A Variational Framework for Simultaneous Motion Estimation and Restoration of Motion-Blurred Video

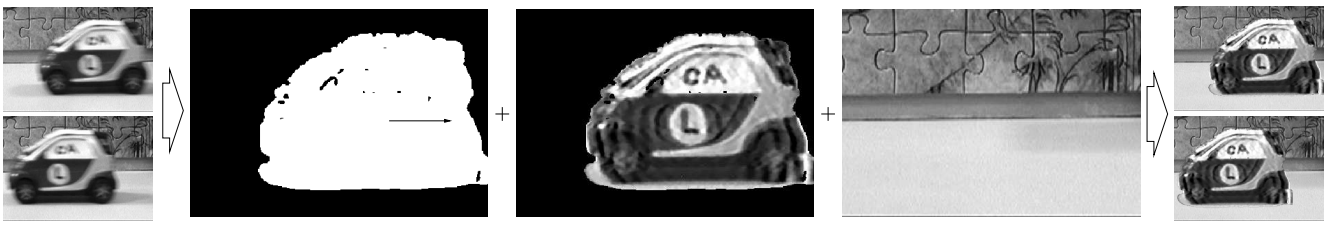

Figure 1. From two real blurred frames (left), we automatically and simultaneously estimate the motion region, the motion vector, and the image intensity of the foreground (middle). Based on this and the background intensity we reconstruct the two frames (right).

Leah Bar
Department of Electrical and Computer Engineering
University of Minnesota, Minneapolis, U.S.A.
barxx002@umn.edu

Benjamin Berkels
Institute for Numerical Simulation
University of Bonn, Germany
benjamin.berkels@ins.uni-bonn.de

Martin Rumpf
Institute for Numerical Simulation
University of Bonn, Germany
martin.rumpf@ins.uni-bonn.de

Guillermo Sapiro
Department of Electrical and Computer Engineering
University of Minnesota, Minneapolis, U.S.A.
guille@umn.edu

Abstract

The problem of motion estimation and restoration of objects in a blurred video sequence is addressed in this paper. Fast movement of the objects, together with the aperture time of the camera, result in a motion-blurred image. The direct velocity estimation from this blurred video is inaccurate. On the other hand, an accurate estimation of the velocity of the moving objects is critical for restoration of motion-blurred video. Therefore, restoration needs accurate motion estimation and vice versa, and a joint process is called for. To address this problem we derive a novel model of the blurring process and propose a Mumford-Shah type of variational framework, acting on consecutive frames, for joint object deblurring and velocity estimation. The proposed procedure distinguishes between the moving object and the background and is accurate also close to the boundary of the moving object. Experimental results both on simulated and real data show the importance of this joint estimation and its superior performance when compared to the independent estimation of motion and restoration.

AMS Subject Classifications: 65M60, 68U10, 65K10

1. Introduction

Motion estimation, that is, the computation of the velocity of moving objects in a given image sequence, is a well known problem in image processing and has received significant attention in recent years. Optical flow computation is one example of a widely used approach to motion estimation. Numerous methods have been developed to determine this flow, e.g., [10, 24]. One commonly known fact is that the clearer the sequence is, the more reliable the motion can be estimated. While certain robustness has been addressed in motion estimation, e.g., under varying illumination, [13], and contrast, [4], simple observation of the state-of-the-art literature in the subject immediately reveals that the videos are quite sharp and in general of sufficiently high quality. In particular, blurred video, see below, is very seldom considered in motion estimation techniques.

There are many real world effects on video footage which make motion estimation more difficult. In this paper, we address how to handle one of these critical effects. Considering video footage from a standard video camera, it is quite noticeable that relatively fast moving objects appear blurred (cf. Fig. 1). This effect is called *motion blur*, and it is caused by the way a camera takes pictures and is linked

to the aperture time of the camera, which roughly integrates information in time. The longer the aperture is open, or the faster the motion, the blurrier moving objects appear.

To improve the accuracy of the motion estimation on a video suffering from motion blur, it would be helpful to remove the motion blur first. On the other hand, if the actual motion is known, the motion blur can be removed by "deconvolution," since the motion gives the velocity of the objects and therefore the exact kernel needed for deconvolution. Realizing that these two problems are intertwined suggests to develop a method to tackle both problems at once.

In this paper we introduce a variational method which jointly handles motion estimation, moving object detection, and motion blur deconvolution (cf. Fig. 1). The proposed framework is a Mumford-Shah type of convex variational formulation, which includes explicit modelling of the motion-blur process as well as shape and image regularization terms, and is solved via efficient regularized decent techniques. The input to the variational formulation are two consecutive frames, while the output are the corresponding reconstructed frames, the segmented moving object, and the actual motion velocity. As demonstrated in this paper, this joint estimation of motion, moving object region, and reconstructed images, outperforms techniques where each individual unknown is individually handled.

Before proceeding with the explicit description of the proposed framework, let us illustrate this last point. For this, we use the image in Fig. 2, which although artificial, is very challenging and appropriate to demonstrate the advantage of joint estimation. In this figure, the Einstein insert f_{obj} is moving (velocity vector $v = (6, 7)$), while the Lena background f_{bg} is fixed. The independently computed velocity from the blurred frames leads to an inaccurate estimate of $v = (5.78, 6.80)$ and of the moving region (level-set of ϕ_0), which results in non-satisfactory restoration of the blurred frames (first image in second row of Fig. 3, see also Fig. 7). With our proposed joint technique, we obtained $v = (5.98, 7.009)$, and both the frames (last row of Fig. 2) and the moving region (blue curve, level line of ϕ in middle row of Fig. 2) are accurately recovered.

The remainder of this paper is organized as follows. After briefly presenting the related literature and a resume of our key contributions, we describe the motion model in Sec. 2 and derive our variational formulation in Sec. 3. Then, in Sec. 4 results of the joint approach are discussed. Section 5 is devoted to a detailed description of the energy minimization algorithm and in Sec. 6 we draw conclusion and give an outlook. The appendix contains a comprehensive collection of gradient components required in the algorithm.

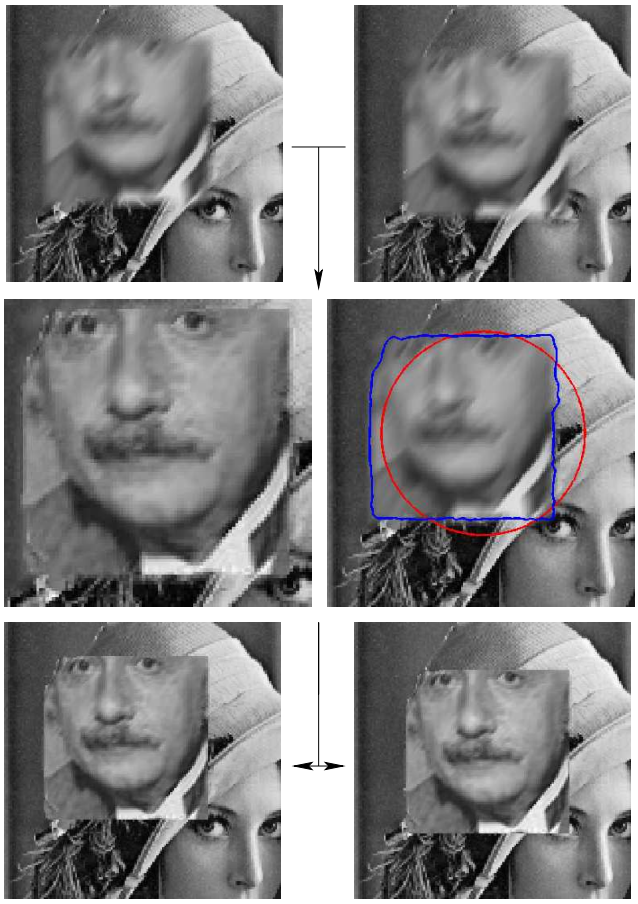

Figure 2. Results on an artificial motion blur sequence showing a square with a picture of Einstein moving on the Lena image as background. The input images g_1 and g_2 (top), the recovered object intensity f_{obj}, the initial boundary contour of the object (red) and the computed contour (blue) (middle row), and finally the recovered frames f_1 and f_2 (bottom) are depicted.

Figure 3. For the example from Fig. 2, intermediate results from our algorithm are depicted. In the top row from left to right the object contour is shown for three iterations from the initialization phase based on motion competition without deblurring. On the bottom three follow–up iterations of the joint method including the restoration of the frame are depicted.

1.1. Related works and key contributions

There exist numerous methods to remove motion blur using a single frame,[1], and these often introduce strong assumptions on the scene and/or blur [8]. As an example, let us mention the recent contribution on blind motion deblurring using image statistics presented in [17], were the author explains, as clear from the results, that while the image often well recovered, the actual motion and region of movement are often quite non-accurate. Another recent approach to motion deblurring [11] uses blending with the background but assumes the shift-variant case. Further [2] tackles piecewise shift-variant deblurring, including a segmentation of the blurred regions. Of more interest to our approach are techniques that use multiple frames, and these (some of them hardware based) are only very few, as summarized in [8]. More on the close connection between our work and [8] will be presented below.

Sequential motion estimation and then deblurring has been reported in [16] (see also [19]), while not addressing a truly joint estimation. The idea of developing joint methods for intertwined problems has become quite popular and successful recently, for example blind deconvolution and denoising [9], segmentation of moving objects in front of a still background and the computation of the motion velocities [14], segmentation and registration using geodesic active contours [12, 23], anisotropic classification and cartoon extraction [3], and optical flow computation and video denoising [18].

Motion deblurring can also be obtained with the so called "super-resolution framework," see [21] and references therein. The basic idea behind these approaches, which often assume that the blurring kernel is provided, is to obtain a higher resolution image from a collection of low-resolution frames. In addition, these techniques often assume that the whole frame suffers motion blur (or attack this with robust norms), and do not explicitly separate the moving object from the background or estimate the motion velocity.

The pioneering work by Favaro and Soatto, [8], is the closest to ours, not only because of the use of multiple frames but also because of the joint estimation. In a separate paper, they also [7] address the problem of simultaneously inferring the depth map, radiance and motion, from motion blurred and defocused. Thus, these works address the same challenges as we do here, which is the joint estimation of motion and scene deblurring from multiple frames. Some differences are that the authors of [8] approximate the motion blur with a Gaussian, rather than the more accurate rectangular filter, described in the next section. This model leads them to an anisotropic diffusion flow, and inverting it is ill posed. On the other hand, the variational formulation we propose here is well-posed and convex. The model in [8] is designed to handle only very little blur (motion), while the proposed method, as illustrated by the real examples below, can handle large velocities and blurs. We also model the crucial blending of the foreground and background, which happens in reality and significantly effects the blur as well as the reconstruction near the boundary of the moving object (see examples in Fig. 2,5,6). Finally, we note that while the proposed formulation could deal with multiple moving objects, in this paper we provide examples with only one, whereas [8] develop their work for multiple moving objects— although they present no examples of this capability with real video data.

To recap, this paper addresses the very important and challenging problem of *joint motion estimation and scene reconstruction from multiple frames*. This problem has been widely ignored in the literature, and ordinary motion estimation techniques assume sharp videos, while deblurring techniques often have other not always realistic assumptions. Furthermore, we incorporate a *motion blur model which is consistent at motion singularities*. The important differences with the only closely related method, proposed [8], are detailed above.

2. Modeling the blurring process

Images from an image sequences captured with a video camera are integrated measurements of light intensity emitted from moving objects over the aperture time interval of the camera. Let $f : [-T, T] \times \Omega;\ (t, x) \mapsto \mathbb{R}$ denote a continuous sequence of scene intensities over a time interval $[-T, T]$ and on a spatial image domain Ω observed via the camera lens. The video sequence recorded with the camera consists of a set of images $g_i : \Omega \to \mathbb{R}$ associated with times t_i, for $i = 1, \cdots, m$, given as the convolution

$$g_i(x) = \frac{1}{\tau} \int_{t_i - \frac{1}{2}\tau}^{t_i + \frac{1}{2}\tau} f(t + s, x)\, \mathrm{d}s \tag{1}$$

over the aperture time τ. For the time integral, we propose a box filter, which realistically approximates the mechanical shutters of film cameras and the electronic read out of modern CCD video recorders. In the simplest case, where the sequence f renders an object moving at constant velocity $v \in \mathbb{R}^2$, i.e. $f(x - sv) = f(t + s, x)$, we can transform integration in time to an integration in space and obtain for the recorded images

$$g_i(x) = \frac{1}{\tau} \int_{-\frac{1}{2}\tau}^{\frac{1}{2}\tau} f(x - sv)\, \mathrm{d}s = (f * h_v)(x), \tag{2}$$

for a one dimensional filter kernel $h_v = \delta_0(\frac{v^\perp}{|v|} \cdot y) h(\frac{v}{|v|} \cdot y)$ with filter width $\tau|v|$ in the direction of the motion trajectory $\{y = x + sv\ :\ s \in \mathbb{R}\}$. Here $v^\perp$ denotes v rotated by

[1] Similarly, the literature on motion estimation is abundant. Here we concentrate only on works addressing blurred video.

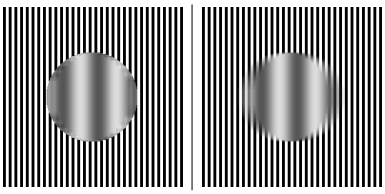

Figure 4. We consider a moving circle with back and white stripes in front of a similarly textured background. For this test case a comparison is shown between the wrong (left) motion blur model which ignores the motion discontinuity at the boundary and our realistic, consistent model (right) given in (4).

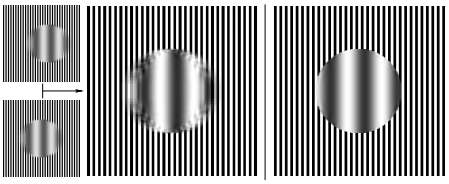

Figure 5. Given two frames for the realistic motion blur showing the moving circle on the texture background from Fig. 4 (left), computational results for the deblurring are depicted based on the wrong motion blur model built into G_i (middle), and on our consistent model (right). This clearly outlines the importance of a proper handling of the motion discontinuity in the considered motion blur model.

90 degrees, δ_0 is the usual 1D Dirac distribution and h the 1D block filter with $h(s) = \frac{1}{\tau|v|}$ for $s \in [-\frac{\tau|v|}{2}, \frac{\tau|v|}{2}]$ and $h(s) = 0$, else. In case of an object moving in front of a (still) background the situation is somewhat more complicated. At a point x close to the boundary of the object, the convolution (1) decomposes into a spatial convolution of object intensities along the motion path for the sub-interval of the aperture interval where the object covers the background at position x, and a retrieval of the background intensity for the remaining opening time of the lens. Figure 4 shows a comparison between the actually observed motion blur and results obtained by a (wrongly) direct application of the spatial convolution formula (2) on a moving circular object in front of a textured background (more specifics on this below). This observation is particularly important for the reliable recovery of boundaries of moving objects from recorded video frames g_i and subsequently for the proper restoration of image frames (cf. Fig. 5 for a corresponding comparison).

In what follows we consider an object moving with speed $v \in \mathbb{R}^2$ in front of a still background $f_{\text{bg}} : \Omega \to \mathbb{R}$ (which simplifies the formulation—see Sec. 6 for remarks on the generalization). The object at time 0 is represented by a intensity function $f_{\text{obj}} : \Omega_{\text{obj}} \to \mathbb{R}$ defined on an object domain Ω_{obj}. From f_{obj} and f_{bg} one assembles the actual scene intensity function

$$\begin{aligned} f(t,x) &= f_{\text{obj}}(x - tv)\chi_{\text{obj}}(x - vt) + \\ &\quad f_{\text{bg}}(x)(1 - \chi_{\text{obj}}(x - vt)) \end{aligned} \tag{3}$$

at time t, where $\chi_{\text{obj}} : \mathbb{R}^2 \to \mathbb{R}$ denotes the characteristic function of Ω_{obj}. Now, inserting (3) in (1) and then using (2) on Ω_{obj}, we deduce the correct formula for the theoretically observed motion blur at time t_i,

$$\begin{aligned} G_i[\Omega_{\text{obj}}, v, f_{\text{obj}}, f_{\text{bg}}](x) := ((f_{\text{obj}}\chi_{\text{obj}}) * h_v)(x - t_i v) + \\ f_{\text{bg}}(x)(1 - (\chi_{\text{obj}} * h_v)(x - t_i v)), \end{aligned} \tag{4}$$

for given object domain Ω_{obj}, motion velocity v, and object and image (background) intensity functions f_{obj} and f_{bg} respectively. If we do not carefully model the observed intensities as the moving object occludes and uncovers the background, we would observe $(f(t,\cdot) * h_v)$ on the object domain and f_{bg} elsewhere (cf. the combination of Eq. (14) and Eq. (3) in [8]). Given the more precise motion blur model proposed here, we now proceed to derive a variational formulation to simultaneously estimate all parameters in this equation based on two consecutive frames.

3. A Mumford–Shah model

Given two frames g_1 and g_2 of a video sequence with motion blur recorded at times t_1 and t_2, respectively, we construct a variational model to extract from these frames the domain Ω_{obj}, the image intensity f_{obj} of a moving object, and the motion velocity v. Here, we propose that the background intensity f_{bg} can a priori be extracted from the video sequence, for example, by averaging pixels with stable values over a sequence of frames. The formulation generalizes easily to include this estimation, as described in Sec. 2. We aim at formulating a joint energy for these degrees of freedom. Modeling this energy we take into account the following observations:

- Given v and intensity maps f_{obj}, $f_{\text{bg}} : \Omega \to \mathbb{R}$ (extended on the whole domain in a suitable way), we phrase the identification problem of the object boundary $\partial\Omega_{\text{obj}}$ in terms of a piecewise constant Mumford–Shah model. This appears to be well-suited in particular because the unknown contour is significantly smeared out due to the motion blur. Hence, a comparison of the expected motion blur G_i with the observed time frames g_i in a least square sense $\int_\Omega \left(G_i[\Omega_{\text{obj}}, v, f_{\text{obj}}, f_{\text{bg}}] - g_i\right)^2 \mathrm{d}x$ is considered as the fidelity energy, where the length of the boundary contour $|\partial\Omega_{\text{obj}}|$ acts as the corresponding prior.
- For known v and Ω_{obj}, we obtain an almost classical deblurring problem for f_{obj} with the modification of the blurring kernel given in (3), which is already reflected in the above fidelity term. We expect f_{obj} to be characterized by edges (cf. Fig. 1, 2, and 6). As a suitable prior for

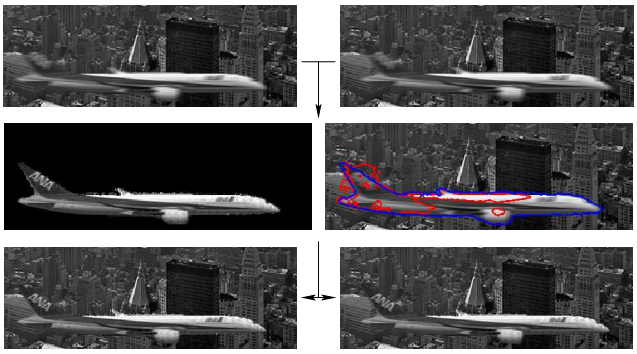

Figure 6. The performance of the joint model is shown in case of 2 consecutive images from an artificially blurred plane sequence. The input images g_1 and g_2 (top), the recovered object intensity f_{obj}, the zero contour of the level set function ϕ at two different relaxation step (in red and blue) of the algorithm (middle row), and finally the recovered frames f_1 and f_2 (bottom) are displayed.

these intensity maps we select the total variation functional $\int_\Omega |\nabla f_{\text{obj}}| \, \mathrm{d}x$ [20], which at the same time guarantees a suitable extension onto the whole space (cf. Fig. 2 for an example of the object intensity f_{obj} which is extended in a total variation consistent way to a neighborhood of the object domain Ω_{obj}).

• Finally, given Ω_{obj} and the two intensities f_{obj}, f_{bg}, the extraction of the motion velocity v is primarily an optical flow problem. The transport of the object intensity f_{obj} from time t_1 to t_2 described in G_1 and G_2 provides us with information on v. In the case of limited intensity modulations on the moving object, it is the comparison of the expected transition profile $\chi_{\text{obj}} * h_v$, encoded in G_i, with the observed profile in g_i that will act as a guidance for the identification of the motion velocity.

Based on these modeling aspects we finally obtain the energy

$$\mathcal{E}[\Omega_{\text{obj}}, v, f_{\text{obj}}] = \sum_{i=1,2} \int_\Omega \left(G_i[\Omega_{\text{obj}}, v, f_{\text{obj}}, f_{\text{bg}}] - g_i\right)^2 \, \mathrm{d}x + \int_\Omega \mu |\nabla f_{\text{obj}}| \, \mathrm{d}x + \nu |\partial \Omega_{\text{obj}}|, \tag{5}$$

and ask for a minimizing set of the degrees of freedom Ω_{obj}, v, and f_{obj}. Once a minimizer is known, we can retrieve the deblurred images $f(t_1, \cdot)$ and $f(t_2, \cdot)$ applying (3).

4. Discussion of the model

In this section we validate the performance of our variational model and discuss results obtained for different applications. Figures 2 and 6 demonstrate the model for two different test cases. In both we see the proper identification of the moving object and estimation of the motion velocity. We obtain an estimated velocity $v = (9.47, -0.007)$ of the plane in Fig. 6, compared to the true velocity $v = (10, 0)$. The joint approach for all three unknowns—the motion velocity v, the object intensity f_{obj} and the object domain Ω_{obj}—turns out to be crucial for a proper reconstruction of blurred video frames. This interdependence is demonstrated by the results in Fig. 7 where we compare our joint approach with a two step method which first tries to identify Ω_{obj} and v based on a motion competition algorithm [6], followed by the actual deblurring in a second step. Note that the proposed method can be regarded as a motion competition method if we skip the convolution with the convolution kernel h_v in the variational formulation. Figure 7 also shows the importance of the consistent motion blur model from (4) for a proper reconstruction in the vicinity of motion singularities. Finally, we have applied our model to a true motion sequence recorded with a hand held video camera. The sequence shows a toy car moving in front of a puzzle (background). We choose a textured object moving in front of a textured background to demonstrate the interplay between the deblurring steered by the fidelity functional $\mathcal{F}^\delta$ (see Eq. 7) and the reconstruction of sharp edges due to the total variation built into the prior $\mathcal{J}^\delta$. Results showing the overall procedure of our approach are also depicted in Fig.1. In Fig. 8 we render a zoom onto the moving object, which demonstrates the interplay of the deblurring and the edge reconstruction.

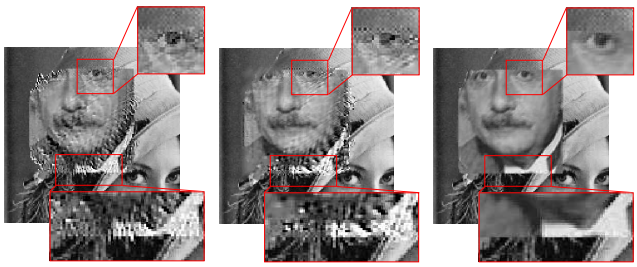

Figure 7. A comparison of our joint method with a non–joint method and with a method not taking into account the consistent motion blur model is shown. A restored frame with two zoom up areas is depicted for a straightforward scale variant motion deblurring, where the contour is extracted a priori based on pure motion competition (left), for the non–consistent motion blur model on the same a priori computed contour (middle), and for the fully joint method with the consistent model (right).

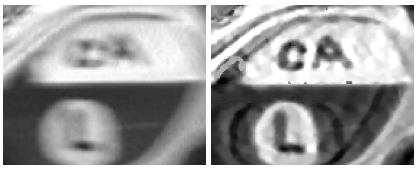

Figure 8. A blow up of the moving object from Fig. 1 is rendered for an original frame with motion blur (left) and for the restored intensity (right) computed by our model.

5. The minimization algorithm

To solve the minimization problem for the energy (5) we consider that the object domain Ω_{obj} is represented by the zero super level set $\{x \in \Omega : \phi(x) > 0\}$ of a level set function and follow the approach proposed by Chan and Vese [5]. The domain splitting into object and background in the different energy terms is encoded via a heaviside function $H(\phi)$ with $H(s) = 1$ for $s > 0$, and 0 elsewhere. Furthermore, the perimeter of the object domain can be rewritten as the total variation of $H(\phi)$, i. e. $|\partial\Omega_{\text{obj}}| = \int_\Omega |\nabla(H(\phi))|\,\mathrm{d}x$ [1]. As in [5] we consider a regularized heaviside function $H_\delta(x) := \frac{1}{2} + \frac{1}{\pi}\arctan\left(\frac{x}{\delta}\right)$ for a scale parameter $\delta > 0$. Let us emphasize that the desired guidance of the initial zero contour to the actual object boundary relies on the non-local support of this regularized heaviside function. Applying this approximation, we get a regularized integrand G_i^δ, representing the expected motion blur at time t_i:

$$\begin{aligned} G_i^\delta[\phi, v, f_{\text{obj}}, f_{\text{bg}}] = &((f_{\text{obj}} H_\delta(\phi)) * h_v)(\cdot - \tau_i v) \\ &+ (f_{\text{bg}}(1 - H_\delta(\phi) * h_v))(\cdot - \tau_i v)\,. \end{aligned} \tag{6}$$

Finally, we obtain an approximate global energy consisting of fidelity term $\mathcal{F}^\delta$ and prior $\mathcal{J}^\delta$

$$\begin{aligned} \mathcal{E}^\delta[\phi, v, f_{\text{obj}}] &= \mathcal{F}^\delta[\phi, v, f_{\text{obj}}] + \mathcal{J}^\delta[\phi, f_{\text{obj}}] \\ &:= \sum_{i=1,2} \int_\Omega \left(G_i^\delta[\phi, v, f_{\text{obj}}] - g_i\right)^2 \mathrm{d}x \\ &\quad + \int_\Omega \mu|\nabla f_{\text{obj}}| + \nu|\nabla H_\delta(\phi)|\,\mathrm{d}x\,. \end{aligned} \tag{7}$$

This expression depends on the motion vector $v \in \mathbb{R}^2$ and two scalar, unknown functions, namely the level set description ϕ of the object domain Ω_{obj} and the object intensity f_{obj}. Now, we take into account discrete intensities for a given video frame resolution of $n \times m$ pixels. We combine this with a finite difference approximation of the energy, and denote by Φ and F_{obj} the corresponding vectors of nodal values in $\mathbb{R}^{nm}$. In what follows, we will outline an energy relaxation method in this already spatially discrete setting based on an operator splitting with step size control and a regularized descent with respect to the level set description.

Initialization. At first, given an initial contour, we select and fix (in a very rough approximation step) F_{obj} as the intensity values of one of the images g_1 and g_2. Then, we relax the functional $\int_\Omega \sum_{i=1,2}(\tilde{G}_i^\delta - g_i)^2 + \nu|\nabla H_\delta(\phi)|\,\mathrm{d}x$, where $\tilde{G}_i^\delta$ is obtained from G_i^δ skipping the motion blur convolution. This initializing step can be regarded as a "motion competition approach" (as in the level set formulation of [6]), and we obtain an initial contour Φ^0 and an initial estimate v^0 for the motion velocity. Now, fixing Φ^0 and v^0, a standard deblurring based on (2) is performed on g_1 and g_2 to obtain an initial estimate for F_{obj}^0.

Gradient descent. We examined experimentally a significantly different *roughness* (difference of gradient directions) of the energy landscape associated with the unknowns Φ, v, and F_{obj}. Hence, an operator splitting strategy which separates these directions and incorporates different time steps for all of them turns out to be appropriate. In any subsequent descent step we pick up the newly computed quantities from the same iteration. As step size control we consider Armijo's rule, [15], separately evaluated for all three components. The descent in the level set description Φ of the object domain Ω_{obj} requires a special treatment.

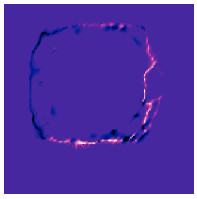

Figure 9. Color coded point-wise gradient $\mathrm{grad}_\Phi \mathcal{E}^\delta$ for one iteration from the application in Fig. 2.

A point-wise evaluation of shape derivatives (here given as variations with respect to the level set function) in the presence of fine scale fluctuation in the corresponding integrand of the shape functional (in our case object and background texture blurred solely in the direction of motion) is questionable (cf. Fig. 9, which shows non–smoothness and concentration of the gradient). Hence, we incorporate a regularized gradient descent in the level set function inspired by the Sobolev active contour approach [22]. It is based on a Gaussian filtering of the descent direction (presented in the appendix) with a filter $\mathcal{G}_\sigma$ of width $\sigma = 0.005$. Let us emphasize that the resulting regularized descent does not affect the energy landscape itself, but solely the descent path towards the set of minimizers.

Stopping criterion. As a stopping criterion we require the offsets in all three unknowns computed in the last time step and measured in the Euclidean norm to be bounded by a threshold parameter ϵ. In our implementation we have chosen $\epsilon = 0.01$.

A plot of the energy decay for the application in Fig. 2 is given in Fig. 10. Finally, let us summarize the algorithm in pseudo code notation:

EnergyDescent(g_1, g_2) {
 initialize $\Phi^0, v^0, F_{\text{obj}}^0$; $k = 1$;
 do {
  $\tau^\Phi = \text{AmijoStepSize}[\mathcal{E}^\delta, \Phi^k]$;
  $\Phi^{k+1} = \Phi^k - \tau^\Phi \mathrm{grad}_\Phi \mathcal{E}^\delta[\Phi^k, v^k, F_{\text{obj}}^k] * \mathcal{G}_\sigma$;
  $\tau^v = \text{AmijoStepSize}[\mathcal{E}^\delta, v^k]$;
  $v^{k+1} = v^k - \tau^v \mathrm{grad}_v \mathcal{E}^\delta[\Phi^{k+1}, v^k, F_{\text{obj}}^k]$;
  $\tau^F = \text{AmijoStepSize}[\mathcal{E}^\delta, F_{\text{obj}}^k]$;
  $F_{\text{obj}}^{k+1} = F_{\text{obj}}^k - \tau^F \mathrm{grad}_{F_{\text{obj}}} \mathcal{E}^\delta[\Phi^{k+1}, v^{k+1}, F_{\text{obj}}^k]$;
  $k = k + 1$;
 } while$(\|\Phi^{k+1} - \Phi^k\|, \|v^{k+1} - v^k\|, \|F_{\text{obj}}^{k+1} - F_{\text{obj}}^k\| \geq \epsilon)$
}

For the convenience of the reader, a comprehensive collection of variations of the different energy contributions

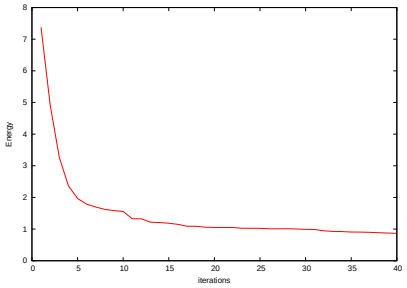

Figure 10. Plot of the energy decay in the descent algorithm.

comprised in the gradient vectors $\mathrm{grad}_\Phi \mathcal{E}^\delta$, $\mathrm{grad}_v \mathcal{E}^\delta$, and $\mathrm{grad}_{F_{\text{obj}}} \mathcal{E}^\delta$ is given in the Appendix.

6. Conclusions and outlook

In this work, we have presented a Mumford-Shah type variational formulation for joint motion estimation and deblurring from video, which includes a segmentation of the moving region. Following the tradition of jointly solving for inter-dependent unknowns, we have shown that this formulation outperforms individual and independent estimates. In particular, we present a consistent motion blur model at motion discontinuities and demonstrate it to be essential for proper deblurring.

Although the presented framework is generic, it was particularly addressed for single moving objects and static background. Handling multiple objects can be simply done by having multiple Ω_{obj} unknown regions in the general introduced formulation (cf. the approach by Chan and Vese [5] for multiple segments). More elegantly, and thereby also permitting dynamic background, we could consider formulations of the type $f(t,\cdot)\chi(\cdot) = f(t+\tau,\cdot)\chi(\cdot - \tau v)$. This constraint means that the function moves with the object, and eliminates the need for having independent f_{obj} and f_{bg} functions. Results using this functional are in progress.

7. Appendix

Here, we discuss the variation of the different energy contributions comprised in the gradient vectors $\mathrm{grad}_\Phi \mathcal{E}^\delta$, $\mathrm{grad}_v \mathcal{E}^\delta$, and $\mathrm{grad}_{F_{\text{obj}}} \mathcal{E}^\delta$ required to reproduce the gradient descent algorithm. In the case of F_{obj} and Φ, these gradients consist of the derivatives of the discrete energy with respect to the nodal values Φ_j and $(F_{\text{obj}})_j$, for $j = 1, \cdots, nm$. To shorten notation, we introduce the residual term

$$r_i(x) := 2\left[G_i^\delta[\phi, v, f_{\text{obj}}, f_{\text{bg}}](x) - g_i(x)\right].$$

First, we consider the derivative with respect to v_1 and v_2 ($v = (v_1, v_2)$), and obtain

$$\begin{aligned}\partial_{v_j}\mathcal{F}^\delta = \sum_{i=1}^{2}\int_\Omega \Big[&\big[\partial_j(f_{\text{obj}}H_\delta(\phi)) * (k_v - \tau_i h_v)\big](x-\tau_i v)\\ &-\big[\partial_j H_\delta(\phi) * (k_v - \tau_i h_v)\big](x-\tau_i v)\\ &\cdot f_{\text{bg}}(x)\Big] r_i(x)\,\mathrm{d}x,\end{aligned}$$

where $k_v(y) = -(y\cdot v)|v|^{-2}h_v(y)$. Here, we rewrite the fidelity term $\mathcal{F}^\delta$ in terms of (2), differentiate this and then convert it back to a spatial integral, instead of differentiating $f * h_v$ directly.

Let us remark that here we do not need to regularize the block filter function h_v to be able to calculate the variation of $\mathcal{F}^\delta$ with respect to v. This approach at the same time leads to significantly more stable results.

To deduce the derivatives with respect to the other unknowns, which represent discrete functions, we begin with the first variation of the fidelity functional $\mathcal{F}^\delta$ and the prior functional $\mathcal{J}^\delta$ in the direction of test functions and discretize afterwards:

$$\begin{aligned}\frac{d}{d\epsilon}\mathcal{F}^\delta[f_{\text{obj}}+\epsilon\vartheta]\Big|_{\epsilon=0} &= \sum_{i=1}^{2}\int_\Omega r_i(x)((\vartheta H_\delta(\phi)) * h_v)(x-\tau_i v)\,\mathrm{d}x\,,\\ \frac{d}{d\epsilon}\mathcal{F}^\delta[\phi+\epsilon\psi]\Big|_{\epsilon=0} &= \sum_{i=1}^{2}\int_\Omega r_i(x)\Big[(f_{\text{obj}}H'_\delta(\phi)\psi) * h_v)(x-\tau_i v)\\ &\qquad - f_{\text{bg}}(x)((H'_\delta(\phi)\psi) * h_v)(x-\tau_i v)\Big]\,\mathrm{d}x,\\ \frac{d}{d\epsilon}\mathcal{J}^\delta[f_{\text{obj}}+\epsilon\vartheta]\Big|_{\epsilon=0} &= -\mu\int_\Omega \mathrm{div}\left(\frac{\nabla f_{\text{obj}}}{|\nabla f_{\text{obj}}|}\right)\vartheta\,\mathrm{d}x\,,\\ \frac{d}{d\epsilon}\mathcal{J}^\delta[\phi+\epsilon\psi]\Big|_{\epsilon=0} &= -\nu\int_\Omega \mathrm{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right)H'_\delta(\phi)\psi\,\mathrm{d}x\,.\end{aligned}$$

Here, we have applied straight forward differentiation and integration by parts. To remove the convolution from the test function ϑ we use the integral transform

$$\int_\Omega f(x+a)(g*h)(x+b)\,\mathrm{d}x = \int_\Omega (f^{a,+} * h^{b,-})(y)g(y)\,\mathrm{d}y\,,$$

where $q^{b,\pm}(x) := q(\pm x + b)$. Now, choosing test functions concentrated at nodes and evaluated for the spatially discretized energy, we finally obtain

$$\begin{aligned}\frac{\partial}{\partial(F_{\text{obj}})_j}\mathcal{F}^\delta &= \sum_{i=1}^{2}(R_i * h_v^{-\tau_i v,-})(x_j)H_\delta(\Phi(x_j))\,,\\ \frac{\partial}{\partial\Phi_j}\mathcal{F}^\delta &= \sum_{i=1}^{2}(R_i * h_v^{(-\tau_i v),-})(x_j)(F_{\text{obj}}H'_\delta(\Phi))(x_j)\\ &\qquad -(R_i F_{\text{bg}} * h_v^{(-\tau_i v),-})(x_j)H'_\delta(\Phi(x_j))\,,\\ \frac{\partial}{\partial(F_{\text{obj}})_j}\mathcal{J}^\delta &= -\mu\,\mathrm{div}\left(\frac{\nabla F_{\text{obj}}(x_j)}{|\nabla F_{\text{obj}}(x_j)|}\right),\\ \frac{\partial}{\partial\Phi_j}\mathcal{J}^\delta &= -\nu H'_\delta(\Phi(x_j))\,\mathrm{div}\left(\frac{\nabla\Phi(x_j)}{|\nabla\Phi(x_j)|}\right),\end{aligned}$$

where $R_i(x) := 2\left[G_i^\delta[\Phi, v, F_{\text{obj}}, F_{\text{bg}}](x) - g_i(x)\right]$ denotes the spatially discrete blurring residual. Note that we use standard difference quotients to numerically evaluate the derivatives appearing above.

Acknowledgement

This work has been partially supported by NSF, NGA, ONR, DARPA, and ARO. Benjamin Berkels was supported by the DFG priority program 1114 on time sequence analysis and image processing.

References

[1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford University Press, 2000. 6

[2] L. Bar, N. Sochen, and N. Kiryati. Restoration of images with piecewise space-variant blur. In *Proceedings of the First International Conference on Scale Space Methods and Variational Methods in Computer Vision*, pages 533–544. Springer, 2007. 3

[3] B. Berkels, M. Burger, M. Droske, O. Nemitz, and M. Rumpf. Cartoon extraction based on anisotropic image classification. In *Vision, Modeling, and Visualization Proceedings*, 2006. 3

[4] V. Caselles, L. Garrido, and L. Igual. A contrast invariant approach to motion estimationa contrast invariant approach to motion estimation. In *Scale Space and PDE Methods in Computer Vision. 5th International Conference, Scale-Space 2005, Hofgeismar, Germany, April 7-9, 2005*, pages 242–253, 2005. 1

[5] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10, no. 2:266 – 277, 2001. 6, 7

[6] D. Cremers and S. Soatto. Motion competition: A variotional approach to piecewiese parametric motion segmentation. *The International Journal of Computer Vision*, 2004. 5, 6

[7] P. Favaro, M. Burger, and S. Soatto. Scene and motion reconstruction from defocused and motion-blurred images via anisotropic diffusion. In *Computer Vision - ECCV 2004, 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part I*, pages 257–269, 2004. 3

[8] P. Favaro and S. Soatto. A variational approach to scene reconstruction and image segmentation from motion-blur cues. In *2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04) - Volume 1*, pages 631–637, 2004. 3, 4

[9] L. He, A. Marquina, and S. Osher. Blind deconvolution using TV regularization and Bregman iteration. *International Journal of Imaging Systems and Technology*, 15(1):74–83, 2005. 3

[10] B. Horn and B. Schunk. Determining optical flow. *Artificial Intelligence*, 17:185–204, 1981. 1

[11] J. Jia. Single image motion deblurring using transparency. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2007. 3

[12] T. Kapur, L. Yezzi, and L. Zöllei. A variational framework for joint segmentation and registration. *IEEE CVPR - MMBIA*, pages 44–51, 2001. 3

[13] Y.-H. Kim, A. M. Martinez, and A. C. Kak. Robust motion estimation under varying illumination. *Image and Vision Computing*, 23(4):365–375, 2005. 1

[14] P. Kornprobst, R. Deriche, and G. Aubert. Image sequence analysis via partial differential equations. *Journal of Mathematical Imaging and Vision*, 11:5–26, 1999. 3

[15] P. Kosmol. *Optimierung und Approximation*. de Gruyter Lehrbuch, 1991. 6

[16] S. H. Lee, N. S. Moon, and C. W. Lee. Recovery of blurred video signals using iterative image restoration combined with motion estimation. In *Proceedings of the 1997 International Conference on Image Processing (ICIP '97) 3-Volume Set-Volume 1 - Volume 1*, pages 755–758, 1997. 3

[17] A. Levin. Blind motion deblurring using image statistics. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Advances in Neural Information Processing Systems 19*. MIT Press, Cambridge, MA, 2007. 3

[18] T. Nir, R. Kimmel, and A. Bruckstein. Variational approach for joint optic-flow computation and video restoration. Technical report, Department of Computer Science, Technion - Israel Institute of Technology, Technion City, Haifa 32000, Israel, 2005. 3

[19] A. Rav-Acha and S. Peleg. Restoration of multiple images with motion blur in differentdirections. In *Fifth IEEE Workshop on Applications of Computer Vision*, pages 22–28, 2000. 3

[20] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992. 5

[21] F. Sina, M. D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *IEEE Transactions on Image Processing*, 13(10):1327–1344, October 2004. 3

[22] G. Sundaramoorthi, J. D. Jackson, A. Yezzi, and A. Mennucci. Tracking with Sobolev active contours. In *Proceedings CVPR*, 2006. 6

[23] G. Unal, G. Slabaugch, A. Yezzi, and J. Tyan. Joint segmentation and non-rigid registration without shape priors. Scr-04-tr-7495, Georgia Institute of Technology, Atlanta, 2004. 3

[24] J. Weickert, A. Bruhn, and C. Schnörr. Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. *International Journal of Computer Vision*, 61(3):211–231, 2005. 1