u = h_\sigma(-x) * h_\sigma(x) * u = h_\sigma * h_\sigma(-x) * u = A^I u \quad (28)$$

Thus  $A^I$  is a self adjoint operator.

Let  $G(\xi)$  and  $H(\xi)$  be the Fourier transforms of  $g(x)$  and  $h_\sigma(x)$  respectively, and let  $\bar{G}(\xi)$ ,  $\bar{H}(\xi)$  be their complex conjugates. For a real function  $g : \Omega \rightarrow \mathbb{R}$ ,  $G(-\xi) = \bar{G}(\xi)$ .

$$\langle g, A^I g \rangle = \int_{\Omega} g(x) [h_\sigma(x) * h_\sigma(-x) * g(x)] dx = [G(\xi) * (H(\xi) \bar{H}(\xi) G(\xi))]_{\xi=0} .$$

Substituting the convolution operator yields

$$\langle g, A^I g \rangle = \left[ \int_{\mathbb{R}^2} G(\xi - \xi') |H(\xi')|^2 G(\xi') d\xi' \right]_{\xi=0} = \int_{\mathbb{R}^2} G(-\xi') G(\xi') |H(\xi')|^2 d\xi'$$

$$= \int_{\mathbb{R}^2} \bar{G}(\xi')G(\xi')|H(\xi')|^2 d\xi' = \int_{\mathbb{R}^2} |G(\xi')|^2 |H(\xi')|^2 d\xi' > 0$$

for all functions  $g$  and kernels  $h_\sigma$  that are not identically zero, which proves that  $A^I$  is positive definite.

For the second part of the operator,

$$\langle g, A^{II} f \rangle = -2\beta \int_{\Omega} g [\nabla \cdot (v^2 \nabla f)] dx .$$

Recall that for a scalar function  $g$  and a vector field  $\phi$ ,

$$g \nabla \cdot \phi = \nabla \cdot (g\phi) - \nabla g \cdot \phi \quad (29)$$

thus

$$\langle g, A^{II} f \rangle = -2\beta \int_{\Omega} \nabla \cdot (gv^2 \nabla f) dx + 2\beta \int_{\Omega} \nabla g \cdot v^2 \nabla f dx$$

Applying the divergence theorem,

$$\langle g, A^{II} f \rangle = -2\beta \int_{\partial\Omega} (gv^2 \nabla f) \cdot dn + 2\beta \int_{\Omega} \nabla g \cdot v^2 \nabla f dx \quad (30)$$

Using the Neumann boundary condition, the first term vanishes, hence

$$\langle g, A^{II} f \rangle = 2\beta \int_{\Omega} \nabla g \cdot v^2 \nabla f dx = 2\beta \int_{\Omega} \nabla f \cdot v^2 \nabla g dx \quad (31)$$

Substituting equation (29) in (31) and using the divergence theorem we obtain

$$\langle g, A^{II} f \rangle = 2\beta \int_{\partial\Omega} (gv^2 \nabla f) \cdot dn - 2\beta \int_{\Omega} f \nabla \cdot (v^2 \nabla g) dx$$

Applying again the Neumann boundary condition the first term vanishes and we observe that

$$\langle g, A^{II} f \rangle = -2\beta \int_{\Omega} \nabla \cdot (v^2 \nabla g) f \, dx = \langle A^{II} g, f \rangle$$

which proves that the operator  $A^{II}$  is self adjoint. We proceed to show that  $A^{II}$  is positive definite:

$$\langle g, A_{II} g \rangle = -2\beta \int_{\Omega} g \nabla \cdot (v^2 \nabla g)$$

Applying Eq. (29) and using the divergence theorem yields

$$\langle g, A_{II} g \rangle = -2\beta \int_{\partial\Omega} (g v^2 \nabla g) \cdot dn + 2\beta \int_{\Omega} \nabla g \cdot (v^2 \nabla g) \, dx$$

The first term vanishes due to the Neumann boundary condition, thus

$$\langle g, A_{II} g \rangle = 2\beta \int_{\Omega} v^2 |\nabla g|^2 \, dx > 0$$

we conclude that  $A^{II}$  is positive definite. Since both  $A^I$  and  $A^{II}$  are self adjoint and positive definite, their sum  $A(v)$  is also self adjoint and positive definite.  $\square$

## References

- [1] L. Ambrosio and V.M. Tortorelli, “Approximation of Functionals Depending on Jumps by Elliptic Functionals via  $\Gamma$ -Convergence”, *Communications on Pure and Applied Mathematics*, Vol. XLIII, pp. 999-1036, 1990.
- [2] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing*, Springer, New York, 2002.

- [3] M. Banham and A. Katsaggelos, “Digital Image Restoration”, *IEEE Signal Processing Mag.*, Vol. 14, pp. 24-41, 1997.
- [4] L. Bar, N. Sochen and N. Kiryati, “Variational Pairing of Image Segmentation and Blind Restoration”, Proc. ECCV’2004, Prague, Czech Republic, Part II: *LNCS #3022*, pp. 166-177, Springer, 2004.
- [5] A. Braides, *Approximation of Free-Discontinuity Problems*, pp. 47-51, *Lecture Notes in Mathematics*, Vol. 1694, Springer, 1998.
- [6] A. S. Carasso, “Direct Blind Deconvolution”, *SIAM J. Applied Math.*, Vol. 61, pp. 1980-2007, 2001.
- [7] A. Chambolle, “Image Segmentation by Variational Methods: Mumford and Shah functional, and the Discrete Approximation”, *SIAM Journal of Applied Mathematics*, Vol. 55, pp. 827-863, 1995.
- [8] T. Chan and C. Wong, “Total Variation Blind Deconvolution”, *IEEE Trans. Image Processing*, Vol. 7, pp. 370-375, 1998.
- [9] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud, “Deterministic Edge-Preserving Regularization in Computed Imaging”, *IEEE Trans. Image Processing*, Vol. 6, pp. 298-311, 1997.
- [10] D. Geman and C. Yang, “Nonlinear Image Recovery with Half-Quadratic Regularization and FFT’s”, *IEEE Trans. Image Processing*, Vol. 4, pp. 932-946, 1995.
- [11] J. Kim, A. Tsai, M. Cetin and A.S. Willsky, “A Curve Evolution-based Variational Approach to Simultaneous Image Restoration and Segmentation”, *Proc. IEEE ICIP*, Vol. 1, pp. 109-112, 2002.

- [12] D. Kundur and D. Hatzinakos, “Blind Image Deconvolution”, *Signal Processing Mag.*, Vol. 13, pp. 43-64, May 1996.
- [13] D. Kundur and D. Hatzinakos, “Blind Image Deconvolution Revisited”, *Signal Processing Mag.*, Vol. 13, pp. 61-63, November 1996.
- [14] D. Mumford and J. Shah, “Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems”, *Communications on Pure and Applied Mathematics*, Vol. 42, pp. 577-684, 1989.
- [15] R. Nakagaki and A. Katsaggelos, “A VQ-Based Blind Image Restoration Algorithm”, *IEEE Trans. Image Processing*, Vol. 12, pp. 1044-1053, 2003.
- [16] T. Richardson and S. Mitter, “Approximation, Computation and Distortion in the Variational Formulation”, in *Geometry-Driven Diffusion in Computer Vision*, B.M. ter Harr Romeny, Ed. Kluwer, Boston, 1994, pp. 169-190.
- [17] L. Rudin, S. Osher and E. Fatemi, “Non Linear Total Variation Based Noise Removal Algorithms”, *Physica D*, Vol. 60, pp. 259-268, 1992.
- [18] L. Rudin and S. Osher, “Total Variation Based Image Restoration with Free Local Constraints”, *Proc. IEEE ICIP*, Vol. 1, pp. 31-35, Austin TX, USA, 1994.
- [19] C. Samson, L. Blanc-Féraud, G. Aubert and J. Zerubia, “Multiphase Evolution and Variational Image Classification”, *Technical Report No. 3662*, INRIA Sophia Antipolis, April 1999.
- [20] M. Sonka, V. Hlavac and R. Boyle, *Image Processing, Analysis and Machine Vision*, PWS Publishing, 1999.

- [21] F. Sroubek and J. Flusser, “Multichannel Blind Iterative Image Restoration”, *IEEE Trans. Image Processing*, Vol. 12, pp. 1094-1106, 2003
- [22] A. Tikhonov and V. Arsenin, “Solutions of Ill-posed Problems”, New York, 1977.
- [23] L.A. Vese and T.F. Chan, “A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model”, *International Journal of Computer Vision*, Vol. 50, pp. 271-293, 2002.
- [24] C. Vogel and M. Oman, “Fast, Robust Total Variation-based Reconstruction of Noisy, Blurred Images”, *IEEE Trans. Image Processing*, Vol. 7, pp. 813-824, 1998.
- [25] E.W. Weisstein *et al*, “Minimal Residual Method”, from *MathWorld—A Wolfram Web Resource*. <http://mathworld.wolfram.com/MinimalResidualMethod.html>
- [26] Y. You and M. Kaveh, “A Regularization Approach to Joint Blur Identification and Image Restoration”, *IEEE Trans. Image Processing*, Vol. 5, pp. 416-428, 1996.