

Joint Segmentation and Shape Regularization With a Generalized Forward–Backward Algorithm

Anca Stefanoiu, Andreas Weinmann, Martin Storath, Nassir Navab, and Maximilian Baust

Abstract—This paper presents a method for the simultaneous segmentation and regularization of a series of shapes from a corresponding sequence of images. Such series arise as time series of 2D images when considering video data, or as stacks of 2D images obtained by slicewise tomographic reconstruction. We first derive a model where the regularization of the shape signal is achieved by a total variation prior on the shape manifold. The method employs a modified Kendall shape space to facilitate explicit computations together with the concept of Sobolev gradients. For the proposed model, we derive an efficient and computationally accessible splitting scheme. Using a generalized forward–backward approach, our algorithm treats the total variation atoms of the splitting via proximal mappings, whereas the data terms are dealt with by gradient descent. The potential of the proposed method is demonstrated on various application examples dealing with 3D data. We explain how to extend the proposed combined approach to shape fields which, for instance, arise in the context of 3D+t imaging modalities, and show an application in this setup as well.

Index Terms—Object segmentation, image sequence analysis.

I. INTRODUCTION

WHEN segmenting multi-dimensional data sets such as data from computed tomography (CT), video data, or even 3D+t acquisitions as for instance arising from magnetic resonance imaging (MRI) data of a beating heart, one usually obtains not only a single, but a whole collection of shapes. These shapes typically enjoy a natural spatial, temporal, or even spatio-temporal ordering. In video sequences for example, the shapes corresponding to a tracked object have a natural temporal ordering. Another example is organ segmentation, where the segmentation is performed slicewise. The resulting shape collection, obtained from outlines of the segmentations, enjoys a natural spatial ordering induced by the ordering of the slices.

In both cases, we have a univariate sequence of shapes called *shape signals*. In case of 3D+t data, we even face a bivariate sequence of shapes which we call *shape fields*. Viewing a shape as a point in a shape manifold, such data can be seen as manifold-valued data sampled on a (regular) one- or two-dimensional grid. For an elaborate discussion on the manifold perspective in data processing we refer to [1]. We emphasize that shape spaces are challenging since they are high-dimensional manifolds (or, even Banach or Hilbert manifolds when considering non-polygonal model spaces).

In this paper, we present a method for the joint segmentation and total variation regularization of shape signals and shape fields. In contrast to an a-posteriori regularization of previously computed shape signals as proposed by Baust et al. [2], we will see that a joint segmentation and regularization is more robust to slice-wise errors while enjoying comparable computational costs.

A. Related Work

There is an extensive body of literature on segmenting spatial, temporal, or spatio-temporal collections of images such as variational approaches, e.g. Yezzi et al. [3], Unger et al. [4], graph-based methods, such as Shi and Malik [5], Tarabalka et al. [6], and Grundmann et al. [7], atlas-based techniques, e.g. Riklin Raviv et al. [8], probabilistic approaches, e.g. Xue et al. [9], methods using super-pixels, e.g. Papazoglou and Ferrari [10], or hybrid approaches such as Ochs et al. [11]. These examples employ spatial and temporal regularity constraints in Euclidean space, where, in contrast, the proposed method utilizes regularity constraints in a shape space which leads to a (non-flat) manifold setting. As the proposed approach involves concepts from active

Manuscript received September 29, 2015; revised March 3, 2016; accepted April 24, 2016. Date of publication May 11, 2016; date of current version June 7, 2016. The work of A. Weinmann was supported in part by the German Science Foundation through the Helmholtz Association within the young investigator Group under Grant VH-NG-526 and Grant WE 5886/3-1 and in part by the Deutsche Forschungs Gemeinschaft within the Young Researchers Network through the Project entitled Mathematics for Magnetic Particle Imaging. The work of M. Storath was supported in part by the European Research Council (ERC) within the Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne, Switzerland, through the European Union's Seventh Framework Programme FP7/2007-2013/ERC under Grant 267439 and in part by the Deutsche Forschungs Gemeinschaft within the Young Researchers Network through the Project entitled Mathematics for magnetic particle imaging. The work of M. Baust and N. Navab was supported by the German Science Foundation within the Collaborative Research Centre SFB 824 (Z2) through the project "Imaging for Selection, Monitoring and Individualization of Cancer Therapies". The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jianfei Cai. (Anca Stefanoiu and Andreas Weinmann contributed equally to this work.)

A. Stefanoiu is with the Computer Aided Medical Procedures and Augmented Reality Group, Institute of Computational Biology, Helmholtz Zentrum München, Technische Universität München, Munich 80333, Germany (e-mail: anca.stefanoiu@tum.de).

A. Weinmann is with the Department of Mathematics and Natural Sciences, Darmstadt University of Applied Sciences, Darmstadt 64295, Germany, and also with the Helmholtz Zentrum München, Department of Mathematics, Institute of Computational Biology, Technische Universität München, Munich 80333, Germany (e-mail: andreas.weinmann@helmholtz-muenchen.de).

M. Storath is with the Image Analysis and Learning Group, Heidelberg Collaboratory for Image Processing, Universität Heidelberg, Heidelberg 69117, Germany (e-mail: martin.storath@epfl.ch).

N. Navab and M. Baust are with the Computer Aided Medical Procedures and Augmented Reality Group, Technische Universität München, Munich 80333, Germany (e-mail: maximilian.baust@tum.de).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2016.2567068

contours (Sobolev-type regularization, active-contour-based tracking), shape spaces (Kendall shapes), and processing of manifold-valued data (cf. previous work [2]) we focus on the discussion of related work with respect to these categories.

1) *Active Contours and Sobolev Gradients*: In the area of active contours, it is a well-known fact that gradients computed from (noisy) image data yield unstable gradient flows, cf. [12]. Thus, regularization has been an important aspect of active contours right from the beginning, cf. Kass et al. [13]. A special regularization concept, which is also employed in this paper, is the one of Sobolev gradients; for an introduction we refer the interested reader to the book of Neuberger [14]. Sobolev gradient methods have been successfully applied to active contour methods by Charpiat et al. [15] and Sundaramoorthi et al. [16], [17] – in particular for tracking applications. However, it should be noted that tracking using Sobolev gradient methods only yields a forward regularity where the segmentation in the current frame is regularized by the previous ones, cf. also the discussion in [2]. We also note that – apart from the already cited works on active-contour-based tracking – there is a large body of literature on integrating various cues, such as shape, motion, or appearance, for making active contours more robust. Examples include, but are not limited to the works of Yezzi and Soatto [3], Cremers et al. [18], Rathi et al. [19], and Jackson et al. [20].

2) *Shape Spaces*: Similar to the recent work of Baust et al. [2], we employ a discrete shape space which is a variant of the one proposed by Kendall [21]. The choice of such a space is motivated by the existence of explicit formulas for computing geodesics, cf. also [2]. In contrast to this, there is also a considerable body of literature on continuous shape spaces, such as the articles of Michor and Mumford [22], Younes et al. [23], Michor and Mumford [24], Srivastava et al. [25], Wirth et al. [26], or Bauer et al. [27], [28] to name just a few. We note that some of these works had also an influence on the development of Sobolev gradient methods for active contours, e.g. [22]. We also mention that, as long as a particular shape representation facilitates the computation of geodesics, the proposed algorithm can be adapted to this shape space. This includes the works of Srivastava et al. [25], Wirth et al. [26], and Bauer et al. [28].

3) *Processing Manifold-Valued Data*: The processing of manifold-valued data has gained considerable interest during the last decade, but total variation regularization for manifold-valued data has only been considered very recently, cf. Lellmann et al. [29], Weinmann et al. [30], and Grohs et al. [31]. Concerning the theoretical foundations of total variation for manifold-valued data we refer the interested reader to Giaquinta and Mucci [32], [33]. Previous algorithmic approaches for total variation regularization of manifold-valued data only considered rather low dimensional manifolds, such as the one and two-dimensional sphere as well as the three-dimensional group of rotations [29], [30], [34], or the space of symmetric positive matrices for processing diffusion tensor data [30], [35]. High-dimensional manifolds, such as shape spaces, have been, for instance, considered by Durrleman et al. [36], Fletcher [37], Samir et al. [38],

Su et al. [39], or Shrivastava et al. [40]. The paper [2] has been the first to consider total variation regularization for shape spaces. In [2], however, only a pure denoising/regularization setup is considered. Moreover, all of these methods take manifold-valued data as an input whereas our method directly works with the given images. In contrast to this, the present work combines the segmentation of individual shapes with the regularization of shape signals or even shape fields. We mention the relation between total variation minimization and iterative median filtering in linear spaces [41]; median filtering for shape spaces has been considered by Berkels et al. [42].

B. Contributions

In this paper, we provide a novel variational formulation for the joint segmentation and total variation (TV) regularization of shapes in a Riemannian manifold setup. A manifold setting is natural since shape spaces are not vector spaces but typically non-linear manifolds. Similar to [2], we choose a variant of the Kendall shape space as underlying manifold, because it allows for an efficient computation of the exponential as well as the inverse exponential map. We develop an algorithm for the joint segmentation and TV regularization of shapes (shape signals and shape fields). We provide a splitting of the model which is computationally accessible using a generalized forward-backward scheme. For the forward part, we employ a Sobolev-gradient-type scheme, cf. Sundaramoorthi et al. [12] and Charpiat et al. [15], in order to allow for a stable numerical treatment. Our approach is efficient since all steps of the algorithm have linear complexity implying a very low overall runtime. Furthermore, we note that the algorithm is parallelizable to a large extent. Finally, we provide an experimental validation of the algorithm. As a baseline for the proposed method, we consider a non-joint two-stage approach. The first step consists of an item-wise plain active contour segmentation method based on Sobolev gradients [12], [15]. The second step is the regularization of the obtained shapes proposed in [2]. We compare the proposed approach with this baseline approach as well as with the state-of-the-art method of Grundmann et al. [7] in order to show its potential. We note that the proposed framework is also applicable to other shape spaces such as the ones discussed in [27].

C. Organization of the Article

In Section II we derive the proposed approach for joint segmentation and shape regularization. After introducing the concept, we derive the proposed model Section II-A and Section II-B. In Section II-C, Section II-D and Section II-E we derive the proposed algorithmic framework and concrete algorithms. Finally, extensions to multivariate domains are discussed in Section II-F.

In Section III we perform numerical experiments to show the potential of the proposed approach. In the main part which is Section III-A we compare the proposed approach with the baseline approach of sequential segmentation and regularization as well as with the state-of-the-art method of Grundmann et al. [7]. In Section III-B, we briefly illustrate the application of the method to shape signals.

II. PROPOSED APPROACH FOR JOINT SEGMENTATION AND SHAPE REGULARIZATION

In order to deal with the competing goals of approximating the data and achieving regularity of the shapes over multiple frames/slices, we consider the variational model

$$\min_{U \in \mathcal{U}} \mathcal{J}_\gamma(U, F), \quad \text{where} \quad (1)$$

$$\mathcal{J}_\gamma(U, F) = \mathcal{D}(\mathcal{A}(U), F) + \gamma \mathcal{TV}(U). \quad (2)$$

Here, \mathcal{D} measures the distance between the measured data F and a nonlinear measuring/modeling operator \mathcal{A} applied to an element U belonging to the signal space modeling \mathcal{U} . The input data $F = (F_i)_i$ is a sequence of images F_i with a (temporal or spatial) ordering, for example a video. The Riemannian manifold \mathcal{U} is a modified Kendall shape space as explained below. The operator \mathcal{A} is a mapping from the Riemannian manifold to the vector space of measurements. More precisely, \mathcal{A} maps a sequence of shapes to a sequence of binary images that represent these shapes. \mathcal{TV} denotes a total variation (TV) regularizer acting on the Riemannian manifold; it is the sum of distances between consecutive shapes. $\gamma > 0$ is a model parameter which controls the balance between data fidelity and regularity.

The remainder of this section is organized as follows. We start with the underlying model space \mathcal{U} which is a (modified) Kendall shape space as explained in Section II-A. Then, we will concentrate on the energy functional acting on \mathcal{U} in Section II-B and derive a generalized forward backward algorithm for the proposed model in Section II-C. The building blocks required for the algorithm are finally derived in Section II-D and Section II-E, respectively. We discuss extensions to multivariate domains in Section II-F.

A. Model Space–Kendall Shape Space

As a model space, we consider a (modified) Kendall shape space which is, due to constraints, a Riemannian manifold rather than a linear space. Besides its simplicity, a major advantage of the employed space is that the Riemannian exponential mapping and its inverse can be implemented using closed form expressions and thus yielding high computational efficiency of the proposed approach. The classical shape space introduced in the context of underlying planar curves by Kendall in [21] is invariant w.r.t. rotations, scalings, and translations. Similar to [2], we here employ a rotation-sensitive variant termed *oriented Kendall shapes*. For the intended application, this representation is particularly suited for the TV regularization of shape signals as demonstrated in [2].

Let us recall the concept: We consider a polygon which discretizes a simply connected planar shape. These are shapes represented by closed curves in two-dimensional space which do not intersect themselves. We obtain an n -gon which can be represented by a vector

$$z = (z_1, \dots, z_n) \in (\mathbb{R}^2)^n, \quad (3)$$

where each entry $z_m \in \mathbb{R}^2$ is a vertex of this polygon. Note that we do not employ a repetition of the first point, i.e., $z_1 \neq z_n$.

We further assume a normalization with respect to translation which means that

$$\sum_{m=1}^n z_m = 0 \in \mathbb{R}^2. \quad (4)$$

This restricts the shape representations to the subspace

$$V_{2n-2} = \{z \in (\mathbb{R}^2)^n : \sum_{m=1}^n z_m = 0\}. \quad (5)$$

Since a shape z can be scaled by any real number $s > 0$, all equivalent shapes (all representations of the same shape) lie on the half-line

$$L_z = \{s \cdot z : s \in \mathbb{R} \setminus \{0\}\}. \quad (6)$$

This means that we obtain a space

$$\mathcal{S} = V_{2n-2}/L \quad (7)$$

isomorphic to the sphere \mathbb{S}^{2n-3} . Using this model, the exponential mapping and its inverse are given by

$$\exp_z(v) = \cos(\phi) \cdot z + \frac{\|z\| \sin(\phi)}{\phi} \cdot v, \quad \phi = \|v\|, \quad (8)$$

and

$$\log_z(y) = \phi \cdot \frac{y - \Pi_z(y)}{\|y - \Pi_z(y)\|}, \quad \phi = \arccos\left(\frac{\langle z, y \rangle}{\|z\| \|y\|}\right). \quad (9)$$

Since these operations only require basic linear algebra operations, they can be implemented efficiently using the BLAS library.

B. Proposed Model

We consider a variational model of the form (2) on the oriented Kendall shape space \mathcal{S} . To explain the model, let us start with a discrete signal $\{U_i\}_i$ of shapes; shape fields $(U_{ij})_{ij}$ are a straightforward generalization. As data term, we consider frame-wise operations of the form

$$\mathcal{D}(\mathcal{A}(U), F) = \sum_i \mathcal{D}(\mathcal{A}(U_i), F_i). \quad (10)$$

In order to prevent the notation from becoming too complicated, we use the symbols \mathcal{A} and \mathcal{D} for the component-wise operations as well. The measurement operator \mathcal{A}_i acting on the i -th component is mapping shapes in \mathcal{S} to implicit curve representations, i.e., images with values in $[0, 1]$ defined on a $k \times l$ -grid. However, being elements of the oriented Kendall shape space, the shapes U_i do not have a particular position or scale. Thus, the measurement operator \mathcal{A} has to endow each representation of the shape with a position x (translational part) and a scaling factor s to obtain a particular implicit representation of the form

$$\mathcal{A}_{x,s}(U_i) = \mu_{\text{in}} \mathbf{1}_{\Omega_{x,s}} + \mu_{\text{out}} \mathbf{1}_{\Omega_{x,s}^c} \quad (11)$$

as illustrated in Fig. 1. $\Omega_{x,s}$ denotes the interior of the shape U_i anchored at x and normalized such that the sum over the squared distances to the anchor equals s^2 . In order to keep the notation simple, we drop the dependency of x and s on i . The symbol $\Omega_{x,s}^c$ denotes the complement of $\Omega_{x,s}$ which is the outer part determined by the curve U_i . Furthermore, μ_{in} and μ_{out} denote the means of intensity values

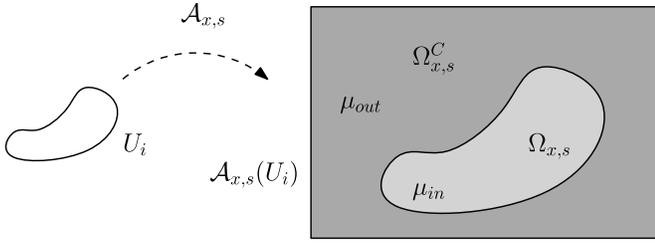


Fig. 1. Action of the Measurement Operator: $\mathcal{A}_{x,s}$ maps a Kendall shape U_i to a translated, i.e., positioned, and scaled version $\mathcal{A}_{x,s}(U_i)$ in the image space. $\Omega_{x,s}$ and $\Omega_{x,s}^C$ denote the areas inside and outside $\mathcal{A}_{x,s}(U_i)$, respectively.

of $I = F_i$ (i th frame) computed w.r.t. $\Omega_{x,s}$ and $\Omega_{x,s}^C$, respectively. The global measurement operator \mathcal{A} now consists of all $\mathcal{A}_{x,s}$.

Using the standard L^2 distance on $[0, 1]^{k \times l}$ we define the data term $\mathcal{D}(\mathcal{A}(\cdot), F_i)$ as

$$\begin{aligned} \mathcal{D}(\mathcal{A}(U_i), F_i) \\ = \min_{x,s} \left(\sum_{j \in \Omega_{x,s}} |(F_i)_j - \mu_{in}|^2 + \sum_{j \in \Omega_{x,s}^C} |(F_i)_j - \mu_{out}|^2 \right). \end{aligned} \quad (12)$$

We note that the term in brackets is similar to the data term employed in the Chan-Vese model [43] which is the two-phase version of the classical Potts model (or piecewise constant Mumford-Shah model) [44]. We further note that, keeping the shape fixed, the minimization problem in (12) is a registration problem w.r.t. position and scale. The minimization over x and s in Eq. (12) reflects the necessary scale selection process as the shapes U_i do not carry the respective information any more. Since already one function evaluation requires solving a registration problem, a naive approach to the corresponding variational problem (2) is not reasonable from a computational perspective. Thus, we propose an approximate strategy which employs the concept of Sobolev gradients as described in Section II-D.

Since the model space is a Riemannian manifold, any regularization involving neighboring data items has to be formulated in terms of the distance induced by the Riemannian metric. We here consider the (discrete) TV regularizer

$$\mathcal{TV}(U) = \sum_i d(U_i, U_{i+1}) = \sum_i \|\log_{U_i}(U_{i+1})\|_{U_i}, \quad (13)$$

where d denotes the distance induced by the Riemannian metric on the considered Kendall shape space as defined in Section II-A. This means that $d(U_i, U_{i+1})$ is the length of the shortest geodesic joining the two shapes U_i and U_{i+1} . As it is well-known in Riemannian geometry, the distance between these shapes agrees with the length of the tangential element $\log_{U_i}(U_{i+1})$ in the tangent space at the shape U_i . This length is measured by the norm $\|\cdot\|_{U_i}$ which is given by the square-root of the Riemannian metric (i.e., a smoothly varying tensor field defining a scalar product at each point U_i) applied to $\log_{U_i}(U_{i+1})$. Here, $\log_{U_i}(U_{i+1})$ denotes the (locally unique) tangent vector at U_i such that the geodesic starting at U_i with this tangent vector reaches U_{i+1} at time $t = 1$.

We briefly explain why we call the manifold object in (13) *total variation*. Note that for a sequence of points $(x_i)_i \subset \mathbb{R}^2$ in the Euclidean plane, i.e., \mathbb{R}^2 with the standard scalar product

Algorithm 1 Proposed Scheme for Problem (1)

input : Video sequence $(F_i)_i$, number of steps N

output: final shape signal $(U_i)_i$

Initialize shape signal $(U_i)_{i=1:K}$;

for $l \leftarrow 1$ **to** N **do**

$\lambda_l \leftarrow \text{CompLambda}(l)$; // compute stepsize

for $i \leftarrow 1$ **to** K **do**

$G_i \leftarrow \text{CompSobolevGradient}(U_i, F_i)$;

 // cf. Sec. 3.4

$U_i \leftarrow \exp_{U_i}(-\lambda_l G_i)$; // forward (descent) step

end

for $i \leftarrow 1$ **to** $K-1$ **do**

$(U_i, U_{i+1}) \leftarrow \text{proxMap}(U_i, U_{i+1}, \gamma \lambda_l)$ // cf. Sec. 3.5

end

end

at each point, we have $\|\log_{x_i}(x_{i+1})\|_{x_i} = \|x_{i+1} - x_i\|_2$. Hence, $\mathcal{TV}((x_i)_i) = \sum_i \|x_{i+1} - x_i\|_2$ is just a discrete version of the total variation, or the total variation of the corresponding polygonal curve. When the underlying manifold is \mathbb{R} , our definition agrees with the total variation of the corresponding Radon measure.

C. A Generalized Forward Backward Approach

We propose an algorithm for problem (2) with data term (12) and TV regularizer (13). We split our target functional in a way such that it is computationally accessible to a generalized forward-backward scheme. The used terminology is in the spirit of Raguet et al. [45] who propose a related splitting scheme for vector valued data. A related approach for diffusion tensor imaging (DTI) can be found in [46].

At first, we split the TV regularizer \mathcal{TV} according to

$$\mathcal{TV}(U) = \sum_i \mathcal{TV}_i(U), \quad \mathcal{TV}_i(U) = d(U_{i+1}, U_i).$$

Using this splitting, we write the functional \mathcal{J}_γ as

$$\mathcal{J}_\gamma(U) = \mathcal{D}(\mathcal{A}(U), F) + \gamma \sum_i \mathcal{TV}_i(U). \quad (14)$$

In the forward step, we treat the data term $\mathcal{D}(\mathcal{A}(U), F)$ in an explicit way, which means that we apply a gradient descent for the data term $\mathcal{D}(\mathcal{A}(U), F)$. To this end, we recall from (12) that the data term is separable, i.e., $\mathcal{D}(\mathcal{A}(U), F) = \sum_i \mathcal{D}(\mathcal{A}(U_i), F_i)$. This means that the forward step reduces to dealing with all individual atoms $\mathcal{D}(\mathcal{A}(U_i), F_i)$ separately, which will be explained in Section II-D.

In the backward step, we treat the non-differentiable atoms \mathcal{TV}_i implicitly. This means that the next iterate is determined by an implicit condition. More precisely, we compute the proximal mappings of these atoms (see Section II-E) and apply them in a cyclic way, i.e., we successively apply the proximal mappings of $\mathcal{TV}_1, \mathcal{TV}_2, \mathcal{TV}_3, \dots$. This corresponds to one step of a cyclic proximal point algorithm [48], [49].

The resulting algorithm performs an iteration of the forward step, using the Sobolev gradient method, and backward steps applying the proximal mappings of the TV atoms cyclically. We summarize the iteration scheme in Algorithm 1. The stepsize (for both explicit and implicit steps) within the iteration

is chosen such that it is square-summable but not summable which is needed for the backward step. Furthermore it is chosen such that it does not increase the energy in the forward step. To this end, we perform a line search to determine the optimal step size for each data term atom. Next, we take the minimum of these step sizes and $3l^{-(0.95+0.5l^{-0.18})}$ (where l denotes the iteration number, cf. [2]) in order to obtain the final step size.

We observe convergence of the proposed algorithm in the performed experiments, cf. Section III. A proof is, however, challenging and topic of future research; convergence in the mathematical sense has not been proven yet. So far, convergence has only been shown in the non-inverse setup for Hadamard spaces [30] as well as for the 1D circle with rather severe restrictions [50].

D. Computing Sobolev Gradients

The data term \mathcal{D} acting on the sequence U may be written as

$$\min_{x_1, \dots, s_1, \dots} \sum_i \sum_j |\mathcal{A}_{x_i, s_i}(U_i)_j - (F_i)_j|^2. \quad (15)$$

Neglecting the regularizer for the moment, the minimization of (2) requires minimizing (15) w.r.t. the shapes U_i in the shape space \mathcal{S} . We note that the triples (V, y, t) of shapes $V \in \mathcal{S}$, translations y and scalings t , parametrize the space of plain curves. Thus minimizing (15) w.r.t. U_i is equivalent to minimizing the sums in (15) with respect to plain curves given by the corresponding triples (U_i, x_i, s_i) . In other words, the minimization of (15) corresponds to a classical active contour evolution.

We recall the standard principle of evolving active contours. Assuming that $\mathcal{A}_{x_i, s_i}(U_i)$ is registered to the image data F_i , one needs to compute a deformation field

$$\vec{V}_{x_i, s_i} = \alpha(z_m) \eta_m, \quad (16)$$

where η_m denotes the (approximate) outer normal at the m -th boundary point z_m ($m = 1, \dots, n$), cf. Fig. 4, and

$$\alpha(z_m) = |F_i(z_m) - \mu_{\text{in}}|^2 - |F_i(z_m) - \mu_{\text{out}}|^2 \quad (17)$$

is the velocity in normal direction evaluated at z_m . In standard active contours, one evolves the contour according to \vec{V}_{x_i, s_i} to improve the segmentation. However, in order to evolve the shape U_i itself, we remove the scale and the translational component of \vec{V}_{x_i, s_i} before applying it to U_i . Denoting this scale and translation adjusted version of \vec{V}_{x_i, s_i} by \vec{V} , the forward step reads

$$\exp_{U_i}(-\tau \vec{V}), \quad (18)$$

where \exp_{U_i} denotes the exponential mapping at U_i . This update step is, of course, also applicable in connection with a regularizing term.

It is important to note that we perform the registration of the shapes in an incremental way: Based on the initialization of U_i we are able to specify (and store) initial values for x_i and s_i . When computing the deformation field \vec{V}_{x_i, s_i} , we simply apply the current values x_i and s_i , estimate the scale change Δs_i as well as the translational change Δx_i caused by \vec{V}_{x_i, s_i} , and update x_i and s_i accordingly.

In practice, \vec{V}_{x_i, s_i} can be severely affected by noise as it inherits all the local behavior of F_i . To deal with this issue, we propose to regularize \vec{V}_{x_i, s_i} based on the concept of Sobolev gradients [12], [15]. Therefore, we interpret the scaled and translated version of U_i as the discretized version of a continuous curve \mathcal{C} in the manifold of smooth immersed curves modulo parametrization, cf. [12]. This allows us to consider \vec{V}_{x_i, s_i} as an element of the tangent space of this manifold at \mathcal{C} . This tangent space can now be endowed with the inner product

$$\langle \vec{V}, \vec{W} \rangle_{H^1} = \text{avg}(\vec{V}) \cdot \text{avg}(\vec{W}) + \theta L \int_0^L \vec{V}_s \cdot \vec{W}_s ds, \quad (19)$$

where s denotes the arc length, L the length of \mathcal{C} , \cdot_s the partial derivative w.r.t. s , and $\text{avg}(\cdot)$ the translational component of the respective argument, cf. [12]. Computing active contour gradients w.r.t. to this inner product yields so-called Sobolev gradients which favor translations over smooth deformations. Fortunately, there is a simple and efficient way of turning \vec{V}_{x_i, s_i} into a Sobolev gradient: we just have to convolve it with the smoothing kernel

$$K_\theta = \frac{1}{L} \left(1 + \frac{(\frac{s}{L})^2 - (\frac{s}{L}) + 1/6}{2\theta} \right) \quad (20)$$

as described in [12]. As suggested in [12] we chose $\theta > 1/24$. Here, we use $\theta = 0.05$ for all experiments. Based on $K_\theta * \vec{V}_{x_i, s_i}$ we can now reliably estimate Δs_i and Δx_i as described in the previous paragraph.

E. Proximal Mappings of the Regularizer

As stated in Section II-C, the proximal mappings of the regularization atoms $\mathcal{TV}_i(U) = d(U_i, U_{i+1})$ can be computed explicitly. More precisely, the corresponding proximal mapping $\text{prox}_{\lambda \mathcal{TV}_i}$ is, for $\lambda > 0$, defined by

$$\text{prox}_{\lambda \mathcal{TV}_i}(U) = \text{argmin}_{U'} \frac{1}{2} d^2(U', U) + \lambda d(U'_i, U'_{i+1}). \quad (21)$$

As shown in [30], these proximal mappings $\text{prox}_{\lambda \mathcal{TV}}$ can be computed explicitly:

$$\begin{aligned} & [\text{prox}_{\lambda \mathcal{TV}_i}(U)]_j \\ &= \begin{cases} U_j, & \text{if } j \neq i, i+1, \\ \exp_{U_i} \left(\min(\lambda, \frac{1}{2}) \exp_{U_i}^{-1} U_{i+1} \right), & \text{if } j = i, \\ \exp_{U_{i+1}} \left(\min(\lambda, \frac{1}{2}) \exp_{U_{i+1}}^{-1} U_i \right), & \text{if } j = i+1. \end{cases} \end{aligned} \quad (22)$$

The relevant terms take two shapes as input, and compute two new shapes which are closer using (9) and (8). These new shapes both lie on the geodesic joining the old shapes. Each new shape always stays nearer to its corresponding old associate than the other new shape with the extreme case of both new shapes being equal. The distances to their old associates are equal for both new shapes.

F. Extension to Shape Fields

In order to extend the approach to shape fields, i.e. two dimensional collections of shapes, we extend the functional (2) such that they deal with arguments defined on a multivariate grid. This is no issue for the separable data term, and we obtain

$$\mathcal{D}(\mathcal{A}(U), F) = \sum_{ij} \mathcal{D}(\mathcal{A}(U_{ij}), F_{ij}). \quad (23)$$

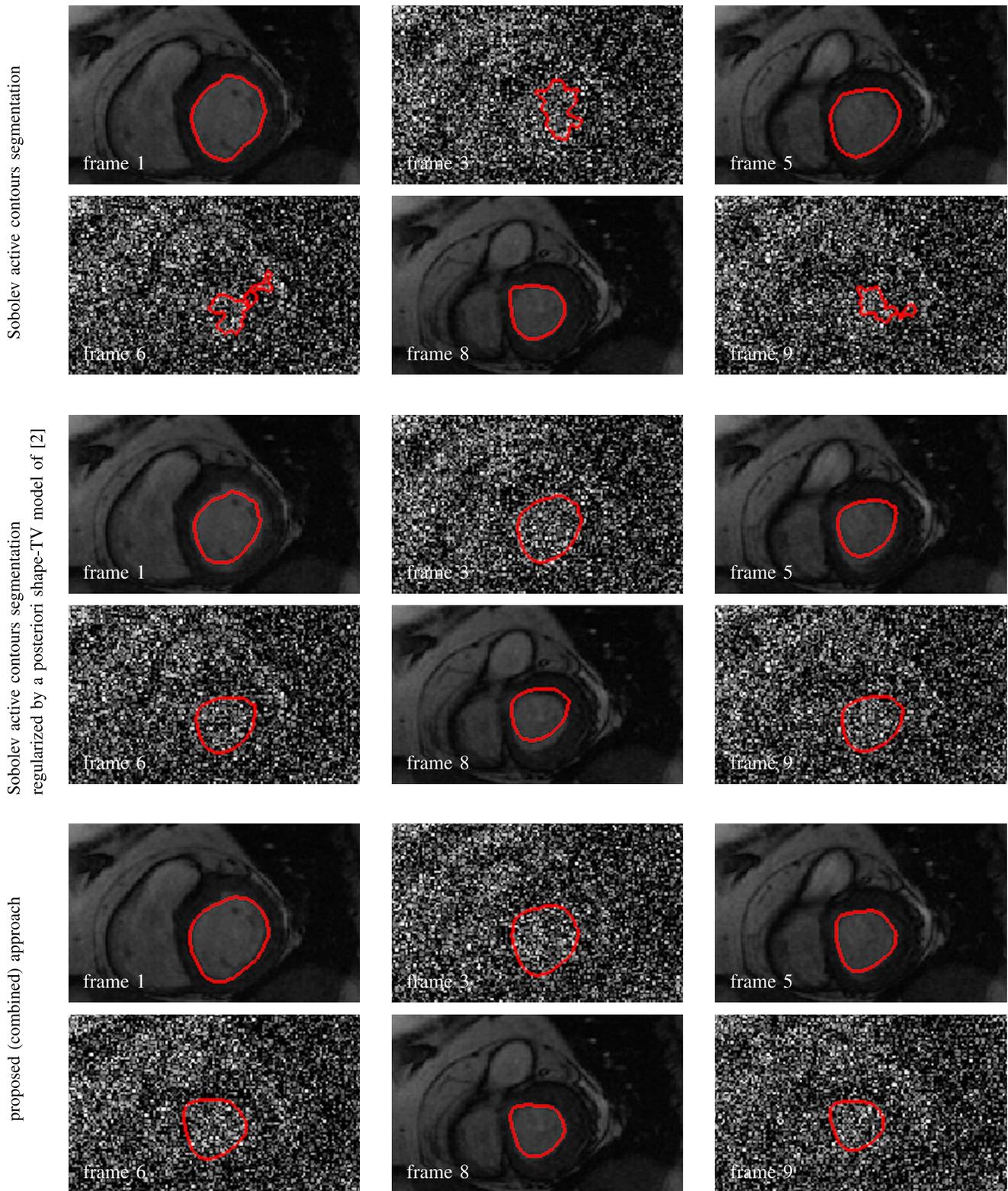


Fig. 2. Benefits of the Joint Approach on a cardiac MRI image sequence (six slices shown) with every third member (artificially) corrupted by noise (image data from [47]). *First and second row.* Segmentation using a Sobolev active contour w.r.t. the piecewise-constant Mumford-Shah model. The segmentation fails on those frames corrupted with noise. *Third and fourth row.* The shape signal of the first row regularized with the method of [2]. Most of the information can be restored, but the faulty segmentations distort the correct ones from the true myocardial boundary in the neighboring slices. *Fifth and last row.* Regularized shape signal obtained with our method. Note that segmentations in the noise-free frames are not affected.

Since the TV term is non-separable, its extension to 2D is a bit more involved. We use the simplest anisotropic variant

$$\mathcal{TV}(U) = \sum_{i,j} d(U_{ij}, U_{i+1,j}) + \sum_{i,j} d(U_{ij}, U_{i,j+1}). \quad (24)$$

To get more isotropic variants, one might also include additional diagonal or knight move directions; see [52], [53]. The proposed algorithm generalizes easily to 2D: We only notice

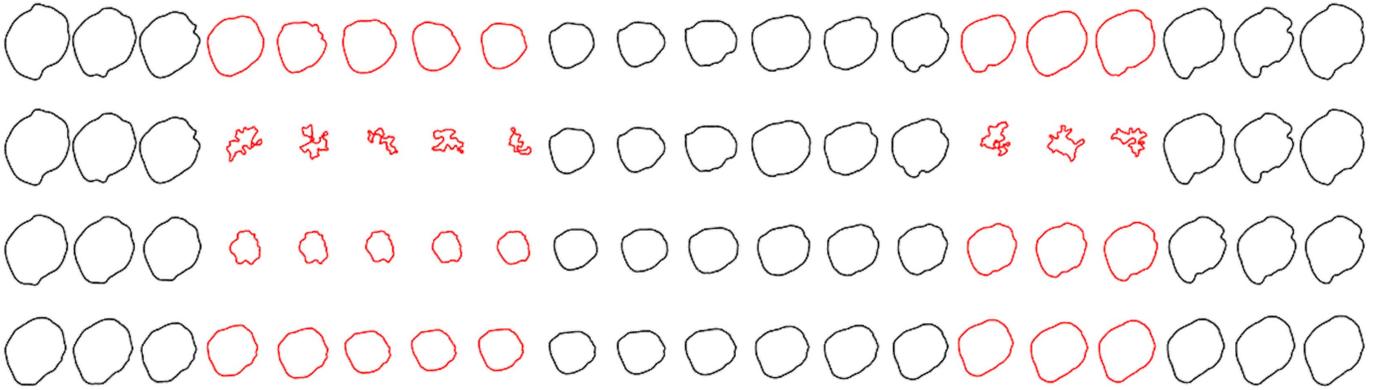


Fig. 3. Robustness w.r.t. to Consecutive Corrupted Frames: In contrast to the experiment in Fig. 2, we corrupted multiple consecutive frames of the left ventricle data set (Fig. 2) with noise – in this case frames 4,5,6,7, and 8 as well as 15,16, and 17. We display the shape signals obtained with different methods, where the shapes of the corrupted frames are highlighted in red. From top to bottom: plain segmentation with noise; Sobolev active contours segmentation regularized by a posteriori shape-TV model of [2]; proposed approach. The proposed approach comes closest to the shape signals obtained from noise-free data.

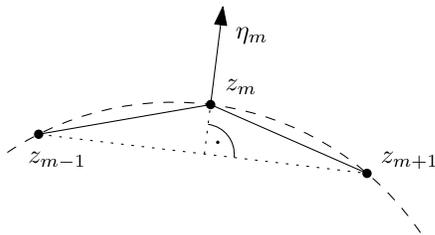


Fig. 4. Estimation of the Outer Normal: The outer normal η at point z_i is estimated via the orthogonal projection of z_i onto the line $[z_{i-1}, z_{i+1}]$.

that proximal mappings are now to be computed for all atoms $d(U_{ij}, U_{i+1,j})$ and $d(U_{ij}, U_{i,j+1})$ in (24).

III. EXPERIMENTS

In order to demonstrate the potential of the proposed approach, we performed a qualitative comparison, see Figs. 2, 3, and 6, as well as a quantitative comparison, see Tab. I and II. The baseline approach consists in segmenting the respective image data using a two-phase piecewise constant model using a Sobolev active contour method (without additional regularization), cf. [12]. We further compare our method to a state-of-the-art unsupervised video segmentation method of Grundmann et al. [7] which in particular takes the temporal regularity of the segments into account. Finally, we provide more applications (Fig. 5) as well as an extension to shape fields (Fig. 8).

A. Comparison

In Fig. 2, we segmented a temporal sequence of one slice of the left ventricular cavity acquired with magnetic resonance imaging (MRI). The first two rows show the segmentations obtained with a plain Sobolev active contour method minimizing the piecewise constant model. One may easily notice that the segmentation fails in those frames which are artificially corrupted by noise. Using this faulty shape sequence as an input for the method proposed in [2] (with regularization

TABLE I
ROBUSTNESS W.R.T. INITIALIZATION: WE ADDED WHITE GAUSSIAN NOISE TO THE INITIALIZATIONS OF THE BRAIN TUMOR SEQUENCE IN FIG. 5 AND COMPUTED THE AVERAGE SURFACE DISTANCE TO THE SEGMENTATIONS OBTAINED WITHOUT NOISE ADDED. WE CONCLUDE THAT THE PROPOSED METHOD IS SLIGHTLY MORE ROBUST W.R.T. THE INITIALIZATION. NOTE THAT THE REGULARIZATION PARAMETERS ARE NOT NECESSARILY COMPARABLE

	$\sigma = 1$	$\sigma = 2$	$\sigma = 3$
plain seg.	0.631	1.042	1.456
plain seg. & [2], $\gamma = 0.5$	0.545	0.964	1.365
plain seg. & [2], $\gamma = 1.0$	0.492	0.929	1.301
plain seg. & [2], $\gamma = 1.5$	0.494	0.910	1.249
plain seg. & [2], $\gamma = 2.0$	0.493	0.910	1.240
plain seg. & [2], $\gamma = 2.5$	0.493	0.910	1.240
plain seg. & [2], $\gamma = 3.0$	0.493	0.910	1.240
proposed meth., $\gamma = 0.5$	0.426	0.889	1.095
proposed meth., $\gamma = 1.0$	0.423	0.879	1.064
proposed meth., $\gamma = 1.5$	0.423	0.879	1.064
proposed meth., $\gamma = 2.0$	0.423	0.879	1.064
proposed meth., $\gamma = 2.5$	0.423	0.879	1.064
proposed meth., $\gamma = 3.0$	0.423	0.879	1.064

parameter $\gamma = 2.0$) we obtain a nicely regularized shape signal, cf. second two rows of Fig. 2. However, as the method of [2] is unaware of the original image data, the corrupted segmentations distort the shapes of the neighboring frames. In contrast to this, the proposed approach, with regularization parameter $\gamma = 1.0$, shown in the last two rows of Fig. 2, does not suffer from this issue as it uses the original image data as “anchors”. This effect becomes even more prominent when corrupting a series of images with noise as done in Fig. 3, where we applied noise to frames 4, 5, 6, 7, and 8 as well as 15, 16, and 17. We observe in Fig. 3 that the proposed approach comes closest to the shape signals obtained from noise-free data.

We also investigated the robustness of the proposed approach w.r.t. to the initialization. Therefore, we added white

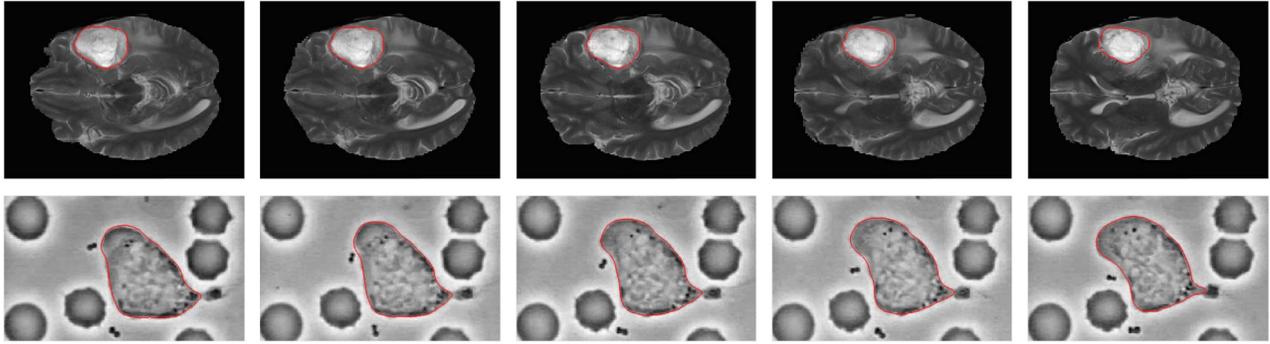


Fig. 5. Further Application Examples: First row: Multiple slices of a brain tumor segmented from MRI data ($\gamma = 2$), which has also been used for the quantitative analysis in Tab. I (image data from [51]). Second row: Several slices of a microscopic video (by David Rogers at Vanderbilt University) showing a leukocyte chasing a bacterium. Both data sets have been used for the quantitative analysis in Tab. II.

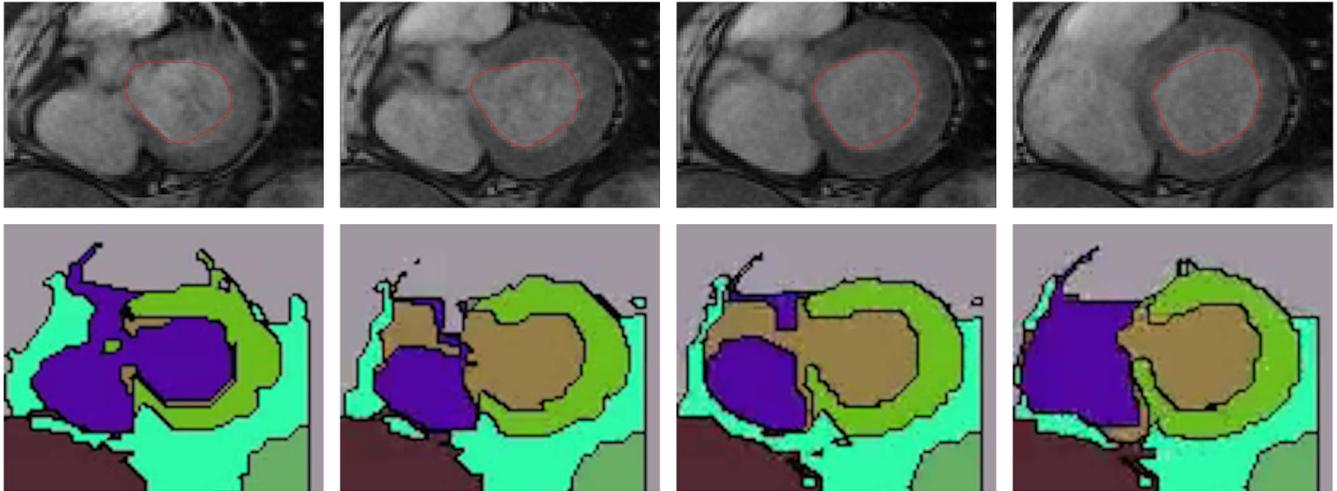


Fig. 6. Comparison to a State-of-the-art Method with Temporal Regularity Constraints [7]: We run our method on 2D+t MRI data (100×100 pixel, 20 frames). *Bottom:* The method of [7] is affected by the adjacent blood pool (result obtained with the implementation provided at www.videosegmentation.com, runtime (only processing) 15 sec.). *Top:* The proposed method segments the left ventricular cavity reliably (runtime 0.1 sec.).

TABLE II

ROBUSTNESS W.R.T. IMAGE NOISE: WE ADDED WHITE GAUSSIAN NOISE TO THE INPUT IMAGES OF THE BRAIN TUMOR SEQUENCE (FIG. 5), THE CELL SEQUENCE (FIG. 5), AND THE LEFT VENTRICLE SEQUENCE (FIG. 2). THEN WE COMPUTED THE AVERAGE SURFACE DISTANCE TO THE SEGMENTATIONS OBTAINED WITHOUT NOISE ADDED. WE CONCLUDE THAT THE PROPOSED METHOD IS MORE ROBUST W.R.T. IMAGE NOISE

		$\sigma = 0.05$	$\sigma = 0.1$	$\sigma = 0.25$
"cell"	plain seg.	0.7893	1.6444	4.2140
	plain seg. & [2]	0.4883	1.2027	3.7199
	proposed meth.	0.3775	1.0889	2.9605
"brain"	plain seg.	0.4973	0.8374	2.3266
	plain seg. & [2]	0.2348	0.5030	1.5585
	proposed meth.	0.2233	0.4252	1.0280
"LV"	plain seg.	1.1563	1.7501	5.2510
	plain seg. & [2]	0.7142	1.1622	4.0512
	proposed meth.	0.6121	0.9789	2.2050

Gaussian noise to the initializations of the brain tumor data of Fig. 5 (first row) and compared the results obtained with (i) the plain segmentation method, i.e., Chan-Vese data term minimized with Sobolev gradient method, (ii) the method proposed in [2] on the results of the plain segmentation, and

(iii) the proposed method applied to the noisy initializations. We varied the regularization parameter γ in the range $[0.5, 3]$ and found that the proposed method is robust w.r.t. to perturbed initializations, as the obtained average surface distances (w.r.t. to the segmentations obtained without noisy initializations) are below the standard deviation of the added noise, cf. Table I. Furthermore, the proposed method achieves slightly better results in terms of initialization robustness than the baseline method, i.e., plain segmentation combined with the regularization method of [2]. Moreover, we applied white Gaussian noise with different variances to three different image sequences and compared the results to the ones obtained from the original sequences. It becomes apparent in Table II that the proposed method is less susceptible to image noise than the baseline approach, i.e. a Sobolev active contour with a posteriori regularization.

In Fig. 6 we compare the proposed method with the video segmentation method of Grundmann et al. [7] when segmenting a basal slice of a left ventricle. For the method of [7], we have used the default parameters where we adapted the label range from 5 (minimum possible) to 10 labels. We observe that the method proposed in [7] erroneously segments parts of the neighboring blood pools as well, cf. Fig. 6. In contrast, the proposed method segments the left

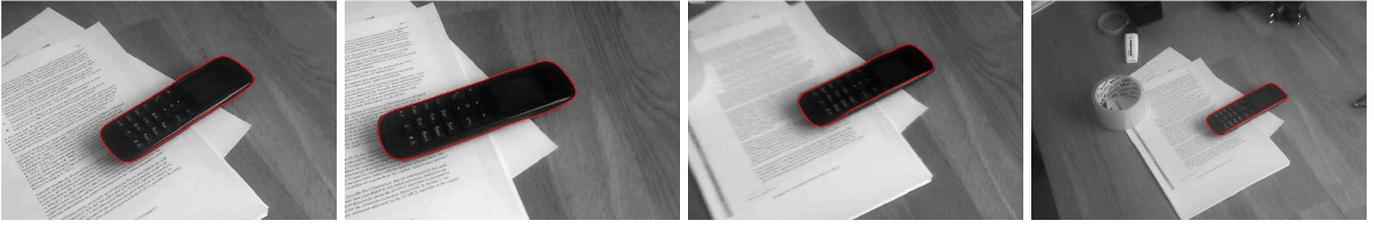


Fig. 7. Robustness to Scale Changes: Our method deals well with large scale changes of the tracked object ($\gamma = 0.5$).

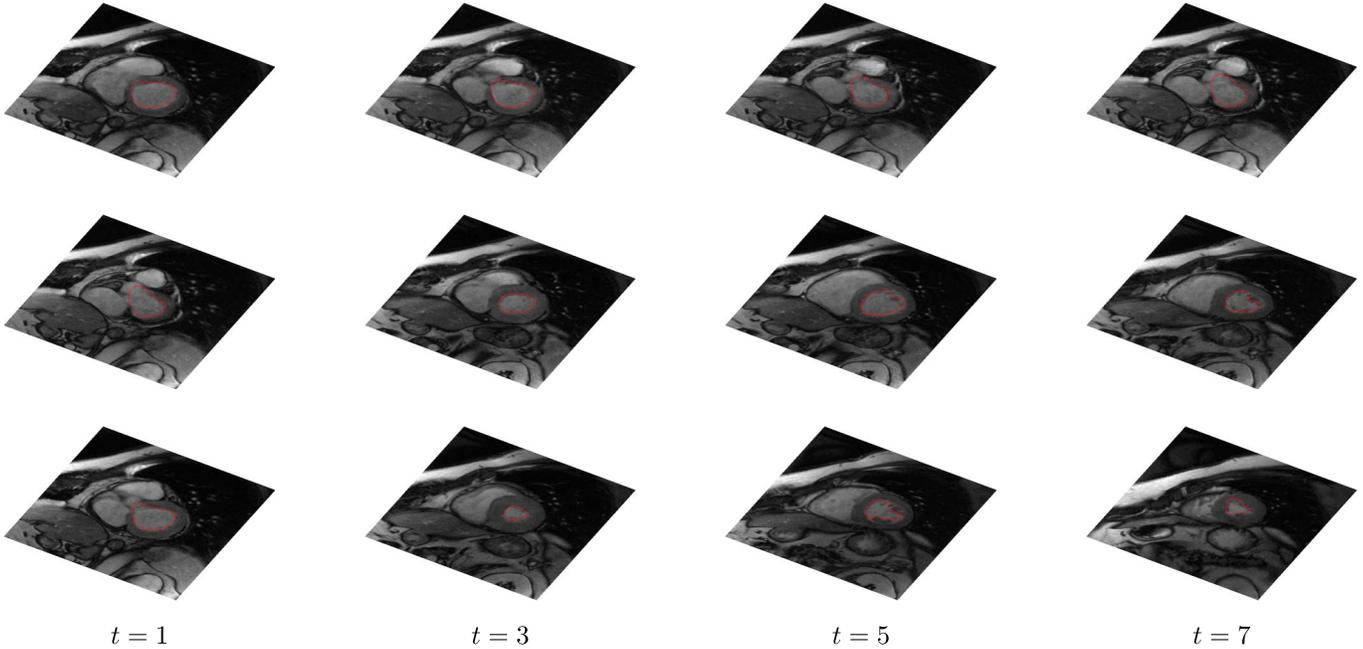


Fig. 8. Joint Segmentation and Regularization of Shape Fields: We segmented the left ventricular cavity from 3D+t MRI data (20 axial slices for 9 time steps; partial data already used in Fig. 6) using the proposed method. We consider this collection of ventricular boundaries as a shape field, because it enjoys a natural spatio-temporal ordering. We segmented and regularized this data set with the proposed method. An important aspect of using the proposed method for shape fields is that one can employ different regularization parameters for the spatial direction ($\gamma = 1$) and the temporal direction $\gamma = 0.1$.

ventricular cavity reliably. Moreover, the runtime of the proposed method (0.1 sec) is significantly lower than that of [7] (15 sec., only processing time without up- and download). Fig. 7 illustrates that the proposed method is robust w.r.t. scale changes of the segmented object. For all frames the number of boundary points has been kept fixed to 200. The runtime is 22 seconds for 480×854 pixels and 50 frames which corresponds to 2.3 frames/sec. We note that these numbers have been obtained using a non-optimized and non-parallelized implementation.

B. Further Examples and Extensions

As discussed in Sec. II-F, it is furthermore possible to extend the proposed approach to shape fields, e.g. collections of shapes which enjoy a spatial and temporal ordering. To demonstrate this possibility, we consider 3D+t cardiac MRI shown in Fig. 8. Note that the apical slice has to be chosen after the papillary muscles and before the cavity ends as suggested in [54]. This example also demonstrates the efficiency of the method: 1000 iterations of our single-threaded C++ implementation, which is partially based on

the Eigen library (eigen.tuxfamily.org), take less than one second on a laptop with 2,8 GHz Intel Core i7 processor.

IV. CONCLUSION

We presented the first Riemannian formulation for joint segmentation and TV regularization of shapes (shape fields and shape signals). Our shape representation employs a modified version of the classical Kendall shape space, which has been proven to be well-suited for the TV regularization of shape signals, see [2]. Using this shape representation for image segmentation is, however, not straightforward as this representation is translation- and scale-invariant. In order to address this problem, we proposed to minimize the functional in the modified Kendall shape space with a generalized forward backward algorithm augmented by a separate evolution of the non-shape characteristics position and scale. Furthermore, we stabilize both the shape and this separate evolution via Sobolev gradients. The resulting method is efficient as demonstrated by the experimental evaluation. The computation of the exponential and the inverse exponential maps as well as the computation of the Sobolev gradients are of linear complexity. Thus, each iteration step has linear complexity yielding a very

low runtime. The proposed method can be easily adapted to more sophisticated segmentation models such as a piecewise smooth model. Future research includes applications to cell tracking as well as the usage of other shape spaces.

REFERENCES

- [1] I. U. Rahman, I. Drori, V. C. Stodden, D. L. Donoho, and P. Schröder, "Multiscale representations for manifold-valued data," *Multiscale Model. Simul.*, vol. 4, no. 4, pp. 1201–1232, 2005.
- [2] M. Baust, L. Demaret, M. Storath, N. Navab, and A. Weinmann, "Total variation regularization of shape signals," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 2075–2083.
- [3] A. J. Yezzi and S. Soatto, "Deformation: Deforming motion, shape average and the joint registration and approximation of structures in images," *Int. J. Comput. Vis.*, vol. 53, no. 2, pp. 153–167, 2003.
- [4] M. Unger, T. Mauthner, T. Pock, and H. Bischof, "Tracking as segmentation of spatial-temporal volumes by anisotropic weighted TV," in *Energy Minimization Methods in Computer Vision and Pattern Recognition* (Lecture Notes in Computer Science), vol. 5681, D. Cremers, Y. Boykov, A. Blake, and F. R. Schmidt, Eds. Berlin, Germany: Springer, 2009, pp. 193–206.
- [5] J. Shi and J. Malik, "Motion segmentation and tracking using normalized cuts," in *Proc. 6th Int. Conf. Comput. Vis.*, 1998, pp. 1154–1160.
- [6] T. Tarabalka, G. Charpiat, L. Brucker, and B. H. Menze, "Spatio-temporal video segmentation with shape growth or shrinkage constraint," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 3829–3840, Sep. 2014.
- [7] M. Grundmann, V. Kwatra, M. Han, and I. Essa, "Efficient hierarchical graph-based video segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2141–2148.
- [8] T. R. Raviv, K. Van Leemput, W. M. Wells, and P. Golland, "Joint segmentation of image ensembles via latent atlases," in *Medical Image Computing and Computer-Assisted Intervention* (Lecture Notes in Computer Science), vol. 5761, G.-Z. Yang, D. Hawkes, D. Rueckert, A. Noble, and C. Taylor, Eds. Berlin, Germany: Springer, 2009, pp. 272–280.
- [9] Z. Xue, K. Wong, and S. Wong, "Joint registration and segmentation of serial lung CT images in microendoscopy molecular image-guided therapy," in *Medical Imaging and Augmented Reality* (Lecture Notes in Computer Science), vol. 5128, T. Dohi, I. Sakuma, and H. Liao, Eds. Berlin, Germany: Springer, 2008, pp. 12–20.
- [10] A. Papazoglou and V. Ferrari, "Fast object segmentation in unconstrained video," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1777–1784.
- [11] P. Ochs, J. Malik, and T. Brox, "Segmentation of moving objects by long term video analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 6, pp. 1187–1200, Jun. 2014.
- [12] G. Sundaramoorthi, A. Yezzi, and A. C. Mennucci, "Sobolev active contours," *Int. J. Comput. Vis.*, vol. 73, no. 3, pp. 345–366, 2007.
- [13] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.
- [14] J. W. Neuberger, *Sobolev Gradients and Differential Equations*. Berlin, Germany: Springer, 1997.
- [15] G. Charpiat, P. Maurel, J.-P. Pons, R. Keriven, and O. Faugeras, "Generalized gradients: Priors on minimization flows," *Int. J. Comput. Vis.*, vol. 73, no. 3, pp. 325–344, 2007.
- [16] G. Sundaramoorthi, A. Yezzi, and A. Mennucci, "Coarse-to-fine segmentation and tracking using Sobolev active contours," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 5, pp. 851–864, May 2008.
- [17] G. Sundaramoorthi, A. Yezzi, A. C. Mennucci, and G. Sapiro, "New possibilities with Sobolev active contours," *Int. J. Comput. Vis.*, vol. 84, no. 2, pp. 113–129, 2009.
- [18] D. Cremers, F. R. Schmidt, and F. Barthel, "Shape priors in variational image segmentation: Convexity, Lipschitz continuity and globally optimal solutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–6.
- [19] Y. Rathi, N. Vaswani, A. Tannenbaum, and A. Yezzi, "Tracking deforming objects using particle filtering for geometric active contours," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 8, pp. 1470–1475, Aug. 2007.
- [20] J. D. Jackson, A. J. Yezzi, and S. Soatto, "Dynamic shape and appearance modeling via moving and deforming layers," *Int. J. Comput. Vis.*, vol. 79, no. 1, pp. 71–84, 2008.
- [21] D. G. Kendall, "Shape manifolds, procrustean metrics, and complex projective spaces," *Bull. London Math. Soc.*, vol. 16, no. 2, pp. 81–121, 1984.
- [22] P. W. Michor and D. B. Mumford, "Riemannian geometries on spaces of plane curves," *J. Eur. Math. Soc.*, vol. 8, no. 1, pp. 1–48, 2006.
- [23] L. Younes, P. Michor, J. Shah, and D. Mumford, "A metric on shape spaces with explicit geodesics," *Rendiconti Lincei Matematica e Appl.*, 2008, doi: 10.4171/rlm/506.
- [24] P. W. Michor and D. Mumford, "An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach," *Appl. Comput. Harmon. Anal.*, vol. 23, no. 1, pp. 74–113, 2007.
- [25] A. Srivastava, E. Klassen, S. H. Joshi, and I. H. Jermyn, "Shape analysis of elastic curves in Euclidean spaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 7, pp. 1415–1428, Jul. 2011.
- [26] B. Wirth, L. Bar, M. Rumpf, and G. Sapiro, "A continuum mechanical approach to geodesics in shape space," *Int. J. Comput. Vis.*, vol. 93, no. 3, pp. 293–318, 2011.
- [27] M. Bauer, M. Bruveris, and P. W. Michor, "Overview of the geometries of shape spaces and diffeomorphism groups," *J. Math. Imag. Vis.*, vol. 50, nos. 1–2, pp. 60–97, 2014.
- [28] M. Bauer, M. Bruveris, S. Marsland, and P. W. Michor, "Constructing reparameterization invariant metrics on spaces of plane curves," *Differ. Geometry Appl.*, vol. 34, pp. 139–165, Jun. 2014.
- [29] J. Lellmann, E. Strelakovsky, S. Koetter, and D. Cremers, "Total variation regularization for functions with values in a manifold," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 2944–2951.
- [30] A. Weinmann, L. Demaret, and M. Storath, "Total variation regularization for manifold-valued data," *SIAM J. Imag. Sci.*, vol. 7, no. 4, pp. 2226–2257, 2014.
- [31] P. Grohs and M. Sprecher, "Total variation regularization by iteratively reweighted least squares on Hadamard spaces and the sphere," Seminar Appl. Math., ETH Zürich, Zürich, Switzerland, Tech. Rep. 2014-39, 2014.
- [32] M. Giaquinta and D. Mucci, "The BV-energy of maps into a manifold: Relaxation and density results," *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, vol. 5, no. 4, pp. 483–548, 2006.
- [33] M. Giaquinta and D. Mucci, "Maps of bounded variation with values into a manifold: Total variation and relaxed energy," *Pure Appl. Math. Quart.*, vol. 3, no. 2, pp. 513–538, 2007.
- [34] M. Storath, A. Weinmann, and M. Unser, "Exact algorithms for L^1 -TV regularization of real-valued or circle-valued signals," *SIAM J. Sci. Comput.*, vol. 38, no. 1, pp. A614–A630, 2016.
- [35] A. Weinmann, L. Demaret, and M. Storath, "Mumford–Shah and Potts regularization for manifold-valued data," *J. Math. Imag. Vis.*, vol. 55, no. 3, pp. 428–445, 2016.
- [36] S. Durrleman, X. Pennec, A. Trounev, G. Gerig, and N. Ayache, "Spatiotemporal atlas estimation for developmental delay detection in longitudinal datasets," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. I*, 2009, pp. 297–304.
- [37] T. P. Fletcher, "Geodesic regression and the theory of least squares on Riemannian manifolds," *Int. J. Comput. Vis.*, vol. 105, no. 2, pp. 171–185, 2013.
- [38] C. Samir, P.-A. Absil, A. Srivastava, and E. Klassen, "A gradient-descent method for curve fitting on Riemannian manifolds," *Found. Comput. Math.*, vol. 12, no. 1, pp. 49–73, 2012.
- [39] J. Su, I. L. Dryden, E. Klassen, H. Le, and A. Srivastava, "Fitting smoothing splines to time-indexed, noisy points on nonlinear manifolds," *Image Vis. Comput.*, vol. 30, nos. 6–7, pp. 428–442, 2012.
- [40] A. Srivastava, P. Turaga, and S. Kurtek, "On advances in differential-geometric approaches for 2D and 3D shape analyses and activity recognition," *Image Vis. Comput.*, vol. 30, nos. 6–7, pp. 398–416, 2012.
- [41] L. Alvarez, F. Guichard, P.-L. Lions, and J.-M. Morel, "Axioms and fundamental equations of image processing," *Arch. Rational Mech. Anal.*, vol. 123, no. 3, pp. 199–257, 1993.
- [42] B. Berkels, G. Linkmann, and M. Rumpf, "An $SL(2)$ invariant shape median," *J. Math. Imag. Vis.*, vol. 37, no. 2, pp. 85–97, 2010.
- [43] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.
- [44] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, no. 5, pp. 577–685, 1989.
- [45] H. Raguette, J. Fadili, and G. Peyré, "A generalized forward-backward splitting," *SIAM J. Imag. Sci.*, vol. 6, no. 3, pp. 1199–1226, 2013.
- [46] M. Baust, A. Weinmann, M. Wiecek, T. Lasser, M. Storath, and N. Navab, "Combined tensor fitting and TV regularization in diffusion tensor imaging based on a Riemannian manifold approach," *IEEE Trans. Med. Imag.*, to be published.
- [47] P. Radau. (2010). *Evaluation of Cardiac MR Segmentation*. [Online]. Available: <http://sourceforge.net/projects/cardiac-mr/files/>

- [48] M. Bačák, "Computing medians and means in Hadamard spaces," *SIAM J. Optim.*, vol. 24, no. 3, pp. 1542–1566, 2014.
- [49] D. P. Bertsekas, "Incremental proximal methods for large scale convex optimization," *Math. Program. B*, vol. 129, no. 2, pp. 163–195, 2011.
- [50] R. Bergmann, F. Laus, G. Steidl, and A. Weinmann, "Second order differences of cyclic data and applications in variational denoising," *SIAM J. Imag. Sci.*, vol. 7, no. 4, pp. 2916–2953, 2014.
- [51] B. H. Menze *et al.*, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015.
- [52] A. Chambolle, "Finite-differences discretizations of the Mumford–Shah functional," *ESAIM, Math. Model. Numer. Anal.*, vol. 33, no. 2, pp. 261–288, 1999.
- [53] M. Storath, A. Weinmann, J. Friel, and M. Unser, "Joint image reconstruction and segmentation using the Potts model," *Inverse Problems*, vol. 31, no. 2, p. 025003, 2015.
- [54] M. D. Cerqueira *et al.*, "Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. A statement for healthcare professionals from the cardiac imaging committee of the council on clinical cardiology of the American heart association," *Circulation*, vol. 105, no. 4, pp. 539–542, 2002.



Anca Stefanoiu received the B.Sc. degree in computer science from the Politehnica University of Timisoara, Romania, in 2013, and the M.Sc. degree in biomedical computing from the Technische Universität München, in 2016. In 2015, as part of her master's thesis, she was with the Institute of Computational Biology, Helmholtz Zentrum München. She joined Computer Aided Medical Procedures and Augmented Reality, Technische Universität München, in 2016. Her current work and interests

are towards multiscale and multimodal biomedical imaging, image formation and processing, direct/inverse problems, and variational optimization.



Andreas Weinmann received the Diploma (Hons.) degree in mathematics and computer science from the Technische Universität München, in 2006, and the Ph.D. (Hons.) degree from the Technische Universität Graz, in 2010. He is currently with the Mathematics Department, Helmholtz Center Munich, Technische Universität München, and with the Darmstadt University of Applied Sciences. His research interests are applied analysis and signal and image processing.



Martin Storath received the Diploma degree in mathematics, the Honors degree in technology management, and the Ph.D. degree in mathematics from the Technische Universität München, in 2008, 2009, and 2013, respectively. From 2010 to 2013, he was a Researcher with the Institute of Biomathematics, Helmholtz Zentrum München. From 2013 to 2016, he was a Post-Doctoral Researcher with the Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne, Switzerland. He currently holds a post-doctoral position with the Image Analysis and Learning Group, Universität Heidelberg. His research interests include signal and image processing, biomedical imaging, variational methods, and inverse problems.



Nassir Navab received the Ph.D. degree from Inria and the University of Paris XI, Paris, France. He spent two years with the MIT Media Laboratory, as a Post-Doctoral Fellow, before joining Siemens Corporate Research (SCR) in 1994. He is currently a Professor and the Director of the Laboratory for Computer Aided Medical Procedures, Technische Universität München, and with Johns Hopkins University, where he also has secondary faculty appointments with both affiliated Medical Schools. He has authored over 100 peer-reviewed scientific papers, fifteen awarded papers, 45 U.S. and over 50 international patents. His current research interests include medical augmented reality, computer-aided surgery, medical robotics, surgical data science, and medical image analysis. He is a member of the board of directors of the MICCAI Society and the Steering committee of the IEEE Symposium on Mixed and Augmented Reality. At SCR, he was a Distinguished Member and received the Siemens Inventor of the Year Award in 2001. In 2012, he was elected as a fellow of the MICCAI Society.



Maximilian Baust received the Ph.D. (Hons.) degree from the Technische Universität München, in 2012. He is currently a Post-Doctoral Fellow with the Laboratory for Computer Aided Medical Procedures, Technische Universität München. He was a Research Engineer, where he was involved in camera-based driver assistance systems for two years. He became a Post-Doctoral Fellow in 2013. His research interests include variational problems, image segmentation, processing of manifold-valued data, and machine learning.