Variational Level Set Segmentation in Riemannian Sobolev Spaces

Maximilian Baust[1]
campar.in.tum.de/Main/MaximilianBaust
Darko Zikic[2]
research.microsoft.com/en-us/people/darko
Nassir Navab[1]
campar.in.tum.de/Main/NassirNavab

[1] Computer Aided Medical Procedures & Augmented Reality, Technische Universität München, Munich, Germany

[2] Machine Learning and Perception, Microsoft Research, Cambridge, GB

Abstract

Gradient flows in the Sobolev space H^1 have been shown to enjoy favorable regularity properties. We propose a generalization of prior approaches for Sobolev active contour segmentation by changing the notion of distance in the Sobolev space, which is achieved through treatment of the function and its derivative in Riemannian manifolds. The resulting generalized Riemannian Sobolev space provides the flexibility of choosing an appropriate metric, which can be used to design efficient gradient flows. We select this metric based on the rationale of preconditioning resulting in a significant improvement of convergence and overall runtime in case of variational level set segmentation.

1 Introduction

The variational level set method [18] is still one of the most widely used methods in computer vision – especially for image segmentation [1]. This popularity might seem surprising, because variational level set segmentation is known to be non-convex, e.g., [3]. All the more, because since the seminal work of Chan et al. [6] a lot of research has been carried out in order to develop efficient methods for solving convex models for image segmentation, cf. [2, 4, 8].

The non-convexity of the variational level set approach is caused by the usage of continuous but non-convex approximations of the Heaviside and Dirac distribution for defining area and boundary integrals. This non-convexity is, however, not always a bane, because variational level set formulations for localized active contours models [9] or image segmentation in the presence of intensity inhomogeneities [11] make extensive usage of smeared-out Heaviside and Dirac distributions. As a consequence, it is still of interest to develop efficient methods for the non-convex variational level set method for image segmentation, which is the goal of this paper. Thereby, we will consider so-called Sobolev gradient flows, which have recently been shown to be superior to classical L^2-based gradient flows [7, 17]. Inspired by [19], we extend these approaches by changing the notion of distance in H^1. The proposed generalization consists of endowing H^1 with a different inner product based on Riemannian metrics, instead of the Euclidean one, which is used in the standard approaches [17]. As a

[1] This fact is proven by 579 hits since 2012 obtained by a Google scholar search for the query "'variational level set' + 'segmentation'" performed on 9th of May 2014.

(a) (b) (c)

Figure 1: **The proposed generalization** (a) results in efficient Riemannian Sobolev flows, which provide accurate results (b), however with significantly improved convergence and overall runtime (c). Every 5th iteration is marked with a $+$.

consequence of this choice, the minimizing gradient flow in the Riemannian Sobolev space exhibits a significantly improved convergence, compared to gradient flow in H^1. This advantage in convergence translates directly to a notable improvement of the overall runtime, cf. Fig. 1 and Fig. 2.

2 Level Set Segmentation in Riemannian Sobolev Spaces

We start in Sec. 2.1 by describing gradient flows in a general Hilbert spaces. In Sec. 2.2 we discuss possible choices for the Hilbert space and the resulting gradient descent flows. After that, we generalize previous considerations by introducing Riemannian Sobolev spaces in Sec. 2.3. Finally, we use this generalization to increase the convergence rate of the resulting flow in Sec. 2.4.

2.1 Gradient Flows in Hilbert Spaces

In this work we will consider the segmentation problem in a variational manner, which means that the level set function $\phi : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ is (a local) minimizer of the problem

$$\min_{\phi \in \mathcal{H}} \mathcal{E}(\phi), \tag{1}$$

where $\mathcal{E}$ is the energy characterizing the optimal configuration of ϕ. $\mathcal{H} = \mathcal{H}(\Omega)$ is a suitable Hilbert space endowed with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ which induces a norm on $\mathcal{H}$ by

$$\|\phi\|_{\mathcal{H}} = \langle \psi, \psi \rangle_{\mathcal{H}}^{\frac{1}{2}}. \tag{2}$$

A gradient flow in $\mathcal{H}$ results in

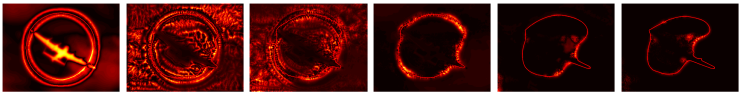
$$\partial_t \phi = -\nabla_{\mathcal{H}} \mathcal{E}(\phi), \tag{3}$$

where $\nabla_{\mathcal{H}} \mathcal{E}(\phi) \in \mathcal{H}$ represents the first variation of $\mathcal{E}$ in $\mathcal{H}$ as follows:

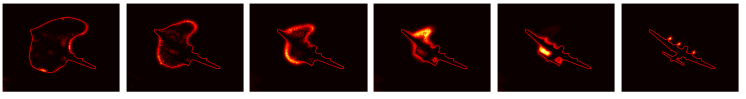
$$\frac{d}{dh} \mathcal{E}(\phi + h \cdot \psi)\Big|_{h=0} = \langle \nabla_{\mathcal{H}} \mathcal{E}(\phi), \psi \rangle_{\mathcal{H}}. \tag{4}$$

Eq. (4) reveals that the choice of $\mathcal{H}$ plays a crucial role for the definition of the actual gradient. It is important to note that the choice of $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ determines $\mathcal{H}$, because we have that

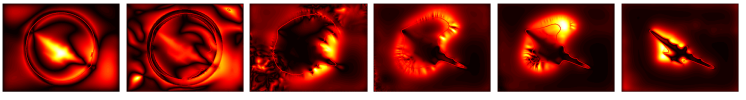
$$\mathcal{H} = \{\phi : \|\phi\|_{\mathcal{H}} < \infty\}. \tag{5}$$



(a) SF iter. 1 (b) SF iter. 4 (c) SF iter. 5 (d) SF iter. 7 (e) SF iter. 20 (f) SF iter. 40



(g) SF iter. 60 (h) SF iter. 95 (i) SF iter. 105 (j) SF iter. 109 (k) SF iter. 112 (l) SF iter. 129



(m) RSF iter. 1 (n) RSF iter. 2 (o) RSF iter. 6 (p) RSF iter. 9 (q) RSF iter. 11 (r) RSF iter. 15

Figure 2: **Gradient flows** for the classical Sobolev space approach (SF) and the proposed Riemannian Sobolev space flow (RSF). We visualize the norms of the updates $\tau\nabla_{\mathcal{E}}H^1$ and $\tau\nabla_{\mathcal{E}}H_M$ during iterations of the experiment from Fig. 1 (colorbar range is $[0,1]$). The improved convergence of the proposed approach results from the structure of the updates, with more equally distributed magnitudes, compared to the classical approach.

2.2 Gradient Flow in the Sobolev Space H^1

For many years $\mathcal{H} = L^2$ has been the only choice considered for $\mathcal{H}$. The gradient flow in L^2 is exposed to the risk of getting stuck in undesired local minima. Moreover, the numerical treatment of L^2 gradient requires regularization methods such as re-initialization [16] or signed distance regularization [10] for maintaining numerical stability.

Due to the drawbacks associated to L^2 flows Sundaramoorthi et al. [17] suggested to employ the Sobolev space H^1 which is based on the inner product

$$\langle \phi, \psi \rangle_{H^1} = \langle \phi, \psi \rangle_{L^2} + \alpha \langle \nabla\phi, \nabla\psi \rangle_{L^2}, \tag{6}$$

where $\alpha > 0$ and

$$\langle \phi, \psi \rangle_{L^2} = \int_\Omega \phi(x)\psi(x)\, dx. \tag{7}$$

The choice $\mathcal{H} = H^1$ is justified by the fact that H^1 contains smoother functions than L^2 [12] - namely those that have also derivatives with bounded L^2-norm - and thus an H^1 gradient flow is less sensitive to undesired local minima.

In order to implement the H^1 gradient flow, it is helpful represent $\nabla_{H^1}\mathcal{E}$ in terms of the L^2 gradient $\nabla_{L^2}\mathcal{E}$. As (4) is valid for any choice of $\mathcal{H}$ we obtain

$$\frac{d}{dh}\mathcal{E}(\phi + h\cdot\psi)\bigg|_{h=0} = \langle \nabla_{L^2}\mathcal{E}(\phi), \psi \rangle_{L^2} = \langle \nabla_{H^1}\mathcal{E}(\phi), \psi \rangle_{H^1}. \tag{8}$$

Using the definition of $\langle \cdot, \cdot \rangle_{H^1}$ and assuming Neumann boundary conditions we can integrate by parts:

$$\langle \nabla_{H^1}\mathcal{E}(\phi), \psi \rangle_{H^1} = \langle (I - \alpha\Delta)\nabla_{H^1}\mathcal{E}(\phi), \psi \rangle_{L^2} = \langle \nabla_{L^2}\mathcal{E}(\phi), \psi \rangle_{L^2}. \tag{9}$$

Finally, applying the fundamental lemma of the calculus of variations yields

$$\nabla_{H^1}\mathcal{E}(\phi) = \mathcal{S}_\alpha^{-1}\nabla_{L^2}\mathcal{E}(\phi), \tag{10}$$

where

$$\mathcal{S}_\alpha = I - \alpha\Delta \tag{11}$$

is called Sobolev operator [15]. The Sobolev operator allows us now to identify the H^1 gradient flow as

$$\partial_t\phi = -\mathcal{S}_\alpha^{-1}\nabla_{L^2}\mathcal{E}(\phi). \tag{12}$$

Thereby $\mathcal{S}_\alpha^{-1}$ can be considered as a smoothing operator which ensures that ϕ stays in H^1 during the evolution.

2.3 Gradient Flow in the Riemannian Sobolev Space H_M

In Sec. 2.2 we have seen that the smoothing operator S_α^{-1} is an important part of the H^1 gradient flow. In this subsection we will show that there are more possible smoothing operators than just S_α^{-1} which still yield an H^1 gradient flow. Consider the inner product

$$\begin{aligned}\langle\phi,\psi\rangle_M &= \langle\phi,\psi\rangle_{M_0} + \langle\nabla\phi,\nabla\psi\rangle_{M_1} &(13)\\ &= \langle\phi,M_0\psi\rangle_{L^2} + \langle\nabla\phi,M_1\nabla\psi\rangle_{L^2}, &(14)\end{aligned}$$

where $M_0 \in \mathbb{R}$ and $M_1 \in \mathbb{R}^{d\times d}$ are two Riemannian metric tensors, and for $M_0 = 1$ and $M_1 = \alpha$ we obtain the H^1 scalar product in (6). As M_0 and M_1 are Riemannian, which means that they are symmetric and positive definite, there exist two constants $0 < c, C < \infty$ such that $\|\cdot\|_{H^1}$ and $\|\cdot\|_M$ are equivalent norms:

$$c\cdot\|\phi\|_{H^1} \leq \|\phi\|_M = \langle\phi,\phi\rangle_M^{\frac{1}{2}} \leq C\cdot\|\phi\|_{H^1}. \tag{15}$$

This means that the Riemannian Sobolev space $H_M = \{\phi : \|\phi\|_M < \infty\}$ contains the same functions as H^1, but the notion of distance is different.

Now we want to study the H_M gradient flow and therefore we represent $\nabla_M\mathcal{E}$ via $\nabla_{L^2}\mathcal{E}$. Repeating the considerations of Sec. 2.2 we obtain

$$\langle\nabla_M\mathcal{E}(\phi),\psi\rangle_M = \langle\nabla_{L^2}\mathcal{E}(\phi),\psi\rangle_{L^2} \tag{16}$$

and integration by parts yields finally

$$\langle(M_0 - \operatorname{div}(M_1\nabla))\nabla_M\mathcal{E}(\phi),\psi\rangle_{L^2} = \langle\nabla_{L^2}\mathcal{E}(\phi),\psi\rangle_{L^2}. \tag{17}$$

Thus $\nabla_M\mathcal{E}(\phi) = (M_0 - \operatorname{div}(M_1\nabla))^{-1}\nabla_{L^2}\mathcal{E}(\phi)$ and the H_M gradient flow reads

$$\partial_t\phi = -\mathcal{S}_M^{-1}\nabla_{L^2}\mathcal{E}(\phi), \tag{18}$$

where

$$\mathcal{S}_M = M_0 - \operatorname{div}(M_1\nabla). \tag{19}$$

In the next section we will choose M_0 and M_1 such that the gradient flow (18) converges significantly faster than the standard H^1 flow in (12).

(a) well-conditioned problem (b) ill-conditioned problem

Figure 3: **2D Illustration of Well- and Ill-conditioned Problems:** The convergence rate of a well-conditioned problem (a) does not depend on the initial position or the particular dimension – in contrast to the one of an ill-conditioned problem (b), for which the convergence rate can be quite different for individual problem dimensions.

2.4 Metric Selection Based on the Preconditioning Rationale

We have seen in the last subsection that replacing the standard H^1 inner product by $\langle\cdot,\cdot\rangle_M$ still yields a smooth gradient flow, but with a different notion of distance depending on the choice of M_0 and M_1. Now we will use the freedom of choosing these metric tensors to speed up the gradient flow. We will select the metric tensors based on the idea of pre-conditioning.

Pre-conditioning is a standard technique for improvement of convergence properties of solvers for linear systems [13]. To this end, consider the second-order Taylor approximation of a function f in an arbitrary Hilbert space $\mathcal{H}$

$$f(x+\varepsilon h) = f(x) + \varepsilon\langle h, \nabla_{\mathcal{H}} f(x)\rangle_{\mathcal{H}} + \frac{\varepsilon^2}{2}\langle h, H_{\mathcal{H}}(f)h\rangle_{\mathcal{H}} + \mathcal{O}(\varepsilon^3)\ . \tag{20}$$

For a critical point x' with $\nabla f(x') = 0$, the first order term in (20) disappears and H dominantly describes the shape of f about x', so that the condition of H has a direct impact on the convergence of gradient-based methods, see also [15]. The idea of pre-conditioning is to multiply (20) by a certain pre-conditioner P, which approximates H^{-1} since this improves the condition of the resulting problem. It is in general important that P can be efficiently computed, so that the improvements in convergence can be translated into effective runtime.

One major observation is that for ill-conditioned problems, the convergence rates can vary significantly for the individual dimensions of the problem space, cf. Fig. 3. Inspecting the level set function, as well as the corresponding gradient for the standard H^1 flow, we see that the magnitude of the gradient at single spatial positions vary significantly, cf. Fig. 2. As the single spatial positions correspond to the dimensions of our problem, this results in varying convergence rates for each pixel. We thus conclude, based on the idea of pre-conditioning, that the adaptation of the magnitude of the gradient at the single spatial points should provide an improvement of convergence properties. To derive a Riemannian metric tensor based on the above argument, we start with the analysis of the level set function. Image segmentation can be considered as a binary decision for every pixel, which is reflected by the assumption

$$\phi(x) = \begin{cases} > 0, & x \in \text{foreground}, \\ \leq 0, & x \in \text{background}. \end{cases} \tag{21}$$

Thus we can assume that ϕ is a basin-shaped function such as $(G * \chi_\delta)(x)$, where G is a

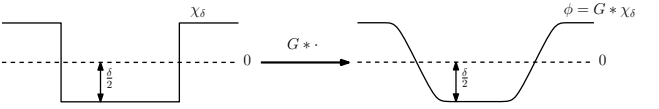


Figure 4: **Construction of a basin-shaped level set function ϕ.**

Gaussian smoothing kernel and

$$\chi_\delta(x) = \begin{cases} \frac{\delta}{2}, & x \in \text{foreground}, \\ -\frac{\delta}{2}, & x \in \text{background}, \end{cases} \tag{22}$$

where $\delta > 0$ is the *basin depth* as depicted in Fig. 4. In order to choose M_0 we first consider the L^2 gradient which is the core component of the H^1 as well as the H_M gradient flow (cf. (12) and (18)). Thus, the L^2 gradient is a function of x, I, ϕ, and $\nabla\phi$:

$$\nabla_{L^2}\mathcal{E}(\phi(x)) = F(x, I(x), \phi(x), \nabla\phi(x)). \tag{23}$$

This formula again illustrates why the L^2 gradient flow is very sensitive to local minima, because F depends on the local behavior of I at a single point x. As $\mathcal{S}_\alpha^{-1}$ smooths the L^2 gradient, the H^1 gradient flow is less sensitive to this local behavior. However, there is still a problem associated to the L^2 and the H^1 gradient flow. By taking a closer look at the first two rows of Fig. 2, we can observe that the local magnitude of the H^1 depends significantly on the position x resulting in a space-dependent convergence rate, which is an indication for an ill-conditioned problem as illustrated in Fig. 3. The reversion of this argument leads us to the idea that a pointwise normalization such as

$$\frac{1}{|\nabla_{L^2}\mathcal{E}(\phi(x))|/\delta}\nabla_{L^2}\mathcal{E}(\phi(x)) \tag{24}$$

could provide us with a uniform update size of δ resulting in a well-conditioned problem. Unfortunately, a simple rescaling of $\nabla_{L^2}\mathcal{E}(\phi)$ would not lead to a smooth gradient flow. Thus we propose to incorporate this normalization into the operator $\mathcal{S}_M$ by defining

$$M_0(x) = |\nabla_{L^2}\mathcal{E}(\phi(x))|/\delta + \gamma, \tag{25}$$

where $\gamma > 0$ is a small constant to ensure the positivity of M_0. As the inverse of S_M is applied to $\nabla_{L^2}\mathcal{E}(\phi)$ in (18), M_0 also gets inverted, resulting in a normalization of $\nabla_{L^2}\mathcal{E}(\phi)$. Further, by choosing $M_1 = \alpha > 0$, we maintain the smoothing behavior of $\mathcal{S}_M^{-1}$ which guarantees that we stay in H^1 during the evolution. Finally, $\mathcal{S}_M$ has the form

$$\mathcal{S}_M = \left(\frac{|\nabla_{L^2}\mathcal{E}(\phi(x))|}{\delta} + \gamma\right) - \alpha\Delta, \tag{26}$$

where $\delta = 1$ and $\gamma = 0.005\alpha$ have proved to be a good choice in all our experiments.

3 Experiments

We performed several experiments on images of the Weizmann database [1], because it provides also manual ground truth labelings. Before we discuss the results of our experiments in Sec. 3.3, we briefly describe the employed variational formulation in Sec. 3.1 and give all important details on the implementation in Sec. 3.2.

3.1 Variational Formulation

Denoting the image with $I : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ we consider the following energy to be minimized:

$$\mathcal{E}(\phi) = \lambda_G \mathcal{G}(\phi) + \lambda_P \mathcal{P}(\phi), \tag{27}$$

where

$$\mathcal{G}(\phi) = \int_\Omega \delta(\phi)\,|\nabla\phi|\,g\,dx, \tag{28}$$

is the *geodesic active contour model* [3] employing an *edge indicator function* such as

$$g(x) = \frac{1}{1+|\nabla I(x)|^2} \tag{29}$$

and the piecewise constant Chan-Vese model [5]:

$$\mathcal{P}(\phi) = \int_\Omega H(-\phi)(I-\mu_i)^2 + H(\phi)(I-\mu_o)^2\,dx. \tag{30}$$

Thereby $\mu_i, \mu_o \in \mathbb{R}$ denote the mean intensity values in and outside the contour and $\lambda_G, \lambda_P \in \mathbb{R}$ are the weights for each data term. In all our experiments we set $\lambda_P = 1$ and $\lambda_G = 0.3$. The resulting L^2 gradient then reads

$$\nabla_{L^2}\mathcal{E}(\phi) = \delta(\phi)\left[\lambda_G \mathrm{div}\left(g\frac{\nabla\phi}{|\nabla\phi|}\right) + \lambda_P\left((I-\mu_o)^2 - (I-\mu_i)^2\right)\right]. \tag{31}$$

3.2 Implementation

We use a forward Euler time discretization for the H^1 gradient flow in (12)

$$\phi^{t+\tau} = \phi - \tau\mathcal{S}_\alpha^{-1}\nabla_{L^2}\mathcal{E}(\phi^t) \tag{32}$$

as well as the H_M gradient flow (18)

$$\phi^{t+\tau} = \phi - \tau\mathcal{S}_M^{-1}\nabla_{L^2}\mathcal{E}(\phi^t). \tag{33}$$

For the inversion of $\mathcal{S}_\alpha$ and $\mathcal{S}_M$ we need to solve a sparse equation system. However, $\mathcal{S}_\alpha$ can also be inverted by computing the impulse response of its inverse on a small domain of size $2\lceil 4\sqrt{\alpha}\rceil \times 2\lceil 4\sqrt{\alpha}\rceil$ pixels and using it as a convolution mask. In general this decreases the runtime. In order to allow for a fair choice of the update size in every iteration step, we chose $\tau = \|\nabla_{L^2}\mathcal{E}\|_\infty^{-1}$. By employing this time step selection we ensure that the convergence rate for each method only depends on the structure of the update steps and not their size. This can be observed in Fig. 2.

In all our experiments we have normalized the image intensities such that $I(x) \in [0,1]$. The operators ∇ and div in (31) have been approximated with forward and backward finite differences, respectively. Further, it is useful to replace ∇I in (29) with $\nabla G * I/\sigma$, where $\sigma = 0.01$ and G is a truncated Gaussian kernel (window size 3 and standard deviation 0.5). For approximating $\delta(\phi)$ in (31) we employ its smeared-out version as suggested by Osher and Fedkiw in [14]:

$$\delta_\varepsilon(\phi) = \frac{1}{2\varepsilon}(1+\cos(\frac{\pi\phi}{\varepsilon})) \tag{34}$$

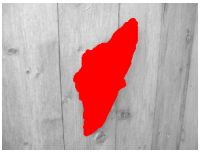
for $|\phi| \leq \varepsilon$,where $\varepsilon = 1.5$, and $\delta_\varepsilon(\phi) = 0$ otherwise.

(a) original image

(b) manual labeling

(c) result RSF

(d) original image



(e) manual labeling

(f) result RSF

Figure 5: **Test Images:** First row: The chopper image (a) segmented with the Riemannian Sobolev flow (RSF) (b). The manual ground truth obtained is given in (c). Second row: The moth image (d) segmented with the Riemannian Sobolev flow (RSF) (e). The manual ground truth obtained is given in (f). Images and the manual labelings are obtained from [1].

3.3 Discussion

We performed a large number of experiments on images of the Weizmann segmentation database [1], which are depicted in Fig. 5. For all experiments we used identical values for λ_G and λ_P. Our experiments reveal that the Riemannian Sobolev Flow (RSF) outperforms the the standard Sobolev flow (SF) as well as the standard Sobolev flow implemented with via convolutions (SFC) concerning both convergence rate and runtime. Moreover, Fig. 6 reveals that the SFC flow, although much faster then the SF flow, is prone to get stuck in local minima. This behavior is not surprising, since a convolution mask with finite size is cannot to fully describe the global behavior of an elliptic differential operator such as $\mathcal{S}_\alpha$. Another benefit of our method is that its behavior is far less dependent on the choice of α. This allows to choose the parameter α depending on the application in question without the necessity to take numerical concerns into account. As we chose the step size τ as well as the parameter γ in (26) automatically our method requires no additional parameter for minimizing a given energy $\mathcal{E}$.

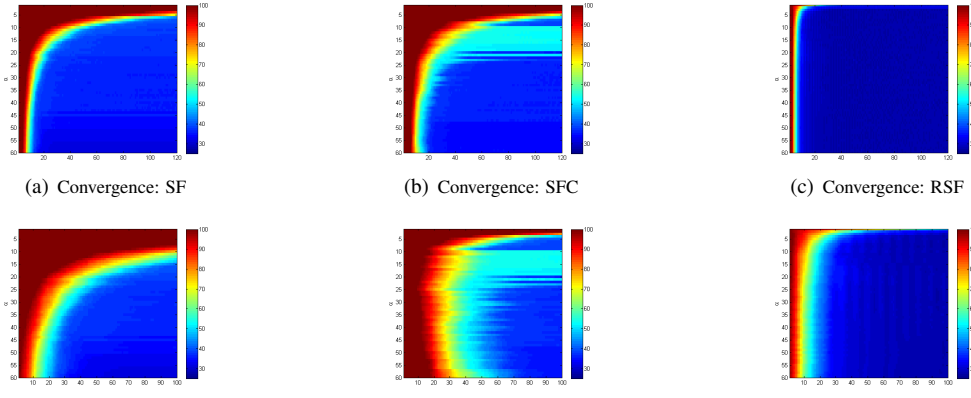
4 Conclusion

We have shown that Riemannian Sobolev spaces are a flexible and powerful tool to influence the properties of the resulting gradient flow for variational image segmentation. Motivated by an analysis of standard Sobolev flows, we propose to select the Riemannian metric tensors based on the idea of pre-conditioning, cf. also [19]. This results in a fast and simple to implement method, which is shown to yield a significant improvement in the convergence rate, while at the same time preserving the smoothness of the level set function. The improvement in convergence translates into effectively much shorter runtimes, which is of importance for practical applications. Furthermore, the proposed method exhibits improved performance, almost independently on the choice of the regularity parameter α, which allows α to be chosen based on the application requirements and without any numerical considerations. Regarding possible applications of the proposed technique one should, however, consider

the following aspects:

- The proposed metric is designed for the usage of basin-shaped level set functions. In case of signed-distance functions, one will have to alter the choice of M_0.
- It should be noted that our choice of the the smeared-out Dirac distribution is different from the one used by Chan and Vese in [6] making the segmentation problem a truly non-convex one – even if the mean values inside and outside the contour are known, cf. [6].

Future work might include a more extensive the evaluation of the proposed method as well as an application of the proposed method to localized models for image segmentation, e.g. [8].



(a) Convergence: SF (b) Convergence: SFC (c) Convergence: RSF

(d) Runtime: SF (e) Runtime: SFC (f) Runtime: RSF

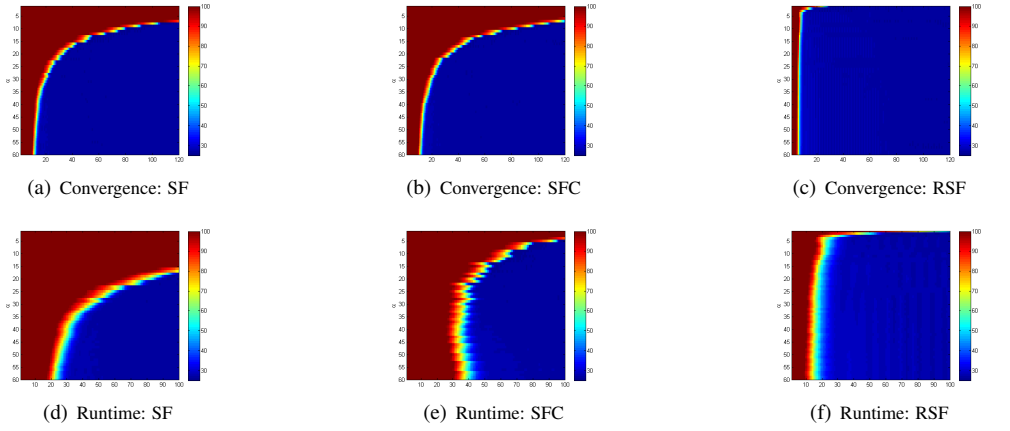
Figure 6: **Moth example continued:** We segmented the moth image in Fig. 5(d) using the H^1 Sobolev Flow (SF), the H^1 Sobolev flow implemented with convolutions (SFC), and the Riemannian Sobolev flow (RSF). The figures in the first row show the evolution of the L^2 error of the segmentation (w.r.t. to Fig. 5(f)) with respect to iterations, for different values of α. The second row plots the evolution of the L^2 error, with respect the runtime. In addition to the results depicted in Fig. 7 we observe that the SFC flow is prone to get stuck in local minima as indicated by the bright stripes in (b) and (e).

References

[1] S. Alpert, M. Galun, R. Basri, and A. Brandt. Image segmentation by probabilistic bottom-up aggregation and cue integration. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

[2] X. Bresson, S. Esedoglu, P. Vandergheynst, J. P. Thiran, and S. Osher. Fast global minimization of the active contour/snake model. *Journal of Mathematical Imaging and Vision*, 28(2):151–167, 2007.

[3] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal on Computer Vision*, 22(1):61–79, 1997.



(a) Convergence: SF (b) Convergence: SFC (c) Convergence: RSF

(d) Runtime: SF (e) Runtime: SFC (f) Runtime: RSF

Figure 7: **Chopper example continued:** We segmented the chopper image in Fig. 5(a) using the H^1 Sobolev Flow (SF), the H^1 Sobolev flow implemented with convolutions (SFC), and the Riemannian Sobolev flow (RSF). We performed the tests for different settings of the regularity parameter α. The first row shows the convergence by plotting the evolution of the L^2 error of the segmentation (w.r.t. to Fig. 5(c)) with respect to the iterations, for different settings of α. The second row shows again the evolution of the L^2 error, but this time with respect the actual runtime. Note that our method is characterized by a significantly higher convergence rate as well as an improved runtime – almost independent of α – in contrast to the other methods.

[4] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.

[5] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[6] Tony F. Chan, Selim Esedo Glu, and Mila Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. Technical report, SIAM Journal on Applied Mathematics, 2004.

[7] G. Charpiat, P. Maurel, J. P. Pons, R. Keriven, and O. Faugerás. Generalized gradients: Priors on minimization flows. *International Journal on Computer Vision*, 73(3):325–344, 2007.

[8] T. Goldstein and S. Osher. The split bregman method for l1-regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343, 2009.

[9] S. Lankton and A. Tannenbaum. Localizing region-based active contours. *IEEE Transactions on Image Processing*, 17(11):2029 –2039, 2008.

[10] C. Li, C. Xu, C. Gui, and M. D. Fox. Level set evolution without re-initialization: a new variational formulation. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2005.

[11] C. Li, R. Huang, Z. Ding, C. Gatenby, D. N. Metaxas, and J. C. Gore. A level set method for image segmentation in the presence of intensity inhomogeneities with application to mri. *IEEE Transactions on Image Processing*, 20(7):2007–2016, 2011.

[12] J. W. Neuberger. *Sobolev gradients and differential equations*. Springer Berlin, 1997.

[13] J. Nocedal and S.J. Wright. *Numerical optimization*. Springer, 2000.

[14] S. Osher and R. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer, 2003.

[15] RJ Renka. A simple explanation of the sobolev gradient method. 2006.

[16] J. A. Sethian. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. Cambridge University Press, 1999.

[17] G. Sundaramoorthi, A. J. Yezzi, and A. Mennucci. Sobolev active contours. *International Journal of Computer Vision*, 73(3):345–366, 2007.

[18] H. K. Zhao, T. Chan, B. Merriman, and S. Osher. A variational level set approach to multiphase motion. *Journal of Computational Physics*, 127(1):179–195, 1996.

[19] Darko Zikic, Maximilian Baust, Ali Kamen, and Nassir Navab. A general preconditioning scheme for difference measures in deformable registration. In *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2011.