A phase field method for joint denoising, edge detection and motion estimation

T. Preusser, M. Droske, C. Garbe, A. Telea, M. Rumpf

No Institute Given

Abstract. The estimation of optical flow fields from image sequences is incorporated in a Mumford–Shah approach for image denoising and edge detection. Possibly noisy image sequences are considered as input and a piecewise smooth image intensity, a piecewise smooth motion field, and a joint discontinuity set are obtained as minimizers of the functional. The method simultaneously detects image edges and motion field discontinuities in a rigorous and robust way and it is able to handle information on motion concentrated on edges. Inherent to it is a natural multi–scale approximation that is closely related to the phase field approximation for edge detection by Ambrosio and Tortorelli. We present an implementation for 2D image sequences with finite elements in space and time. This leads to three linear systems of equations, which have to be solved in a suitable iterative minimization procedure. Numerical results underline the robustness of the presented approach and different applications are shown.

1 Introduction

The task of motion estimation is a fundamental problem in computer vision. In low level image processing, the accurate computation of object motion in scenes is a long standing problem, which has been addressed extensively. In particular, global variational approaches initiated by the work of Horn and Schunck [18] are increasingly popular. Initial problems such as the smoothing over discontinuities or the high computational cost have been resolved successfully [25, 8, 10]. Motion also poses an important cue for object detection and recognition. While a number of techniques first estimate the motion field and segment objects later in a second phase [36], an approach of both computing motion as well as segmenting objects at the same time is much more appealing. First advances in this direction were investigated in [32, 27, 28, 11, 22, 30]. In particular, Kornprobst et al. [21, 3, 4] have considered piecewise smooth motion pattern on image sequences characterized by piecewise smooth objects. Their results are phrased rigorously on the space of functions of bounded variation and they propose suitable approximations for the numerical implementation. Already, in [21] a joint approach for the segmentation of moving objects in front of a still background and the computation of the motion velocities has been proposed. For a given intensity function on an image sequence a total variation (TV) type functional for the motion field - which allows for jumps in the optical flow velocity - is analysed in [4, 3]. Recently, Papenberg et al. [29] considered another TV regularization of the motion field and optical flow constraints involving higher order gradients.

The idea of combining different image processing tasks into a single model in order to cope with interdependencies has drawn attention in several different fields. In image registration, for instance, a joint discontinuity approach for simultaneous registration, segmentation and image restoration has been proposed by Droske & Ring [15] and extended in [5] to incorporate phase field approximations. Yezzi, Zöllei and Kapur [19] and Unal *et al.* [34] have combined segmentation and registration by applying geodesic active contours described by level sets in both images. Vemuri *et al.* have also used a level set technique to exploit a reference segmentation in an atlas [35]. We refer to [14] for further references.

Recently, Keeling and Ring [20] investigated the relation between optimization and optical flow extraction. A first approach which relates optical flow estimation to Mumford–Shah image segmentation was presented by Nesi [26]. Recently, Rathi *et al.* investigated active contours for joint segmentation and optical flow extraction [31]. Cremers and Soatto [13, 12] presented an approach for joint motion estimation and motion segmentation with one functional. Incorporating results from Bayesian inference, they derived an energy functional, which can be seen as an extension to the well known Mumford–Shah [23] approach. Their functional involves the length of boundaries separating regions of different motion as well as a "fidelity-term" for the optical-flow assumption. Brox et al. [9] present a Chan-Vese type model for piecewise smooth motion extraction. For given fixed image data the decomposition of image sequences into regions of homogeneous motion is encoded in a set of level set functions and the regularity of the motion fields in these distinct regions is controlled by a total variation functional. Our approach is in particular inspired by these investigations.

We combine denoising and edge detection with the estimation of motion. This results in an energy functional incorporating fidelity- and smoothness-terms for both the image sequence and the flow field. Our focus is in particular on information on motion concentrated on edges such as in the case of moving object with sharp egde contours but without shading and texture. To cope with this, we phrase the optical flow equations appropriately in regions apart from edges and on the edge set. Moreover, we incorporate an anisotropic enhancement of the flow along the edges of the image in the sense of Nagel and Enkelmann [24]. This effectively allows to spread information on the motion from the edge set onto the whole domain of a moving object. The model is implemented using the phase-field approximation in the spirit of Ambrosio's and Tortorelli's[2] approach for the original Mumford–Shah functional. The identification of edges is phrased in terms of a phase field function, no *a-priori* knowledge of objects is required, as opposed to formulations of explicit contours. *A particular focus is on optical flow constraints which are not only continuously distributed over shaded or textured regions, but might be concentrated on edges, e. g. in case of moving objects without texture and shading.* In contrast to a level set approach, the built-in multi-scale of the phase field model enables a natural cascadic energy relaxation approach and thus an efficient computation. Indeed, no initial guess for the edge set and the motion field will be required. We present here a truly $d + 1$ dimensional algorithm, considering time as an additional dimension to the d-dimensional image data. This fully demonstrates the conceptual advantages of the joint approach. Nevertheless, a transfer of the method for only

two consecutive time frames is possible but not investigated here. The characteristics of our approach are:

- The distinction of smooth motion fields and optical flow discontinuities is directly linked to edge detection, improving the reliability of the motion estimation.
- The denoising and segmentation task will profit from the explicit coupling of the sequence via the brightness constancy assumption.
- The phase field approximation is expected to converge to a limit problem for vanishing scale parameter, with a strict notion of edges and motion field discontinuities not involving any additional filtering parameter.
- The algorithm is based on an iteration. In each step a set of three relatively simple linear systems have to be solved for the image intensity, the edge description via the phase field, and the motion field, respectively. Only a small number of iterations is required.

This paper is organized as follows: In Section 2 Mumford–Shah type image denoising and edge detection is reviewed, in Section 3 we discuss a generalized optical flow equation and in Section 4 the minimization problem is presented. Section 5 shows how to approximate the segmentation in terms of a variational phase field model. Furthermore, we prove existence of solutions of this model and discuss the limit behaviour. Section 6 propounds the corresponding Euler-Lagrange equations, which are discretized applying the usual finite element procedure in Section 7. We conclude with the results in Section 8. Finally, in Appendix A we provide explicit formulas of all matrices and vectors appearing in the implementation to enable readers to reproduce the algorithm.

2 Recalling the Mumford–Shah functional

In their pioneering paper, Mumford and Shah [23] proposed the minimization of the following energy functional:

$$E_{MS}[u,S] = \lambda \int_{\Omega} (u-u_0)^2 \, \mathrm{d}\mathcal{L} + \frac{\mu}{2} \int_{\Omega \setminus S} |\nabla u|^2 \, \mathrm{d}\mathcal{L} + \nu \mathcal{H}^{d-1}(S) \,, \qquad (1)$$

where u_0 is the initial image defined on an image domain $\Omega \subset \mathbb{R}^d$ and λ, μ, ν are positive weights. Here, one asks for a piecewise smooth representation u of u_0 and an edge set S, such that u approximates u_0 in the least-squares sense, u ought to be smooth apart from the free discontinuity set S, and in addition S should be smooth and thus small with respect to the $(d-1)$-dimensional Hausdorff-measure $\mathcal{H}^{d-1}$. Mathematically, this problem has been treated in the space of functions of bounded variation BV, more precisely in the specific subset SBV [1]. In this paper, we will pick up a phase field approximation for the Mumford–Shah functional (1) proposed by Ambrosio and Tortorelli [2]. They describe the edge set S by a phase field ζ which is supposed to be small on S and close to 1 apart from edges, i. e., one asks for minimizers of the energy functional

$$E_{\epsilon}[u,\zeta] = \int_{\Omega} \lambda(u-u_0)^2 + \frac{\mu}{2}(\zeta^2 + k_{\epsilon}) |\nabla u|^2 + \nu\epsilon |\nabla\zeta|^2 + \frac{\nu}{4\epsilon}(1-\zeta)^2 \, \mathrm{d}\mathcal{L} \,, \quad (2)$$

where ϵ is a scale parameter and $k_\epsilon = o(\epsilon) \ll 1$ a small positive regularizing parameter, that mathematically ensures strict coercivity with respect to u. On edges the weight ζ^2 is expected to vanish. Hence, the second term measures smoothness of u but only apart from edges. The last two terms in the integral encode the approximation of the $d - 1$ dimensional area of the edge set and the strong preference for a phase field value $\zeta \approx 1$ apart from edges, respectively. For larger ϵ one obtains coarse, blurred representations of the edge sets and corresponding smoother images u. With decreasing ϵ we successively refine the representation of the edges and include more image details.

3 Generalized optical flow equation

In image sequences we observe different types of motion fields: locally smooth motion visible via variations of object shading and texture in time, or jumps in the motion velocity apparent at edges of objects moving in front of a background. We aim for an identification of corresponding piecewise smooth optical flow fields in piecewise smooth image sequences

$$u : [0, T] \times \Omega \mapsto \mathbb{R} \, ; \quad (t, x) \rightarrow u(t, x)$$

for a finite time interval $[0, T]$ and a spatial domain $\Omega \subset \mathbb{R}^d$ with $d = 1, 2, 3$. In what follows, we assume $\partial \Omega$ to be Lipschitz. The flow fields are allowed to jump on edges in the image sequence. Hence, the derivative Du splits into a singular and a regular part. The regular part is a classical gradient $\nabla_{(t,x)} u$ in space and time, whereas the singular part lives on the singularity set S - the set of edge surfaces in space–time. Time slices of S are the actual image edges S with respect to space–time. The singular part represents the jump of the image intensity on S, i. e., one observes that $D^s u = (u^+ - u^-) n_s$. Here, u^+ and u^- are the upper and lower intensity values on both sides of S, respectively. Now, we suppose that the image sequence u reflects an underlying motion with a piecewise smooth motion velocity v, which is allowed to jump only on S. Thus, S represents object boundaries moving in front of a background, which might as well be in motion. In strict mathematical terms, we suppose that $u, v \in SBV$ (the set of functions of bounded variation and vanishing Cantor part in the gradient) [16, 1]. In this general setting without any smoothness assumption on u and v, we have to ask for a generalized optical flow equation. In fact, apart from moving object edges we derive, as usually, from the brightness constancy assumption $u(t + s, x + s\,v) = const$ on motion trajectories $\{(t + s, x + s\,v) \,|\, s \in [0, T]\}$, that

$$\nabla_{(t,x)} u \cdot w = 0 \, , \tag{3}$$

where $w = (1, v)$ is the space–time motion velocity. On edges, the situation is more complex and in general requires prior knowledge. For instance, a white circular disk moving in front of a black background is visually identical to a black mask with a circular hole moving with the same speed on a white background. Hence, it is ambiguous on which $\pm$ side of the edge $w^\pm$ vanishes and on which side a non–trivial optical flow equation $n_s \cdot w^\pm = 0$ holds. We will not resolve this ambiguity via semantic assumptions. In what follows, we assume instead that locally only one object – in our example

either the circle or the mask – is moving on a stationary remaining background. Hence, we rule out that foreground and background are in motion. In other words, our background is that image part which is not moving, and the foregound correspondingly. Then, one of the two values of w on both sides of the edge vanishes by assumption and we can rewrite the optical–flow–constraint on the edge without identifying foreground or background by

$$n_s \cdot (w^+ + w^-) = 0. \tag{4}$$

This in particular includes the case of a sliding motion without any modification of the object overlap, where $n_s \cdot w^+ = n_s \cdot w^- = 0$.

4 Mumford–Shah approach to optical flow

Now, we ask for a simultaneous denoising, segmentation and flow extraction on image sequences. Hence, we will incorporate the motion field generating an image sequence into a variational method. Let us formulate a corresponding minimization problem in the spirit of the Mumford–Shah model:

Definition 1 (Mumford–Shah type optical flow approach). *Given a noisy initial image sequence $u_0 : D \mapsto \mathbb{R}$ on the space time domain $D = [0, T] \times \Omega$, we define the following energy*

$$E_{\mathrm{MSopt}}[u, w, S] = \int_D \frac{\lambda_u}{2} (u - u_0)^2 \, d\mathcal{L} + \int_{D \setminus S} \frac{\lambda_w}{2} \left(w \cdot \nabla_{(t,x)} u \right)^2 \, d\mathcal{L}$$

$$+ \int_{D \setminus S} \frac{\mu_u}{2} \left| \nabla_{(t,x)} u \right|^2 \, d\mathcal{L} + \int_{D \setminus S} \frac{\mu_w}{q} \left| \nabla_{(t,x)} w \right|^q \, d\mathcal{L} + \nu \mathcal{H}^d(S) \tag{5}$$

for a piecewise smooth image sequence u, and a piecewise smooth motion field $w = (1, v)$ with a joint jump set S. Furthermore, we require the optical flow constraint $n_s \cdot (w^+ + w^-) = 0$ on S from (4). Now, one asks for a minimizer (u, w, S) of the corresponding constraint minimization problem.

The first and second term of the energy are fidelity terms with respect to the image intensity and the regular part of the optical–flow–constraint, respectively. The third and fourth term encode the smoothness requirement of u and w. Finally, the last terms represents the area of the edge surfaces S. The fidelity weights λ_u, λ_w, the regularity weights μ_u, μ_w and the weight ν controlling the phase field are supposed to be positive and $q \geq 2$. Let us emphasize that, without any guidance from the local time–modulation of shading or texture on both sides of an edge, there is still a undecidable ambiguity with respect to foreground and background.

5 Phase field approximation

Similar to the original model for denoising and edge detection the above Mumford–Shah approach with its explicit dependence on the geometry of the edge set is difficult

to implement without any additional strong assumptions either on the image sequence or on the motion field. For a corresponding parametric approach we refer to the recent results by Cremers and Soatto [13, 12]. The level set approach recently presented by Brox et al. [9] does not explicitly encode motion concentrated on edges. We do not aim to impose any additional assumption on the image sequence u and the motion field v and ask for a suitable approximation of the above model. To gain more flexibility and, in addition, to incorporate a simple multi–scale into the model, we propose here a phase-field formulation (2) in the spirit of Ambrosio and Tortorelli [2]. Let us remark, that in [3] Aubert et al. already proposed to consider this type of phase field approximation for the regularization of the motion field. Let us introduce an auxiliary variable ζ – the phase field – describing the edge set S. Apart from S we aim for $\zeta \approx 1$ and on S the phase field should vanish. As in the original Ambrosio–Tortorelli model a scale parameter ϵ controls the thickness of the region with small phase field values. We consider the following energy functionals:

$$E_{\text{fid},u}^{\epsilon}[u] = \int_D \frac{\lambda_u}{2} (u - u_0)^2 \, \mathrm{d}\mathcal{L} \,, \tag{6}$$

$$E_{\text{fid},w}^{\epsilon}[u, w] = \int_D \frac{\lambda_w}{2} \left(w \cdot \nabla_{(t,x)} u \right)^2 \, \mathrm{d}\mathcal{L} \,, \tag{7}$$

$$E_{\text{reg},u}^{\epsilon}[u, \zeta] = \int_D \frac{\mu_u}{2} (\zeta^2 + k_\epsilon) \left| \nabla_{(t,x)} u \right|^2 \, \mathrm{d}\mathcal{L} \,, \tag{8}$$

$$E_{\text{phase}}^{\epsilon}[\zeta] = \int_D \left(\nu\epsilon \left| \nabla_{(t,x)} \zeta \right|^2 + \frac{\nu}{4\epsilon} (1 - \zeta)^2 \right) \, \mathrm{d}\mathcal{L} \tag{9}$$

These energy contributions control the approximation of the initial image u_0 (6) and the optical–flow–constraints (7), the regularity of u (11) , and the shape of the phase field ζ (9). Here, as in the original model $k_\epsilon = o(\epsilon) > 0$ is a "safety" coefficient, which is later on needed to establish existence of solutions of our approximate problem. Still missing is a regularity term for the motion field corresponding to the fourth energy term in the Mumford Shah model (5). If we would consider in a straightforward way the integral

$$\tilde{E}_{\text{reg},w}^{\epsilon}[w, \zeta] \int_D \frac{\mu_w}{q} (\zeta^2 + k_\epsilon) \left| \nabla_{(t,x)} w \right|^2 \, \mathrm{d}\mathcal{L} \tag{10}$$

the motion field will form approximate jumps on S but without any coupling of a concentrated motion constraint on S and the motion field in homogeneous regions on the image sequence. Fig. 1 clearly outlines this drawback in case of a circle with constant white image intensity inside moving on a textured background. As an alternative one might want to decouple the scales for image edges and motion edges introducing a second phase field with a much finer scale parameter $\tilde{\epsilon} \ll \epsilon$ for the representation of motion singularities. But this is not very handsome concerning a suitable discretization on digital images with limited pixel resolution and an parameter ϵ already in the range of the pixel size. Furthermore, following this approach we obtain in case of finite energy motion fields w bounded in $W^{1,2}$, which is not sufficient to ensure compactness of the

optical flow integrant in (7). Thus, to allow for piecewise smooth motion fields and to enable an extension of motion velocities first concentrated on edges via the variational approach, we consider

$$E_{\text{reg},w}^{\epsilon}[w,\zeta] = \int_D \frac{\mu_w}{q} \left| P_\delta[\zeta] \nabla_{(t,x)} w \right|^q \, d\mathcal{L}, \qquad (11)$$

Here, the following desired properties are encoded in the operator $P_\delta[\zeta]$:

- Close to the edges, were $\zeta \leq \theta^-$ for some θ^- with $0 < \theta^- < 1$, $P_\delta[\zeta]$ should behave like the original edge indicator ζ^2 proposed by Ambrosio and Tortorelli [2].
- Apart from the edges, were $\zeta \geq \theta^+$ for $\theta^- < \theta^+ < 1$, $P_\delta[\zeta]$ is expected to be the identity matrix, which enforces an isotropic smoothness modulus for the motion field w.
- In the spirit of the classical approach by Nagel and Enkelmann [24], $P_\delta[\zeta]$ will be an (approximate) projection onto level sets of the phase field function in the intermediate region. These level sets are surfaces approximately parallel to the edge set in space time. Thus, information on the optical flow is mediated along the edge set, without a coupling across edge surfaces.

An explicit definition for $P_\delta[\zeta]$ fulfilling these properties is the following:

$$P_\delta[\zeta] = \alpha(\zeta^2) \left(\mathbb{1} + k_\epsilon - \beta(\zeta^2) \frac{\nabla_{(t,x)}\zeta}{\left|\nabla_{(t,x)}\zeta\right|_\delta} \otimes \frac{\nabla_{(t,x)}\zeta}{\left|\nabla_{(t,x)}\zeta\right|_\delta} \right),$$

where $|z|_\delta = (|z|^2 + \delta^2)^{\frac{1}{2}}$ represents a regularized normal. Furthermore, $\alpha : \mathbb{R} \to \mathbb{R}_0^+$ and $\beta : \mathbb{R} \to \mathbb{R}_0^+$ are continuous blending functions, with $\alpha(s) = \max\left(0, \min\left(1, \frac{s}{\theta^-}\right)\right) + k_\epsilon$ and $\beta(s) = \max\left(0, \min\left(1, \frac{1-s}{1-\theta^+}\right)\right)$. Concerning algebraic notation, $\nabla_{(t,x)}w(t,x)$ is a $(d+1)^2$ matrix and thus $P_\delta[\zeta]\nabla_{(t,x)}w$ represents the matrix product. We consider the Frobenius norm of matrices, given by $|A| = \sqrt{\text{tr}(A^T A)}$. Suitable choices for the parameters are $\theta^+ = 0.8$ and $\theta^- = 0.0025$. For vanishing ϵ and a corresponding steepening of the slope of u, this operator basically leads to a separated diffusionon both sides of S in the relaxation of the energy.

Let us recall, that the energies $E_{\text{reg},u}^{\epsilon}$, $E_{\text{phase}}^{\epsilon}$, and the term $E_{\text{fid},u}^{\epsilon}$ are identical to those in the original Ambrosio–Tortorelli approach (see above). In addition, we ask for an optical flow field w according to the optical–flow–constraint encoded in $E_{\text{fid},w}^{\epsilon}$ (cf. Fig. 1 for a first test case). At the same time, this term implies a strong coupling of the image intensities along motion trajectories – which turns into a flow-aligned diffusion in the corresponding Euler–Lagrange equations – for the benefit of a more robust denoising and edge detection. Figure 2 shows an example where a completely destroyed time step in the image sequence is recovered by this enhanced diffusion along motion trajectories. Due to the regularity energy $E_{\text{reg},w}^{\epsilon}$ this motion field is isotropically smooth apart of the approximate jump set of u and the smoothness modulus is characterized by a successively stronger anisotropy along level sets of u while approaching the approximate jump set. The energy term $E_{\text{reg},w}^{\epsilon}$ (8) we consider for the regularization of the

motion field is very similar to the corresponding smoothness term in the classical approach by Nagel and Enkelmann [24], where tangential diffusion is steered by the local structure tensor. In the above multi–scale approach no additional pre-filtering of the image sequence in terms of a structure tensor is required.

The projection operator $P_\delta[\zeta]$ couples the smoothness of the motion field w to the image geometry, which in fact is very beneficial for the purpose of piecewise smooth motion extraction. The reverse coupling, which would try to align tangent spaces of level sets of u to the motion field, is not required and might even be misleading for our actual goal. The optical flow term in the fidelity energy $E^\epsilon_{\text{fid},w}$ already couples image sequence gradients to the motion field in a direct way. Hence, we don't ask for global minimizers of the sum of all energies but formulate the phase field approximation problem as follows:

Definition 2 (Solution of the phase field model). *For given noisy space time image $u_0 : D \mapsto \mathbb{R}$ and boundary data $v_\delta \in W^{1,q}(D, \mathbb{R}^d)$ for the velocity field, we denote a space time image $u \in W^{1,2}(D, \mathbb{R})$, a motion field $w = (1, v + v_\delta)$, with $v \in W^{1,q}_0(D, \mathbb{R}^d)$, and a phase field $\zeta \in W^{1,2}(D, \mathbb{R})$ a solution of the phase field model, if u and ζ minimizes the restricted energy*

$$E^\epsilon_{\text{fid},u}[u] + E^\epsilon_{\text{fid},w}[u, w] E^\epsilon_{\text{reg},u}[u, \zeta] + E^\epsilon_{\text{phase}}[\zeta], \tag{12}$$

for fixed w in $W^{1,2}(D, \mathbb{R}^{d+1})$, and the motion field w minimize the global energy

$$E^\epsilon_{\text{global}}[u, w, \zeta] = E^\epsilon_{\text{fid},u}[u] + E^\epsilon_{\text{fid},w}[u, w] + E^\epsilon_{\text{reg},u}[u, \zeta] + E^\epsilon_{\text{reg},w}[w, \zeta] + E^\epsilon_{\text{phase}}[\zeta] \tag{13}$$

for fixed u, $\zeta \in W^{1,2}(D, \mathbb{R})$.

Remark: The definition of u and ζ as the minimizer of a restricted functional is not only sound with respect to the applications. Indeed, a simultaneous relaxation of the global energy with respect to all unknowns is theoretically questionable. $E^\epsilon_{\text{reg},w}$ is not convex in ζ and we can not expect this energy contribution to be lower semi–continuous on a suitable set of admissible functions. With the above notion of solutions the direct method in the calculus of variations can be applied and in particular one observes compactness of the sequence of phase fields associated with a minimizing sequence of image sequences and motion fields (cf. the proof below).

Theorem 1 (Existence of solutions). *Suppose $d + 1 < q < \infty$, $\lambda_u, \lambda_w, \mu_u, \mu_w, \nu, \epsilon > 0$ and let $k_\epsilon > 0$. Then there exist a solution (u, w, ζ) of the phase field problem introduced in Definition 2.*

Proof. At first, we rewrite the phase field approach as an energy minimization problem, which later allows us to apply the direct method from the calculus of variations. For fixed w the energy functional

$$E_w[u, \zeta] := E^\epsilon_{\text{fid},u}[u] + E^\epsilon_{\text{fid},w}[u, w] + E^\epsilon_{\text{reg},u}[u, \zeta] + E^\epsilon_{\text{phase}}[\zeta]$$

is strictly convex and by the direct method we obtain a unique minimizer. Thus, let us denote by $(u[w], \zeta[w])$ this minimizer in $W^{1,2}(D, \mathbb{R}) \times W^{1,2}(D, \mathbb{R})$ of the quadratic energy functional $E_w[u, \zeta]$ for fixed $u \in W^{1,2}(D, \mathbb{R})$. The minimizing phase field is given as the weak solution of the corresponding Euler Lagrange equation

$$-\epsilon \Delta \zeta + \frac{1}{4\epsilon} \zeta = f[u, \zeta] := \frac{1}{4\epsilon} - \frac{\mu_u}{2\nu} \left| \nabla_{(t,x)} u \right|^2 \zeta. \tag{14}$$

Applying the weak maximum principle we observe that $\overline{\zeta} \equiv 1$ is a super solution and $\underline{\zeta} \equiv 0$ a sub solution. Thus, $\zeta[w]$ is uniformly bounded, i. e. $0 \le \zeta[w] \le 1$.

Given $(u[w], \zeta[w])$ we consider the global energy $E_{\text{global}}^{\epsilon}$ solely as a functional of the motion field $w = (1, v)$:

$$E[w] = E_{\text{global}}^{\epsilon}[u[w], w, \zeta[w]]$$

on the admissible set

$$\mathcal{A} := \{ w \mid w = (1, v + v_\delta), \ v \in W_0^{1,q}(D, \mathbb{R}^{d+1}) \}$$

and define $\underline{E} := \inf_{w \in \mathcal{A}} E[w]$. Testing the energy at $u \equiv 0$, $\zeta \equiv 0$, and $w = (1, v_\delta)$ we observe that $\underline{E} \le \frac{\lambda_u}{2} |u_0|_{L^2}^2 + \frac{\mu_w}{q} \left| \nabla_{(t,x)} v_\delta \right|_{L^q}^q < \infty$. Let us consider a minimizing sequence $\left(w^k \right)_{k=1\cdots,\infty}$ in $\mathcal{A}$ with $E[w^k] \to \underline{E}$ for $k \to \infty$. We set $u^k = u[w^k]$ and $\zeta^k = \zeta[w^k]$ and estimate the energy E_{global} as follows

$$E_{\text{global}}[u, w, \zeta] \ge \frac{\lambda_u}{4} \left(|u|_{L^2}^2 - 2 |u_0|_{L^2}^2 \right) + \frac{\mu_u k_\epsilon}{2} \left| \nabla_{(t,x)} u \right|_{L^2}^2 + \frac{\mu_w k_\epsilon}{q} \left| \nabla_{(t,x)} w \right|_{L^q}^q$$

$$+ \frac{\nu}{4\epsilon} \left(|\zeta|_{L^2}^2 - 2\mathcal{L}(D) \right) + \nu\epsilon \left| \nabla_{(t,x)} \zeta \right|_{L^2}^2 ,$$

where $\mathcal{L}(D)$ denotes the Lebesgue measure of D. From this, we deduce that $(u^k)_k$ and $(\zeta^k)_k$ are bounded in $W^{1,2}(D, \mathbb{R})$ and taking into account the boundary conditions that $(w^k)_k$ is bounded in $W^{1,q}(D, \mathbb{R})$. Hence, we can extract a weakly converging subsequence again denoted by $(u^k, w^k, \zeta^k)_k$ with weak limits (u, w, ζ). From the Sobolev embedding theorem and the assumption $q > d + 1$ we derive that w^k strongly converges in L^∞. Furthermore, the corresponding sequence $(\zeta^k)_k$ of phase field functions $\zeta^k := \zeta[w^k]$ are weak solutions of $-\epsilon \Delta \zeta^k + \frac{1}{4\epsilon} \zeta^k = f^k$ (cf. (14)). From the bounds on ζ^k in L^∞ and on u^k in $W^{1,2}$ we obtain that $f^k = f[u^k, \zeta^k]$ is uniformly bounded in L^1. This observation allows to apply a compensated compactness result to verify that $\nabla_{(t,x)} \zeta^k$ converges to $\nabla_{(t,x)} \zeta$ a.e., which is proven for the equation $-\Delta \zeta = f$ on the space $W_0^{1,2}$ in [33, Chapter I, Thm 3.4], but can easily be generalized for equations of the type $-\Delta \zeta + \zeta = f$ on $W^{1,2}$. The matrix valued function $P_\delta[\cdot]$ is continuous and bounded. Hence, we obtain that $P_\delta[\zeta^k] \to P_\delta[\zeta]$ a.e. for $k \to \infty$. For later use, we define the constants $C_u = \sup_{k=1,\cdots,\infty} \left| \nabla_{(t,x)} u^k \right|_{L^2}$ and $C_w = \sup_{k=1,\cdots,\infty} \max \left\{ \left| w^k \right|_{L^\infty}, \left| \nabla_{(t,x)} w^k \right|_{L^q} \right\}$.

Next, we verify that $u = u[w]$ and $\zeta = \zeta[w]$. Indeed, taking into account the lower semi–continuity of E_w and the modulus of continuity with respect to w we can estimate

$$
\begin{aligned}
E_w[u, \zeta] &\leq \liminf_{k \to \infty} E_{w^k}[u^k, \zeta^k] \\
&\leq \liminf_{k \to \infty} E_{w^k}[\tilde{u}, \tilde{\zeta}] \\
&\leq E_w[\tilde{u}, \tilde{\zeta}] + \liminf_{k \to \infty} \left(\left| w^k \cdot \nabla_{(t,x)} \tilde{u} \right|_{L^2}^2 - \left| w \cdot \nabla_{(t,x)} \tilde{u} \right|_{L^2}^2 \right) \\
&\leq E_w[\tilde{u}, \tilde{\zeta}] + 2C_w \left| \nabla_{(t,x)} \tilde{u} \right|_{L^2}^2 \liminf_{k \to \infty} \left| w - w^k \right|_{L^\infty}
\end{aligned}
$$

for any $\tilde{u}, \tilde{\zeta} \in W^{1,2}(D, \mathbb{R})$. From the L^∞ convergence of w^k to w, we immediately obtain that $E_w[u, \zeta] \leq E_w[\tilde{u}, \tilde{\zeta}]$. Thus, by definition $u = u[w]$ and $\zeta = \zeta[w]$. Based on these preliminaries, we are able to prove weak lower semi–continuity of the energy. For this we assume without loss of generality that

$$
E[w^k] \leq \underline{E} + \rho, \quad \left| P_\delta[\zeta^k] - P_\delta[\zeta] \right|_{L^\infty} \leq \rho, \quad \left| w^k - w \right|_{L^\infty} \leq \rho
$$

for a fixed and small constant $\rho > 0$. Applying Mazur's lemma we obtain a sequence of convex combinations

$$
\left(\sum_{i=1,\cdots,k} \lambda_i^k (u^i, w^i, \zeta^i) \right)_k, \quad \text{with} \sum_{i=1,\cdots,k} \lambda_i^k = 1, \quad \lambda_i^k \geq 0,
$$

converging strongly to (u, w, ζ) in $W^{1,2}(D, \mathbb{R}) \times W^{1,2}(D, \mathbb{R}^{d+1}) \times W^{1,2}(D, \mathbb{R})$. Finally, taking into account convexity properties of the integrands, Fatou's lemma and the modulus of continuity of $E_{\text{fid},u}$, $E_{\text{fid},w}$, and $E_{\text{reg},w}$ with respect to w and $P_\delta[\zeta]$,

respectively, we estimate (using Einstein's summation convention)

$$
\begin{aligned}
E[w] = E_{\text{global}}[u[w], w, \zeta[w]] &= E_{\text{global}}[u, w, \zeta] \\
&= \int_D \frac{\lambda_u}{2} \left(\liminf_{k\to\infty} \lambda_i^k u^i - u_0 \right)^2 + \frac{\lambda_w}{2} \left| \liminf_{k\to\infty} w \cdot \left(\lambda_i^k \nabla_{(t,x)} u^i \right) \right|^q \, \mathrm{d}\mathcal{L} \\
&\quad + \int_D \frac{\mu_u}{2} (\zeta^2 + k_\epsilon) \left| \liminf_{k\to\infty} \lambda_i^k \nabla_{(t,x)} u^i \right|^2 + \frac{\mu_w}{q} \left| \liminf_{k\to\infty} P_\delta[\zeta] \lambda_i^k \nabla_{(t,x)} w^i \right|^q \, \mathrm{d}\mathcal{L} \\
&\quad + \int_D \left(\nu\epsilon \left| \liminf_{k\to\infty} \lambda_i^k \nabla \zeta^i \right|^2 + \frac{\nu}{4\epsilon} (1 - \liminf_{k\to\infty} \lambda_i^k \zeta^i)^2 \right) \mathrm{d}\mathcal{L} \\
&\leq \int_D \frac{\lambda_u}{2} \liminf_{k\to\infty} \lambda_i^k \left(\left(u^i - u_0 \right)^2 + \frac{\lambda_w}{2} \left| w \cdot \left(\nabla_{(t,x)} u^i \right) \right|^q \right) \mathrm{d}\mathcal{L} \\
&\quad + \int_D \liminf_{k\to\infty} \lambda_i^k \left(\frac{\mu_u}{2} (\zeta^2 + k_\epsilon) \left| \nabla_{(t,x)} u^i \right|^2 + \frac{\mu_w}{q} \left| P_\delta[\zeta] \nabla_{(t,x)} w^i \right|^q \right) \mathrm{d}\mathcal{L} \\
&\quad + \int_D \liminf_{k\to\infty} \lambda_i^k \left(\nu\epsilon \left| \nabla_{(t,x)} \zeta^i \right|^2 + \frac{\nu}{4\epsilon} (1 - \zeta^i)^2 \right) \mathrm{d}\mathcal{L} \\
&\leq \liminf_{k\to\infty} \lambda_i^k E[u^i, w^i] + \frac{\lambda_w}{2} \sup_{i=1,\cdots,\infty} \left(\left| w^i \cdot \nabla_{(t,x)} u^i \right|_{L^2}^2 - \left| w \cdot \nabla_{(t,x)} u^i \right|_{L^2}^2 \right) \\
&\quad + \frac{\mu_w}{q} \sup_{i=1,\cdots,\infty} \left(\left| P_\delta[\zeta^i] \nabla_{(t,x)} w^i \right|_{L^q}^q - \left| P_\delta[\zeta] \nabla_{(t,x)} w^i \right|_{L^q}^q \right) \\
&\leq \underline{E} + \lambda_w C_w C_u^2 \sup_{i=1,\cdots,\infty} \left| w^i - w \right|_{L^\infty} + \mu_w C_w^q \sup_{i=1,\cdots,\infty} \left(\left| P_\delta[\zeta^i] - P_\delta[\zeta] \right|_{L^\infty} \right) \\
&\leq \underline{E} + \rho + \lambda_w C_w C_u^2 \rho + \mu_w C_w^q \rho,
\end{aligned}
$$

This estimate holds for any $\rho \geq 0$. Thus, we obtain $E[w] \leq \underline{E}$, which implies that w is a minimizer of the energy E and hence (u, w, ζ) a solution of our phase field problem.

Remark: The above problem formulation is not only sound with respect to the actual modeling, but it will allow a simple relaxation approach (see below). Indeed, on all tested data sets we obtain convergence in very few iterations $(5 - 10)$.

Applying formal asymptotics, one observes that the phase field approach proposed here indeed converges to the above Mumford–Shah model. For small ϵ we expect a steeping of the gradient u on a strip of thickness ϵ around the edge set. The phase field ζ will approximate 1 apart from a decreasing neighbourhood of the edge surface. For $\epsilon \to 0$ we expect to observe convergence of $E_{\text{reg},u}$ and $E_{\text{reg},w}$ to $\int_{D\setminus S} \frac{\mu_u}{2} \left| \nabla_{(t,x)} u \right|^2 + \frac{\mu_w}{q} \left| \nabla_{(t,x)} u \right|^q \mathrm{d}\mathcal{L}$ and of $E_{\text{phase}}^\epsilon$ to $\mathcal{H}^d(S)$. Under these assumptions on the qualitative behavior $\int_D (w \cdot \nabla_{(t,x)} u)^2 \, \mathrm{d}\mathcal{L}$ converges to the second term of E_{MSopt}, whereas on the edge surface one observes a concentration of energy on the jump set scaling like $O(\epsilon^{-1})$. Thus, we observe that in the limit we reproduce our optical–flow–constraint $n_S \cdot (w^+ + w^-) = 0$ from the sharp interface Mumford–Shah approach. A rigorous validation of this limit behaviour in terms of Γ convergence is still open. For results on Γ convergence for the optical flow problem in the context of total variation type models we refer to [4, 21].

6 Variations of the energy and an algorithm

In what follows, we will consider the Euler–Lagrange equations of the above energies. Thus, we need to compute the variations of the energy contributions with respect to the involved unknowns u, w, ζ. The variation of an energy E in direction ζ with respect to a parameter function z will be denoted by $\langle \delta_z E, \zeta \rangle$. For the ease of implementation we confine here to the case $q = 2$. Using straightforward differentiation for sufficiently smooth u, w, ζ and initial data u_0 we obtain

$$\langle \delta_u E_{\text{fid},u}^\epsilon[u], \vartheta \rangle = \int_D \lambda_u (u - u_0) \, \vartheta \, \mathrm{d}\mathcal{L},$$

$$\langle \delta_u E_{\text{fid},w}^\epsilon[u, w], \vartheta \rangle = \int_D \lambda_w (\nabla_{(t,x)} u \cdot w)(\nabla_{(t,x)} \vartheta \cdot w) \, \mathrm{d}\mathcal{L},$$

$$\langle \delta_w E_{\text{fid}}^\epsilon[u, w], \psi \rangle = \int_D \lambda_w (\nabla_{(t,x)} u \cdot w)(\nabla_{(t,x)} u \cdot \psi) \, \mathrm{d}\mathcal{L},$$

$$\langle \delta_u E_{\text{reg},u}^\epsilon[u, \zeta], \vartheta \rangle = \int_D \mu_u (\zeta^2 + k_\epsilon) \nabla_{(t,x)} u \cdot \nabla_{(t,x)} \vartheta \, \mathrm{d}\mathcal{L}, \tag{15}$$

$$\langle \delta_\zeta E_{\text{reg},u}^\epsilon[u, \zeta], \xi \rangle = \int_D \mu_u \zeta \left| \nabla_{(t,x)} u \right|^2 \xi \, \mathrm{d}\mathcal{L},$$

$$\langle \delta_w E_{\text{reg},w}^\epsilon[w, \zeta], \psi \rangle = \int_D \mu_w P_\delta[\zeta] \nabla_{(t,x)} w : \nabla_{(t,x)} \psi \, \mathrm{d}\mathcal{L},$$

$$\langle \delta_\zeta E_{\text{phase}}^\epsilon[\zeta], \xi \rangle = \int_D 2\nu\epsilon \nabla_{(t,x)} \zeta \cdot \nabla_{(t,x)} \xi \, \mathrm{d}\mathcal{L} + \int_D \frac{\nu}{2\epsilon} (\zeta - 1) \xi \, \mathrm{d}\mathcal{L},$$

for scalar test functions ξ, ϑ and velocities type test functions ψ with the structure $\psi = (0, \pi)$. Here, we use the notation $A : B := \mathrm{tr}(B^T A)$. Now, summing up the different terms as in (12) and integrating by parts we end up with the following system of PDEs

$$-\mathrm{div}_{(t,x)}\left(\frac{\mu_u}{\lambda_u}(\zeta^2 + k_\epsilon)\nabla_{(t,x)}u + \frac{\lambda_w}{\lambda_u}w(\nabla_{(t,x)}u \cdot w)\right) + u = u_0 \tag{16}$$

$$-\epsilon \Delta_{(t,x)}\zeta + \left(\frac{1}{4\epsilon} + \frac{\mu_u}{2\nu}\left|\nabla_{(t,x)}u\right|^2\right)\zeta = \frac{1}{4\epsilon} \tag{17}$$

$$-\frac{\mu_w}{\lambda_w}\mathrm{div}_{(t,x)}\left(P_\delta[\zeta]\nabla_{(t,x)}v\right) + (\nabla_{(t,x)}u \cdot v)\nabla_x u = 0 \tag{18}$$

as the Euler–Lagrange equations characterizing the necessary conditions for a solution (u, w, ζ) of the above stated phase field approach. Let us emphasize, that the full Euler–Lagrange equations, characterizing a global minimizer of the energy would in addition involve variations of $E_{\text{reg},w}$ with respect to ζ.

The last equation (18) is degenerate in directions tangential to edges of u, because the optical flow constraint acts only in directions of the image gradient. To remedy this degeneracy we consider a gradient descent of (18)

$$\partial_s v - \frac{\mu_w}{\lambda_w}\mathrm{div}_{(t,x)}\left(P_\delta[\zeta]\nabla_{(t,x)}v\right) + (\nabla_{(t,x)}u \cdot w)\nabla_x u = 0. \tag{19}$$

Consequently the matrices resulting from a discretization are well conditioned and the corersponding systems can be solved easily (see Section 7).

Inspired by Ambrosio and Tortorelli, we propose the following iterative algorithm for the solution of the phase field problem with $q = 2$:

Step 0. Initialize $u = u_0$, $\zeta \equiv 1$, and $w \equiv (1,0)$.
Step 1. Solve (16) for fixed w, ζ.
Step 2. Solve (17) for fixed u, w.
Step 3. Compute one step of the gradient descent (19) for fixed u, ζ.
Step 4. Return to Step 1.

Steps 1 and 2 of the algorithm consist of a consecutive solution of linear partial differential equations. We consider in Step 3 a gradient descent in particular because for Neumann boundary conditions (which we actually consider in the application) the Euler Lagrange equation in w is not guaranteed to be coercive in $W^{1,q}$. Indeed, the optical flow terms $w \cdot \nabla_{(t,x)} u$ represents a pointwise rank–1 condition and a priori it is not known that "sufficiently many" of these conditions, in the sense of the Lebesgue measure, are assembled in the image while integrating this term. Let us note that we use a time step control for Step 3, i.e. the gradient descent. Alternatively we might iterate first Step 1 and 2 until convergence and then in an other iteration we would consider the identification of the motion field w. Even though, this second variation seems to be closer to our definition of solutions of the phase field problem, the above algorithm converges to the same solution in the applications we have considered. Our algorithm can be seen as a diagonal scheme, where the iteration of Steps 1/2 and the gradient descent iteration in Step 3 are intertwined.

7 Finite Element Discretization

We proceed similarly to the Finite Element method proposed by Bourdin and Chambolle [6, 7] for the phase field approximation of the Mumford–Shah functional. To solve the above system of PDEs we suppose $[0, T] \times \Omega$ to be overlaid by a regular hexahedral grid. In the following, the spatial and temporal grid width are denoted by h and τ respectively. Hence, image frames are at a distance of τ and pixels of each frame are placed on a regular mesh with grid size h.

On this hexahedral grid we consider the space of piecewise trilinear continuous functions $\mathcal{V}$ and ask for discrete functions $U, Z \in \mathcal{V}$ and $V \in \mathcal{V}^2$, such that discrete and weak counterparts of the Euler Lagrange equations (16), (17) and (18) are fulfilled. This leads to the solution of systems of linear equations for the vectors of the nodal values of the unknowns U, Z, V. We refer to the appendix for a detailed description of the matrices and the resulting systems of equations. A careful implementation is required to ensure an efficient method. For a time-space volume of K time steps and images of $N \times M$ pixels, the finite element matrices for U and Z have $N\,M\,K\,C$ entries, where $C = 27$ is the number of nonzero entries per row, equal to the number of couplings of a node. The finite element matrix for V has four times more elements, as V is a two-dimensional vector. Using an efficient custom-designed compressed row sparse matrix storage, we can treat datasets of up to $K = 10$ frames of $N = 500, M = 320$ pixels in less than 1GB memory. The linear systems of equations are solved applying a classical conjugate gradient method. For the pedestrian sequence (Fig. 5), one such iteration

takes 47 seconds on a Pentium IV PC at 1.8 GHz running Linux. The complete method converges after 10-15 such iterations. To treat large video sequences, we typically consider a window of $K = 6$ frames, to avoid boundary effects, and then shift this window successively in time.

In Figure 7 we have depicted the progression of the various components of the energy $E_{\text{global}}^{\epsilon}$ for the taxi sequence shown in Figure 6. The rapid decay of the global energy in the first steps of the algorithm is clearly visible. While the image fidelity $E_{\text{fid}}^{\epsilon}$ and its regularity $E_{\text{reg},u}^{\epsilon}$ decay, the other parts of the energy increase. Obviously this is the case, because we are starting with constant initial values $\zeta = 0$ and $w = (1, 0)$.

8   Results and Discussion

We present here several results of the proposed method for two dimensional image sequences. In the considered examples, the parameter setting $\epsilon = h/4$, $\mu_u = h^{-2}$, $\mu_w = \lambda_u = 1$, $\lambda_w = 10^5 h^{-2}$ and $k_\epsilon = \epsilon$, $\delta = \epsilon$ has proven to give good results. We first consider a simple example of a white disk moving with constant speed $v = (1, 1)$ on a vaguely textured, low-contrast, dark background (Fig. 1). Let us first consider the top row in Fig. 1 which corresponds to the energy formulation without the projection component. A limited amount of smoothing results from the regularization energy $E_{\text{reg},u}^{\epsilon}$ (Fig. 1(a)), which is desirable to ensure robustness in the resulting optical flow term $\nabla_{(t,x)} u \cdot w$ and removes noisy artifacts in real-world videos, e.g. Fig. 5 and Fig. 8. The phase field clearly captures the moving object's contour. The optical flow is depicted in Fig. 1(c) by color coding the vector directions as shown by the lower-right color wheel. Clearly, the method is able to extract the uniform motion of the disc's boundary, which has a high image contrast. The optical flow information, available only on the motion edges (black in Fig. 1(b)), is propagated only to a limited extent into the information-less area inside the moving disk. Indeed, we notice that the model with the standard regularity term for w (10) is not able to diffuse the optical flow information, concentrated on the motion edges, in order to completely and uniformly fill in the moving circle.

In the bottom row of Fig. 1, the same example is shown, this time run with the energy formulation including the projection term. We now clearly see a perfect reconstruction of the optical flow (Fig. 1 c bottom row) also inside the non-textured moving disc.

In the next example we revisit this simple image sequence of the moving circle. This time we have added noise to the sequence. At the same time we have completely destroyed the information of one frame of the sequence (Fig. 2). In Figure 2 we show the results for frames 3, and $9 - 11$ where frame 10 has been completely destroyed. From the images we see that the phase-field detects the missing circle in the destroyed frame as a temporal edge surface in the sequence. Indeed the ζ drops down to zero in the temporal vicinity of the destroyed frame. This is still visible in the previous and the next timestep, shown in the second and third row. But it does not hamper the restoration of the correct optical flow field shown in the fourth column. This is due to the anisotropic smoothing of information from the surrounding frames into the destroyed frame. For this example we have chosen $\epsilon = 0.4h$.
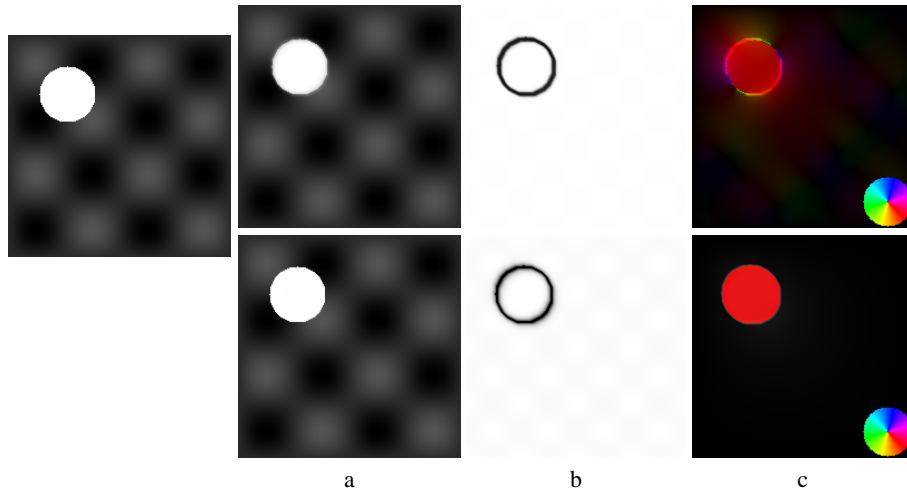
Fig. 1. One frame of the test sequence (left) and corresponding smoothed images (a), phase field (c) and optical flow (color coded) (c). Top row: Energy formulation without projection. Bottom row: energy formulation with projection.

Another synthetic example is shown in Fig. 3. This example is from the publicly available dataset collection at [17]. Here, a textured sphere spins on a textured background (Fig. 3(a)). Again, the method is able to clearly segment the moving object from the background, even though the object doesn't change position. We used a phase field parameter $\epsilon = 0.15h$. The extracted optical flow clearly shows the spinning motion (Fig. 3(d)) and the discontinuous motion field.

A first example on real video data is shown in Figure 8. The video shows a tennis player whose body moves to the right while the hand goes down as he strikes the ball. This motion is well captured in the flow field (Fig. 8 c).

Furthermore, we consider a complex, higher resolution video sequence, taken under outdoor conditions by a monochrome video camera. The sequence shows a group of walking pedestrians (Fig. 5 (top)). The human silhouettes are well extracted and captured by the phase field (Fig. 5(middle)). We do not visualize a vector plot of the optical flow, as it is hard to interpret visually at the video sequence resolution of 640 by 480 pixels. However, the color-coded optical flow plot (Fig. 5(bottom)) shows how the method is able to extract the moving limbs of the pedestrians. The overall red and blue color corresponds to the walking directions of the pedestrians. The estimated motion is smooth inside the areas of the individual pedestrians and not smeared across the motion boundaries. In addition, the algorithm nicely segments the different moving persons. The cluttered background poses no big problem to the segmentation, nor are the edges of occluding and overlapping pedestrians, who are moving at almost the same speed.

Finally, let us note on a limitation of the approach we consider here. Let us consider the well-known Hamburg taxi video sequence. The sequence is available from `csd.uwo.ca` under the directory `pub/vision`. Figure 6 shows the taxi sequence processed both with the classical AT energy component (top row) and with our pro-
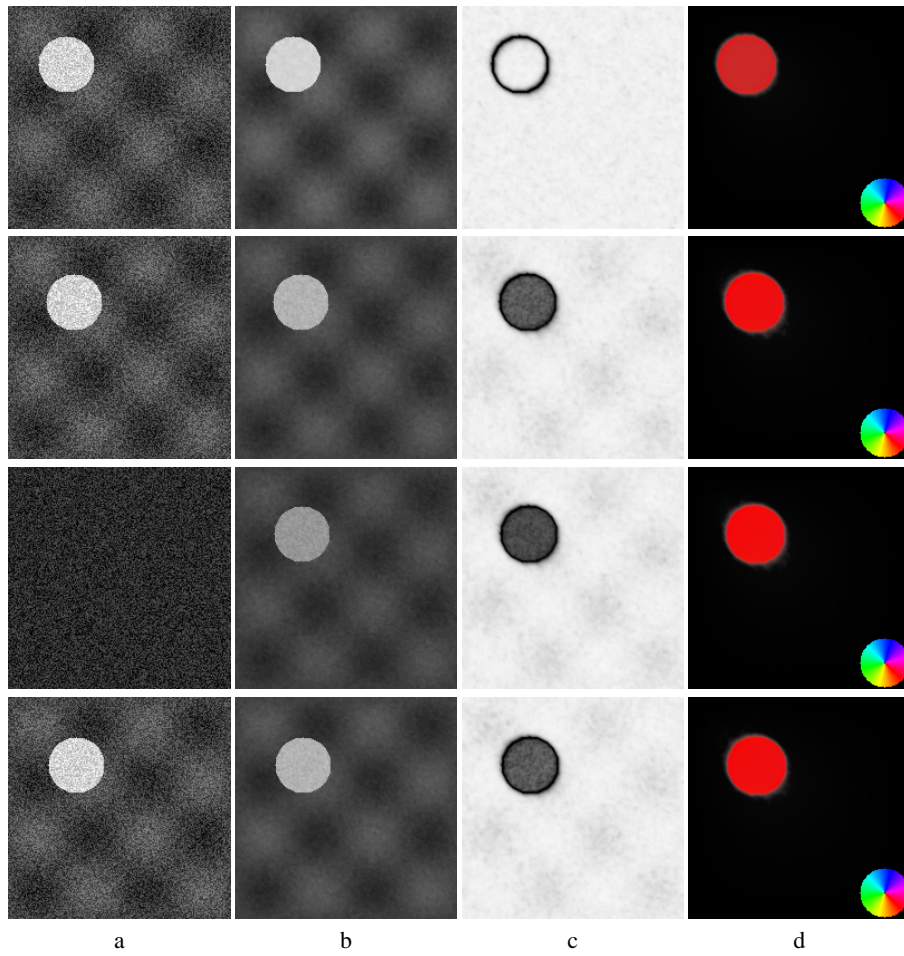
Fig. 2. Noisy test sequence: From top to bottom frames 3 and $9-11$ are shown. (a) original image sequence, (b) smoothed images, (c) phase field, (d) estimated motion (color coded)
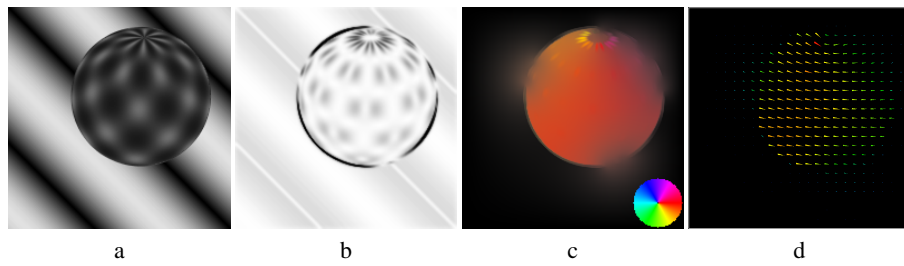

Fig. 3. Rotating sphere: smoothed image (a), phase field (b), optical flow (color coded) (c), optical flow (vector plot, color coded magnitude) (d)
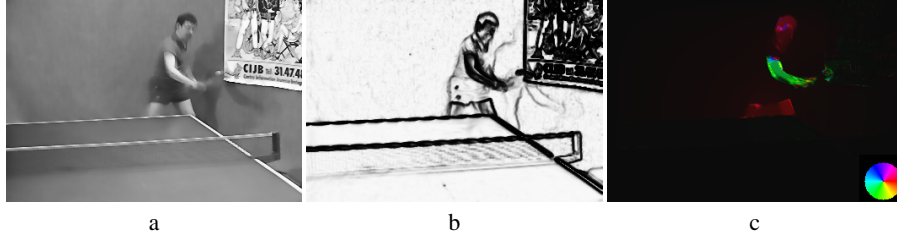
Fig. 4. Table tennis sequence: smoothed image (a), phase field (b) and optical flow (c)

jection operator (bottom row). The progression of the various energy contributions is shown in Figure 7. Here we start with $u = 0$, i.e. a black image, and a zero velocity field $v = 0$. In this sequence, cars of different image contrasts are moving. When the projection operator P_δ in our model is used (bottom row), only the central, high-contrast moving car is captured. When the operator is not used (top row), motion edges corresponding to low-contrast image edges also determine the phase field, hence the other oppositely moving cars in the bottom part of the image are captured as well and the corresponding optical flow is extracted. For all the cars in this example, the motion field is determined largely by the low-contrast shading and not only by high-contrast image edges, as it was the case in the synthetic example in Figure 1.

A   Algorithmic building blocks

In this appendix we would like to focus on the discrete version of the Euler–Lagrange equations resulting from (12). Let us denote by $\{\Psi_i\}_{i=1,...,N}$ the usual nodal basis of $\mathcal{V}$ (cf. Section 7). The corresponding basis of the vector valued discrete functions $\Psi \in \mathcal{V}^2$ is given by $\{\Psi_i e_1\}_i \cup \{\Psi_i e_2\}_i$ where $e_{1,2}$ are the standard basis vectors of $\mathbb{R}^2$: $e_1 = (1,0)$, $e_2 = (0,1)$. For any discrete function $Q \in \mathcal{V}$ we denote by $\overline{Q}$ the corresponding nodal vector. For discrete vector valued functions we order the coefficients such that the e_1 coefficients are followed by the e_2 coefficients. Hence, the systems of discrete equations to be solved in the above algorithm are given in matrix vector notion as follows. We ask for solution vectors $\overline{U}, \overline{Z} \in \mathbb{R}^N$ and $\overline{V} \in \mathbb{R}^{2N}$, such that denoting $\overline{W} = (1, \overline{V})$ we have

$$(\mathbf{L}_u[W, \zeta] + \mathbf{M})\overline{U} = R_u, \tag{20}$$

$$(\mathbf{L}_\zeta + \mathbf{M}_\zeta[U])\overline{Z} = R_\zeta, \tag{21}$$

$$(\mathbf{L}_w[Z] + \mathbf{M}_w[U])\overline{V} = R_w. \tag{22}$$

These systems contains the matrices $\mathbf{L}_u[W, Z], \mathbf{L}_\zeta, \mathbf{M}, \mathbf{M}_\zeta[U] \in \mathbb{R}^{N \times N}$, $\mathbf{L}_w[Z], \mathbf{M}_w[U] \in \mathbb{R}^{2N \times 2N}$, $R_u, R_\zeta \in \mathbb{R}^n$ and finally $R_w \in \mathbb{R}^{2N}$, which can easily be
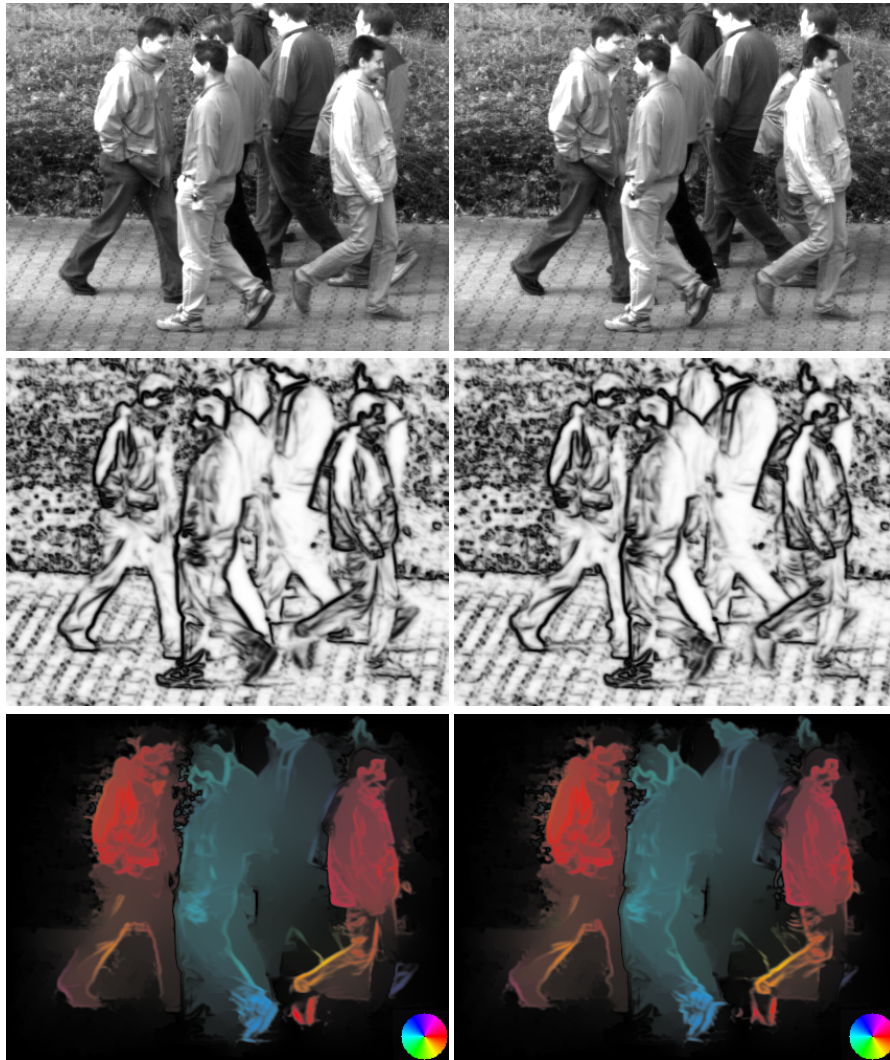
Fig. 5. Pedestrian video: frames from original sequence (top); phase field (middle); optical flow, color coded (bottom)
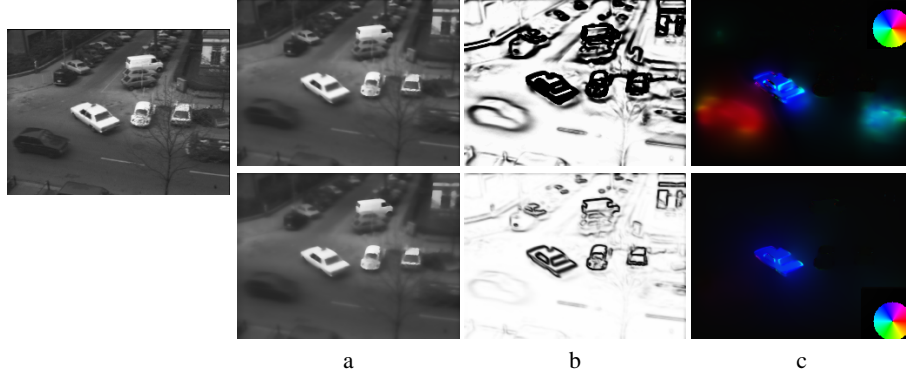
Fig. 6. The taxi sequence. Original image (left). Flow extraction without the projection operator (top row) and with projection (bottom row). Smoothed image (a), phase field (b) and optical flow, color coded (c)

derived from the variations of the energy (13). We have

$$(\mathbf{L}_u[W,Z])_{ij} = \int_D \frac{\mu_u}{\lambda_u}(Z^2 + k_\epsilon)\nabla_{(t,x)}\Psi_i \cdot \nabla_{(t,x)}\Psi_j \; + \frac{\lambda_w}{\lambda_u}(\nabla_{(t,x)}\Psi_i \cdot W)(\nabla_{(t,x)}\Psi_j \cdot W)\,\mathrm{d}\mathcal{L},$$

$$\mathbf{M}_{ij} = \int_D \Psi_i \Psi_j \,\mathrm{d}\mathcal{L},$$

$$R_u = \mathbf{M}\overline{\mathcal{I}_h u_0},$$

as well as

$$(\mathbf{L}_\zeta)_{ij} = \epsilon \int_D \nabla_{(t,x)}\Psi_i \cdot \nabla_{(t,x)}\Psi_j \,\mathrm{d}\mathcal{L},$$

$$(\mathbf{M}_\zeta[U])_{ij} = \int_D \left(\frac{\mu_u}{2\nu} \left| \nabla_{(t,x)}U \right|^2 + \frac{1}{4\epsilon} \right) \Psi_i \Psi_j \,\mathrm{d}\mathcal{L},$$

$$(R_\zeta)_i = \frac{1}{4\epsilon} \int_D \Psi_i \,\mathrm{d}\mathcal{L},$$

and

$$(\mathbf{L}_w[Z])_{ikjl} = \int_D \mu_w P_\delta[Z]\nabla_{(t,x)}\Psi_i \cdot \nabla_{(t,x)}\Psi_j \delta_{kl} \,\mathrm{d}\mathcal{L},$$

$$(\mathbf{M}_w[U])_{ikjl} = \int_D \lambda_w \partial_{x_k}U \partial_{x_l}U \, \Psi_i \Psi_j \,\mathrm{d}\mathcal{L},$$

$$(R_w)_{ik} = - \int_D \lambda_w \partial_t U \partial_{x_k}U \Psi_i \,\mathrm{d}\mathcal{L}.$$

(a) image fidelity $E_{\mathrm{fid},u}^{\epsilon}$

(b) flow fidelity $E_{\mathrm{fid},w}^{\epsilon}$

(c) image regularity $E_{\mathrm{reg},u}^{\epsilon}$

(d) flow regularity $E_{\mathrm{reg},w}^{\epsilon}$

(e) phase field energy $E_{\mathrm{phase}}^{\epsilon}$

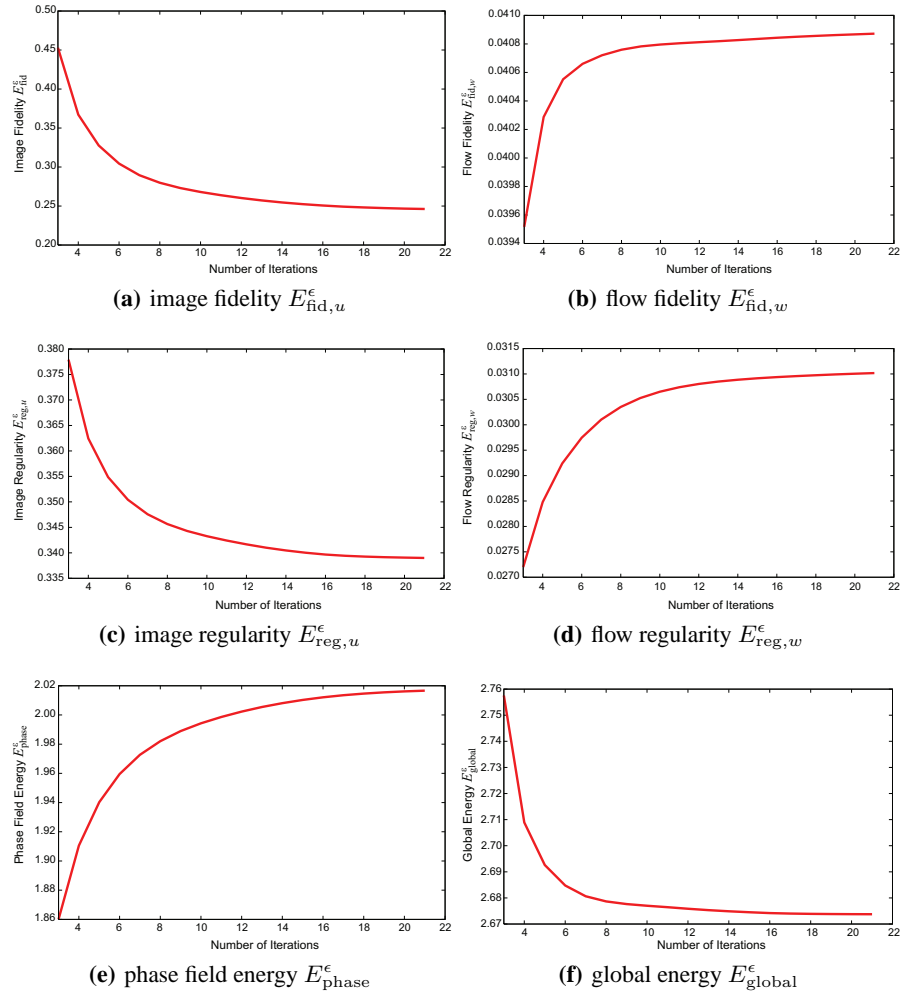**(f)** global energy $E_{\mathrm{global}}^{\epsilon}$

Fig. 7. For the example shown in Figure 6 (bottom row) we show the progression of the various energy contributions during the solution iteration. The decay of the global energy is shown in the lower right plot (f).

Here, δ_{kl} is the usual Kronecker symbol, which is 1 if $k = l$ and otherwise 0. Let us remark that the integrands are piecewise polynomials of degree ≤ 2. We use a suitable quadrature rule on the heaxhedra, which ensures exact integration.

References

1. L. AMBROSIO, N. FUSCO, AND D. PALLARA, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, 2000.
2. L. AMBROSIO AND V. M. TORTORELLI, *On the approximation of free discontinuity problems*, Boll. Un. Mat. Ital. B, 6 (1992), pp. 105–123.
3. G. AUBERT, R. DERICHE, AND P. KORNPROBST, *Computing optical flow via variational techniques*, SIAM J. Appl. Math., 60 (1999), pp. 156–182.
4. G. AUBERT AND P. KORNPROBST, *A mathematical study of the relaxed optical flow problem in the space $bv(\omega)$*, Siam J. Math. Anal., 30 (1999), pp. 1282–1308.
5. AUTHORS, *Mumford-shah based registration*, Computing and Visualization in Science manuscript, (2005). submitted.
6. B. BOURDIN, *Image segmentation with a Finite Element method*, ESIAM: Mathematical Modelling and Numerical Analysis, 33 (1999), pp. 229–244.
7. B. BOURDIN AND A. CHAMBOLLE, *Implementation of an adaptive Finite-Element approximation of the Mumford-Shah functional*, Numer. Math., 85 (2000), pp. 609–646.
8. T. BROX, A. BRUHN, AND J. WEICKERT, *Variational motion segmentation with level sets.*, in Computer Vision ? ECCV 2006., A. P. H. Bischof, A. Leonardis, ed., vol. Vol. 3951 of Lecture Notes in Computer Science, Springer, 2005.
9. ——, *Variational motion segmentation with level sets*, in COMPUTER VISION - ECCV 2006, vol. 1 of Lecture notes in computer science, 2006, pp. 471–483.
10. A. BRUHN AND J. WEICKERT, *A confidence measure for variational optic flow methods*, Tech. Rep. 106, Department of Mathematics, Saarland University, 2004. In: R. Klette, R. Kozera, L. Noakes, J. Weickert (Eds.): Geometric Properties from Incomplete Data, 283-297, Springer, Dordrecht, 2006.
11. V. CASELLES AND B. COLL, *Snakes in movement*, SIAM Journal on Numerical Analysis, 33 (1996), pp. 2445–2456.
12. D. CREMERS, T. KOHLBERGER, AND C. SCHNÖRR, *Nonlinear shape statistics in mumford-shah based segmentation*, in 7th European Conference on Computer Vision, Copenhagen, Springer LNCS, A. H. et al., ed., vol. 2351, 2002, pp. 93–108.
13. D. CREMERS AND S. SOATTO, *Motion competition: A variotional approach to piecewiese parametric motion segmentation*, The International Journal of Computer Vision, (2004).
14. C. A. DAVATZIKOS, R. N. BRYAN, AND J. L. PRINCE, *Image registration based on boundary mapping*, IEEE Trans. Medical Imaging, 15 (1996), pp. 112–115.
15. M. DROSKE AND W. RING, *A Mumford-Shah level-set approach for geometric image registration*, SIAM Appl. Math., (2005). to appear.
16. L. C. EVANS AND R. F. GARIEPY, *Measure Theory and Fine Properties of Functions*, CRC Press, 1992.
17. C. V. R. GROUP, *Optical flow datasets*, Univ. of Otago, New Zealand, `www.cs.otago.ac.nz/research/vision`, (2005).
18. B. HORN AND B. SCHUNK, *Determining optical flow*, Artificial Intelligence, 17 (1981), pp. 185–204.
19. T. KAPUR, L. YEZZI, AND L. ZÖLLEI, *A variational framework for joint segmentation and registration*, IEEE CVPR - MMBIA, (2001), pp. 44–51.

20. S. L. KEELING AND W. RING, *Medical image registration and interpolation by optical flow with maximal rigidity*, Journal of Mathematical Imaging and Vision, 23 (1) (2005), pp. 47–65. to appear.

21. P. KORNPROBST, R. DERICHE, AND G. AUBERT, *Image sequence analysis via partial differential equations*, Journal of Mathematical Imaging and Vision, 11 (1999), pp. 5–26.

22. E. MEMIN AND P. PEREZ, *A multigrid approach for hierarchical motion estimation*, in ICCV, 1998, pp. 933–938.

23. D. MUMFORD AND J. SHAH, *Optimal approximation by piecewise smooth functions and associated variational problems*, Communications on Pure Applied Mathematics, 42 (1989), pp. 577–685.

24. H.-H. NAGEL AND W. ENKELMANN, *An investigation of smoothness constraints for the estimation of dispalcement vector fields from image sequences*, IEEE Transaction on Pattern Analysis and Machine Intelligence, 8 (1986), pp. 565–593.

25. H. H. NAGEL AND W. ENKELMANN, *An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences*, IEEE Trans. Pattern Anal. Mach. Intell., 8 (1986), pp. 565–593.

26. P. NESI, *Variational approach to optical flow estimation managing discontinuities*, Image and Vision Computing, 11 (7) (1993), pp. 419–439.

27. J.-M. ODOBEZ AND P. BOUTHEMY, *Robust multiresolution estimation of parametric motion models*, Journal of Visual Communication and Image Representation, 6 (1995), pp. 348–365.

28. ———, *Direct incremental model-based image motion segmentation for video analysis*, Signal Processing, 66 (1998), pp. 143–155.

29. N. PAPENBERG, A. BRUHN, T. BROX, S. DIDAS, AND J. WEICKERT, *Highly accurate optic flow computation with theoretically justified warping*, International Journal of Computer Vision, 67 (2006), pp. 141–158.

30. N. PARAGIOS AND R. DERICHE, *Geodesic active contours and level sets for the detection and tracking of moving objects*, IEEE Transaction on Pattern Analysis and Machine Intelligence, 22 (2000), pp. 266–280.

31. Y. RATHI, N. VASWANI, A. TANNENBAUM, AND A. YEZZI, *Particle filtering for geometric active contours with application to tracking moving and deforming objects*, in CVPR, 2005, pp. 2–9.

32. C. SCHNÖRR, *Segmentation of visual motion by minimizing convex non-quadratic functionals*, in 12th International Conference on Pattern Recognition, Jerusalem, Israel, October 1994.

33. M. STRUWE, *Variational Methods: Applications to Nonlinear Partial Differential Equations and Hamiltonian Systems*, Springer Verlag, 2000.

34. G. UNAL, G. SLABAUGCH, A. YEZZI, AND J. TYAN, *Joint segmentation and non-rigid registration without shape priors*, scr-04-tr-7495, Georgie Institute of Technology, Atlanta, 2004.

35. B. VEMURI, J. YE, Y. CHEN, AND C. LEONARD, *Image registration via level-set motion: Applications to atlas-based segmentation*, Medical Image Analysis, 7 (2003), pp. 1–20.

36. J. Y. A. WANG AND E. H. ADELSON, *Representating moving images with layers*, IEEE Transaction on Image Processing, 3 (1994), pp. 625–638.

B Variant I

$$E_{\text{reg},w}^{\epsilon}[u, w] = \int_D \frac{\mu_w}{q} (\zeta^2 + k(\epsilon)) \left|\nabla_{(t,x)} w\right|^q \, d\mathcal{L},$$

$$\langle \delta_w E_{\text{reg},w}^{\epsilon}[u, w], \psi \rangle = \int_D \mu_w (\zeta^2 + k(\epsilon)) \nabla_{(t,x)} w : \nabla \psi, \tag{23}$$

$$\tag{24}$$

for velocities $\psi = (0, \pi)$.

Phase field approach. *For given noisy image sequence u_0, find a regularized image sequence u, a motion field $w = (1, v)$, and a phase field ζ from corresponding sets of admissible functions, such that for fixed w, u and ζ minimize $E^{\epsilon}[u, w, \zeta]$.*

Euler–Lagrange system:

$$-\text{div}\left(\frac{\mu_u}{\lambda_u}(\zeta^2 + k_\epsilon)\nabla_{(t,x)}u + \frac{\lambda_w}{\lambda_u}w(\nabla_{(t,x)}u \cdot w)\right) + u = u_0 \tag{25}$$

$$-\epsilon\Delta_{(t,x)}\zeta + \left(\frac{1}{4\epsilon} + \frac{\mu_u}{2\nu}\left|\nabla_{(t,x)}u\right|^2 + \frac{\mu_w}{2\nu}\left|\nabla_{(t,x)}w\right|^2\right)\zeta = \frac{1}{4\epsilon} \tag{26}$$

$$-\text{div}\left(\mu_w(\zeta^2 + k(\epsilon))\nabla_{(t,x)}w\right) + \lambda_w(\nabla_x u \cdot v)\nabla_x u = -\lambda_w \partial_t u \nabla_x u \tag{27}$$

as the Euler–Lagrange equations characterizing the necessary conditions for a solution (u, w, ζ) of the above stated phase field approach. Let us emphasize, that the full Euler–Lagrange equations, characterizing a global minimizer of the energy in addition would involve variations of $E_{\text{reg},w}$ with respect to u.