Bayesian Estimation Based Mumford-Shah Regularization for Blur Identification and Segmentation in Video Sequences

Hongwei Zheng and Olaf Hellwich
Computer Vision and Remote Sensing, Berlin University of Technology
Franklinstrasse 28/29, Sekretariat FR 3-1, D-10587 Berlin
{hzheng,hellwich}@cs.tu-berlin.de

Abstract

We present an extended Mumford-Shah (MS) regularization for blind image deconvolution and segmentation in the context of Bayesian estimation. The extended MS functional is added to have costs for the identification of blur via a newly introduced prior solution space. The functional is minimized using \(\Gamma\)-convergence approximation by projecting iterations onto a newly designed embedded alternating minimization within Neumann conditions. Image segmentation is closely related to accurate blur identification and restoration, that is, the problem of estimating an image based on its degraded observation. Experiments show that the proposed algorithm is efficient and robust in that it can handle images that are formed in different environments with different types and amounts of blur and noise.

1. Introduction

The challenge of blind image deconvolution is to uniquely define the optimized signals from the degraded image with unknown blur information in an ill-posed inverse condition. The ideal image \(f\) in the object plane is normally degraded by a linear space-invariant point spread function (PSF) \(h\) with an additive white Gaussian noise \(n\) using the lexicographic notation, \(g = hf + n\). The equation provides a good working model for image formation.

A general regularization method proposed by Mumford and Shah [9] has been formulated in an energy minimization approach. Currently, most Mumford-Shah (MS) based segmentation approaches combining with the level set method (LST) are intensively tested on the influences of noises or occlusions [4] and get successful results. However, the method is not suitable to segment blurred objects due to unstable and weak differences of gradients between foreground objects and cluttered background, i.e., Fig. (1).

Moreover, variational regularization needs effective prior information or constraints to yield a unique solution to the corresponding optimization procedure. The Bayesian estimation provides a structured way to include prior knowledge concerning the quantities to be estimated. The Bayesian approach is, in fact, the framework in which most recent restoration methods have been introduced [3][5][6]. Furthermore, [8] have also reported that the obtained statistic estimates for the PSF could vary significantly, depending on the initialization.

Combination of blur identification, image restoration and segmentation in the extended MS functional is a reasonable strategy due to the mutual support within a regularized Bayesian approach. This has some important effects: firstly, Bayesian MAP estimation supports a good initial value for the optimization to the soft extended Mumford-Shah regularization. Secondly, it becomes possible to get edge-preserving image restoration in the extended MS regularization via a \(\Gamma\)-convergence approximation [1]. Finally, an embedded alternate minimization method can achieve the outputs without scale problems. One output of the finite sets of curves and object boundaries can be considered as discrete analogues of graphs. A graph-theory is integrated to the MS functional for partitioning and grouping these different gradient edges. The experimental results show that the method yields good segmentation results as well as edge-preserving restoration.

The paper is organized as follows. In Sect. (2), the \(\Gamma\)-convergence approximated MS functional is extended for blur estimation. The Bayesian MS functional is described in Sect. (3). Experimental results are shown in Sect. (4). Conclusions are summarized in Sect. (5).

2. Extended \(\Gamma\)-Convergence MS Functional

The basic idea of the Mumford-Shah functional is to subdivide an image into many meaningful regions (objects). It means to find a decomposition into regions \(\Omega_i\) of an image \(\Omega\) and an optimal piecewise smooth approximation \(f\) given
an observed image \( g \). Thus, the estimated \( f \) varies smoothly within each \( \Omega_i \), and discontinuously across the boundaries of \( \Omega_i \). The formula was considered as an energy \( E(f, C) \) minimization problem,

\[
E(f, C) = \int_{\Omega} (f - g)^2 \, dA + \alpha \int_{\Omega_0} |\nabla f|^2 \, dA + \beta |C|
\]

where, \( dA = dx \, dy \), \( \Omega \subset \mathbb{R}^2 \) is a connected, bounded and open subset \( \mathbb{R}^2 \), \( f \) is the smoothed image \( \subset \Omega \setminus C \); \( g : \Omega \to R \) is a bounded image-function with uniform feature intensity, \( C \subset \Omega \) is a finite set of segmenting curves and unit of object boundaries, \( |C| \) is the length of curve of \( C \).

By minimization of the MS functional \( E(f, C) \), it is difficult to derive the set \( C \) numerically, keep track of possible changes of its topology, and calculate its length. Also, the number of possible discontinuity sets is enormous even on a small grid. To solve such difficulties, Ambrosio and Tortorelli [1] applied a \( \Gamma \)-convergence approximation to the Mumford-Shah functional which means to replace \( C \) by a continuous variable \( \nu \) in the third term \( |C| = \int_{\Omega} (\nu |v|^2 + \frac{(\nu - 1)^2}{4\epsilon}) \, dA \). An irregular functional \( E(f, C) \) is then approximated by a sequence \( E_\nu(f) \) of regular functionals, \( \lim_{\nu \to 0} E_\nu(f) = E(f, C) \) and the minimization of \( E_\nu \) approximates the minimization of \( E \). The edge set is represented by a characteristic function \( (1 - \chi_C) \) which is approximated by an auxiliary function \( \nu(x) \) of the gradient edge integration map, i.e., and \( \nu(x) \approx 1 \) for getting edges.

Based on the \( \Gamma \)-convergence of MS equation, We use the degradation model \( h * f \) instead of \( f \) in the first fidelity term. The estimates to the original image and the blur kernel are denoted as \( \hat{f} \) and \( \hat{h} \) separately. The functional thus takes the form

\[
E_\nu(\hat{f}, \hat{h}, \nu) = \int_{\Omega} (\hat{f} * \hat{h} - g)^2 \, dA + \alpha \int_{\Omega} \nu^2 |\nabla \hat{f}|^2 \, dA + \beta \int_{\Omega} \left( \nu |\nabla \hat{v}|^2 + \frac{(\nu - 1)^2}{4\epsilon} \right) \, dA + \gamma \int_{\Omega} |\nabla \hat{h}|^2 \, dA + \delta |\hat{h} - \hat{h}_f|^2
\]

where \( \hat{h}_f \) is the final estimated PSF in each iteration and \( \hat{h} \) is the current estimated PSF. The fourth term \( \gamma \int_{\Omega} |\nabla \hat{h}|^2 \, dA \) represents the regularization of the blur kernel. This term is necessary to reduce the ambiguity in the division of the apparent blur between the recovered image and the blur kernel. The flexibility of the last term \( \delta |\hat{h} - \hat{h}_f|^2 \) is the PSF learning error which can adjust and incorporate the relevance of the parametric model of PSF throughout the blur identification and image restoration.

3. Extended MS in Bayesian Estimation

The Bayesian MAP estimation is utilized to get a solution integrating from some prior knowledge. Following Bayesian paradigm, the estimated image \( \hat{f} \), the estimated PSF \( \hat{h} \) and the observed image \( g \) are based on,

\[
p(\hat{f}, \hat{h} | g) = p(g | \hat{f}, \hat{h}) p(\hat{f}, \hat{h}) / p(g)
\]

Applying the Bayesian paradigm to the blind deconvolution problem, we try to get convergence values from Eq. (1) with respect to the estimated image \( f \) and the estimated PSF \( h \). This Bayesian MAP approach can also be seen as a regularization approach which combines optimization method for the minimization of two proposed cost functions in the image domain and the PSF domain. The cost function \( E \) in maximum a posterior (MAP) estimation of the restored image \( f \) and PSF \( h \) from Eq. (1) deduced as follows,

\[
E(\hat{f}(\nu, h), \hat{h}) \propto p(g | \hat{f}, \hat{h}) p(\hat{f})
\]

\[
E(\hat{h}(\nu, f), \hat{f}) \propto p(g | \hat{f}, \hat{h}) p(\hat{f})
\]

Some constraints are assumed for the application of these equations due to the fact that image pixels are independent identically distributed without influencing pixel correlations. The proposed prior solution space supports Bayesian estimation of parametric PSFs. The reason is that PSFs of numerous real blurred images satisfy up to a certain degree parametric structure. Also, most PSFs exist in the form of low-pass filters. The proposed method attempts to address these asymmetries by integrating parametric blur knowledge into the scheme of the extended Mumford-Shah regularization.

3.1. Solution Space of Blur Kernel Priors

We define a set \( \Theta \) as a solution space of Bayesian estimation which consists of primary parametric PSF models as \( \Theta = \{h_i(\theta), i = 1, 2, 3, \ldots, N \} \). \( h_i(\theta) \) represents the \( i \)th parametric PSF with general parameters \( \theta \), and \( N \) is the number of PSFs.

\[
h_i(\theta) =
\begin{cases}
  h_1(\theta) \propto h(x, y, \theta_1, \theta_2) = 1/\theta_1, & \text{if } |\theta_1| \leq \theta_1 \text{ and } |\theta_2| \leq \theta_2 \\
  h_2(\theta) \propto h(x, y, \theta_3) = K \exp(-\frac{x^2 + y^2}{\theta_3^2}) & \text{if } \sqrt{x^2 + y^2} \leq \theta_3 \\
  h_3(\theta) \propto h(x, y, \theta_4) = 1/\theta_4, & \text{if } \theta_4 \leq \theta_4 \\
  h_4(\theta) = K \text{ a box filter with a length of radius } K \\
  h_5(\theta) \text{ is a Gaussian PSF and can be characterized by parameters with its variance } \sigma^2 \text{ and a normalization constant } K \\
  h_6(\theta) \text{ is a simple linear motion blur PSF with a camera direction motion } d \text{ and a motion angle } \phi 
\end{cases}
\]
blur structures like out-of-focus and uniform 2D blur are also built in the solution space as a priori information. The solution space is then constructed by a set of predefined parametric PSFs for estimation in the Bayesian MAP Estimation.

3.2. Estimation in the Image Domain

The fundamental goal is to find a unique restored image \( f \) given an observed image \( g \) and the estimated PSF \( \hat{h} \) in the cost value of \( E \). The cost in the image domain can be minimized using the following function, \( p(g|\hat{f},\hat{h}) \) follows a fidelity distribution in the MS regularization and a prior \( p(\hat{f}) \) follows an image smoothing term. The cost function can be deducted as the following,

\[
E_c(\hat{f}|g,\hat{h}) = \arg \max_{\hat{f}} [p(g|\hat{f},\hat{h}) p(\hat{f})]
\]

\[
= \frac{1}{2} \int_\Omega (g - \hat{h} \star \hat{f})^2 dA + \alpha \int_\Omega |\nabla \hat{f}|^2 dA
\]

where \( p(g|\hat{f},\hat{h}) \propto \exp \left\{ -\frac{1}{2} \int_\Omega (g - \hat{h} \star \hat{f})^2 dA \right\} \) and \( p(\hat{f}) \) is prior information for the image, then \( p(\hat{f}) \propto \exp \left\{ \alpha \int_\Omega |\nabla \hat{f}|^2 dA \right\} \).

3.3. Estimation in the PSF Domain

The estimation of PSF is the starting point for the image estimation. In the PSF domain, the PSF can be seen as the maximization of conditional probability. The cost function of the PSF is described using the extended Mumford-Shah functional,

\[
E_v(\hat{g}|\hat{f},\hat{h}) = \arg \max_{\hat{h}} \left\{ p(\hat{g}|\hat{h},\hat{f}) p_{\hat{h}}(\hat{h}) \right\}
\]

\[
= \frac{1}{2} \int_\Omega (g - \hat{h} \star \hat{f})^2 dA + \gamma \int_\Omega |\nabla \hat{h}|^2 dA + \delta(\hat{h} - \hat{h}_f)^2
\]

where \( p(\hat{g}|\hat{h},\hat{f}) \propto \exp \left\{ -\frac{1}{2} \int_\Omega (g - \hat{h} \star \hat{f})^2 dA \right\} \), \( p_{\hat{h}}(\hat{h}) \propto \exp \left\{ \gamma \int_\Omega |\nabla \hat{h}|^2 dA + \delta(\hat{h} - \hat{h}_f)^2 \right\} \) is the prior knowledge. Since both the original and observed image represent intensity distributions that cannot take negative values, the PSF coefficients are always nonnegative, \( h(x) \geq 0 \). Furthermore, since image formation systems normally do not absorb or generate energy, the PSF should satisfy \( \sum_{x \in \Omega} h(x) = 1.0 \). \( x \in \Omega, \Omega \subset \mathbb{R}^2 \) is known.

3.4. From Learnt PSF Statistics to PSF Estimation

During the estimation process, each estimated PSF \( \hat{h} \) for a sampling image region normally has some noise error \( h_j = \hat{h} + \text{noise} \). To achieve a better estimated PSF with less noise, the estimated PSF is convolved with several low-pass filters to remove noise and get several PSF neighbors \( h_j, j \in \{1, ..., K\} \). In order to study the interaction between statistical blur kernel knowledge and blur degraded image information, we define the likelihood \( P(\hat{h}_j) \) of the estimated PSF \( \hat{h}_j \) of an observed image in resembling the \( i \)th parametric model \( h_i(\theta) \) in a multivariate Gaussian distribution,

\[
P(\hat{h}_j) = P \left( h_i(\theta) | \hat{h}_j \right) = \exp \left[ -\frac{1}{2} \left( h_i(\theta) - \hat{h}_j \right)^T \sum_{dd}^{-1} \left( h_i(\theta) - \hat{h}_j \right) \right]
\]

\[
(2\pi)^k \sum_{dd}^{-\frac{1}{2}}
\]

The first subscript \( i \) denotes the index of blur kernel. The modeling error \( d = h_i(\theta) - \hat{h}_j \) is assumed to be a zero-mean homogeneous Gaussian distributed white noise process with covariance matrix \( \sum_{dd} = \sigma^2 I \) independent of image. \( LB \) is assumed support size of blur. Then the Gaussian probability corresponds to a PSF learning:

\[
l_i(\hat{h}_j) = \frac{1}{2} \exp \left( (h_i(\theta) - \hat{h}_j)^T \sum_{dd}^{-1} (h_i(\theta) - \hat{h}_j) \right)
\]

In reality, most of blurs satisfy up to a certain degree parametric structures. A best fit model \( h_i(\theta) \) for \( \hat{h}_j \) is selected according to the Gaussian distribution and a weighted mean filter. The mean value of PSF learning likelihood \( l_i(\hat{h}_j) \) is that \( l_i(\hat{h}_j) \) is weight divided by \( d_i(\hat{h}_j) \). \( d_i(\hat{h}_j) \) is the Euclidean distance between \( \hat{h}_j \) and its neighbor \( h_j \), \( d_i(\hat{h}_j) = \sum_{i} E(\hat{h}_j) d^2(\hat{h}_j) / |d^2(\hat{h}_j)| \).

The weighted mean likelihood \( l_i(\hat{h}_j) \) depends on two conditions. The first condition is the likelihood value of the blur manifold \( l_i(\hat{h}_j) \), and the second is the distance between \( \hat{h}_j \) and its neighbor \( h_j \). The estimated output blur \( \hat{h}_f \) is obtained from the parametric blur models using

\[
\hat{h}_f = [\hat{h}_0(\hat{h}) + \sum_{i=1}^{C} l_i(\hat{h}) P(\hat{h}_j)/\sum_{i=1}^{C} l_i(\hat{h})]
\]

where \( \hat{h}_0(\hat{h}) = 1 - \max(l_i(\hat{h})), i = 1, ..., C \).

The main objective of this equation is to assess the relevance of current estimated PSF \( \hat{h} \) with respect to parametric PSF models, and integrates such knowledge progressively into the computation scheme. If the current estimated PSF \( \hat{h} \) is closely with the estimated PSF \( \hat{h}_f \), that means \( \hat{h} \) belongs to a predefined parametric blur structure. Otherwise, if \( \hat{h} \) differs from \( \hat{h}_f \) significantly, the current blur \( \hat{h} \) may not belong to the predefined PSF priors.

4. Experiments

4.1. Embedded Alternate Minimization

To achieve the results from the extended MS functional, a scale problem arises between the minimization of the PSF
and the image via gradient descent. The reason is that the \(\partial E_{\kappa}/\partial h\) is \(\sum_{x=0} f(x)\) times larger than \(\partial E_{\kappa}/\partial \hat{f}\). Also, the dynamic range of the image \([0, 255]\) is larger than the dynamic range of the PSF \([0, 1]\). The scale factor changes dynamically with space coordinates \((x, y)\). To avoid the scale problem, an alternate minimization method \([11]\) is introduced following the idea of coordinate descent \([7]\).

To minimize the energy cost \(E_{\kappa}\), three outputs of the ideal image \(\hat{f}\), the edge integration map \(v\) and the PSF \(h\) are computed for getting an optimized value from the partial differential equation of \(E_{\kappa}\). The minimization of this equation with respect to \(v\), \(h\) and \(\hat{f}\) is carried out based on Euler-Lagrange equations. We can observe that the equation \(\partial E_{\kappa}/\partial v\) is a strictly convex and lower bounded with respect to the functions \(\hat{f}\) and \(v\), while \(h\) is fixed. We have designed an embedded alternating minimization algorithm to get local minimum values simultaneously based on three partial derivatives,

\[
\frac{\partial E_{\kappa}}{\partial \hat{f}} = (\hat{h} - g)\hat{f}(x, y) - 2\gamma D\text{iv}(\nabla \hat{h}) - 2\delta|\hat{h} - \hat{h}| f
\]

\[
\frac{\partial E_{\kappa}}{\partial h} = (\hat{f} - g)\hat{f}(x, y) - 2\gamma D\text{iv}(\nabla \hat{h}) - 2\delta|\hat{h} - \hat{h}| f
\]

\[
\frac{\partial E_{\kappa}}{\partial v} = 2\alpha\sigma|\nabla \hat{f}|^2 - 2\alpha\sigma^2|\nabla v|^2
\]

where \(\partial E_{\kappa}\) is \(\partial E_{\kappa}(\hat{f}, \hat{h}, v)\). For solving these three equations, the Neumann conditions \(\partial E_{\kappa}/\partial v = 0\), \(\partial E_{\kappa}/\partial \hat{h} = 0\) and \(\partial E_{\kappa}/\partial \hat{f} = 0\) correspond to the reflection of the image across the boundary with the advantages of not imposing any value on the boundary. \(\varepsilon\) is a small positive constant for discrete implementation. The small positive constant can help to keep the estimation of image relatively stable in the minimization process. Based on an initial PSF value \(h_0(x)\), the estimation of the ideal image \(\hat{f}\) is initialized by the observed image \(g\), edge parameter \(v = 1\). The algorithm is summarized as:

**Initialization:**

\(f_0(x) = g(x), v = 1, h_0(x)\) is random numbers while \(\text{mse}_1 >\) \(c_1\)

1. **Initial:** \(f_n(x) = \arg\min(\hat{f}h_{n-1}, g), \text{fix} \hat{h}(x)\) while \(\text{mse}_2 > c_2\)

   (i) \(v = \arg\min(v|f)| = \partial E_{\kappa}/\partial v, \text{fix f}(x)\)

   (ii) \(f = \arg\min(f|v) = \partial E_{\kappa}/\partial \hat{f}, \text{fix v}(x)\)

2. \((n + 1)\) **Initial:** \(h_{n+1} = \arg\min(h|\hat{f}, g), \text{fix f}(x)\)

Global convergence can be reached given small positive thresholds \(\varepsilon_1\) and \(\varepsilon_2\) due to the nonnegativity of the image and the PSF. We use normalized mean square values of the PSF and the image to measure the minimization with respect to the thresholds.

**4.2. Results on Synthetic and Real Data**

In the first experiment, we have compared the results of image restoration using the weighted \(L^2\) Tikhonov regularization and the extended MS regularization. The parameters in the Mumford-Shah functional are tuned for the best performance \(\alpha = 10^{-4}, \beta = 10^{-8}, \varepsilon = 10^{-3}\) with an estimated PSF from regularization. In Fig. (2), (a) is motion blur with 20 dB Gaussian noise; (b) is the original PSF; (c) is the estimated PSF in 2D; (d) is the restored image using the \(L^2\) Tikhonov regularization and the estimated PSF; (e) is the restored image using the Mumford-Shah regularization and estimated PSF. From the results, we can easily observe that the result from extended Mumford-Shah functional is sharper and has less ringing artifacts compared to the weighted \(L^2\) Tikhonov regularization using the same estimated PSF. It highlights that the restoration using extended MS regularization is inclined towards edge-
preserving restoration.

This experiment is tested on real-world video data. For cluttered images in Fig. (1), some of closed contour curves using curve evolution technique [4] easily disappear due to some boundary leaks, non-average gradient differences and cluttered background. Segmentation of a blurred, noisy video sequence has good performance using the suggested method shown in Fig. (3). One advantage of our method is that the discontinuity set is not restricted to isolated closed contours. Cluttered background objects with stronger gradients do not influence the segmentation of blurred objects with unstable and lower gradients. The MS functional can achieve accurate edge detection, as initialized as 1, the edges are computed after a few iterations. These edges driven from the extended Mumford-Shah functional are grouped and thinned via an extended graph-grouping and partitioning $N_{cuts}$ [10] method into numerical labeled groups. The segmentation result is labeled and color filled following the partitioned regions.

5. Conclusions

Blind image deconvolution is an ill-posed inverse problem. Searching for the solution in a larger space is not a good strategy. To utilize accurate prior information directly in the computation is an excellent strategy since the approach improves the accuracy of initial values. The $T$-convergence approximated MS functional is extended to include cost terms for the estimation of blur kernels. The estimated PSF is not only based on the Bayesian MAP estimation but also optimized alternatingly in the regularization. The estimated image, the estimated PSF and edge curves are generated simultaneously from the extended MS functional. Furthermore, a graph-cuts method is integrated to group edges derived from the extended MS functional. These lower gradient blurred objects can then be segmented accurately without any prior knowledge. It is clear that the method is instrumental in image restoration and segmentation and can easily be extended in practical environments.

References