Total Variation Regularization of Shape Signals

Maximilian Baust[1] Laurent Demaret[2] Martin Storath[3] Nassir Navab[1,4] Andreas Weinmann[2]

[1]Computer Aided Medical Procedures and Augmented Reality, Technische Universität München,

[2]Institute of Computational Biology, Helmholtz Zentrum München, [3]Biomedical Imaging Group, EPFL, Lausanne, [4]Johns Hopkins University, Baltimore.

[maximilian.baust,nassir.navab]@tum.de, [laurent.demaret, andreas.weinmann]@helmholtz-muenchen.de, martin.storath@epfl.ch

Abstract

This paper introduces the concept of shape signals, i.e., series of shapes which have a natural temporal or spatial ordering, as well as a variational formulation for the regularization of these signals. The proposed formulation can be seen as the shape-valued generalization of the Rudin-Osher-Fatemi (ROF) functional for intensity images. We derive a variant of the classical finite-dimensional representation of Kendall, but our framework is generic in the sense that it can be combined with any shape space. This representation allows for the explicit computation of geodesics and thus facilitates the efficient numerical treatment of the variational formulation by means of the cyclic proximal point algorithm. Similar to the ROF-functional, we demonstrate experimentally that ℓ_1-type penalties both for data fidelity term and regularizer perform best in regularizing shape signals. Finally, we show applications of our method to shape signals obtained from synthetic, photometric, and medical data sets.

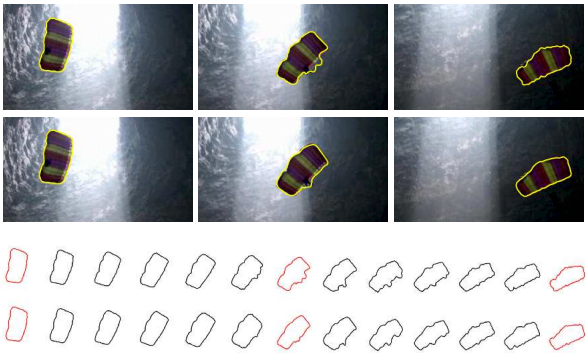


Figure 1: **Regularization of Shape Signals:** Objects segmented from video data enjoy a natural temporal ordering and thus form a *shape signal*. First row: Three frames of the "parachute" sequence from [28] segmented with the method proposed by [18]. Second row: Regularized shapes obtained with our method. Third and fourth row: Shape signal of the original segmentations (third row) and regularized shape signal (fourth row). Only a few shapes of this sequence are shown for better visibility. Shapes corresponding to the selected frames are highlighted in red.

1. Introduction

In this paper we wish to introduce the concept of *shape signals*, i.e., collections of shapes which appear in a spatial or temporal context. An obvious example is object tracking in video sequences where all shapes have a natural temporal ordering, cf. Fig. 1. Another example is organ segmentation from tomographic imaging modalities such as computed tomography or magnetic resonance imaging, where a three-dimensional organ can be segmented by obtaining its two-dimensional outlines from all containing slices. In both scenarios it is possible to arrange the obtained shapes in a one dimensional grid with (often) equidistant spacing. The resulting shape signal is a mapping from a discrete index, which corresponds to a time stamp or a slice number for instance, to some shape space.

1.1. Motivation

The idea of introducing shape signals is inspired by the observation of Rahman *et al*. [19], who noticed that many signals in science and engineering are actually *manifold-valued* signals as well as some recent efforts of total variation (TV) regularization of such data, cf. Lellmann *et al*. [14] and Weinmann *et al*. [29]. A typical example is pose tracking data, e.g., acquired with an optical tracking system, which can be represented as a series of rigid transformation matrices acquired at equally spaced points in time. While these approaches focus, however, on rather low dimensional manifolds, e.g. $SE(3)$ which is useful for regularizing pose data, we would like to go beyond these applications and consider *shapes* and *shape spaces* which are high-dimensional by nature – typical polygonal shape rep-

978-1-4673-6964-0/15/$31.00 ©2015 IEEE

resentations used in this paper comprise hundreds of boundary points. In general, there a two main approaches for obtaining a suitable shape representation or shape space:

1. The continuous approach: A shape can for instance be considered as an element of the manifold of simple closed smooth and unparametrized curves in $\mathbb{R}^2$ or as an element of the orbifold of immersions from S^1 to $\mathbb{R}^2$ modulo all diffeomorphisms of S^1, cf. Michor and Mumford [17] or Bauer *et al*. [2].

2. The discrete approach: A shape can be interpreted as a simple and closed n-gon in $\mathbb{R}^2$ or the complex plane $\mathbb{C}$, cf. Kendall [13].

In this paper, we focus on the discrete approach – in particular the one of Kendall [13] – for mainly two reasons: Firstly, algorithms for processing manifold-valued data, require the implementation of the exponential as well as the inverse exponential mapping of the respective manifold, cf. Sec. 3. These mappings can be implemented very efficiently in the case of Kendall shapes. Secondly, Kendall shapes are *only* invariant w.r.t. translations, rotations, and global scalings. Although a lot of research has been carried out in regard to shape representations which are invariant to large classes of transformations, e.g., affine transformations, we will demonstrate in Sec. 5.2 that ensuring too many invariance properties is not always desired in the context of shape regularization. The observation that too many invariances can be disadvantageous is not at all obvious and it even motivated us to propose so-called *oriented Kendall shapes* which are only invariant w.r.t. translations and scalings, cf. Sec. 4.3.

2. Related Work

We discuss related work concerning TV regularization for manifold-valued data and comment on the connections to two other important fields of research: active contours and continuous shape spaces.

2.1. TV Regularization for Manifold-valued Data

From a theoretical point of view, TV minimization for manifold-valued data has been considered by Giaquinta and Mucci [9, 10]. Algorithms for TV regularization for manifold-valued data have been developed by Lellmann *et al*. [14] and by Weinmann *et al*. [29]. Applications considered in [14] are the one and two-dimensional sphere as well as the three-dimensional group of rotations. The algorithm in [14] is based on rewriting the problem as a labeling problem with the label space consisting of a discretization of the manifold. Unfortunately, the label space grows rapidly with the dimension of the manifold which makes application to higher dimensional manifolds, such as shape spaces, hard. In contrast to this, the algorithms described by Weinmann *et al*. [29], i.e., the cyclic and the parallel proximal point algorithm, do not require such a discretization. Furthermore, the computational times do not increase more than linearly with the dimension of the manifold. Application areas in [29] also include the six-dimensional space of diffusion tensor data. It should be noted that there are also some works on fitting curves on manifolds which share the same goal but essentially consider ℓ_2-regularized scenarios, e.g., [8, 21, 24, 23].

2.2. Active Contours

The idea of active-contour-based tracking has already been described in the seminal work of Kass *et al*. [12]. It is based on the assumption that the shape of the tracked object varies only gently between two consecutive frames. Thus, the segmentation from one frame can be used as an initialization for the next frame, which greatly benefits the non-convexity of active contours. Since the original work of Kass *et al*. a lot of improvements have been proposed which help to make this method more robust w.r.t. large deformations, occlusions, background clutter, etc. As it is beyond the scope of this paper to discuss all of them we refer the interested reader to the exhaustive overview of Cremers *et al*. [7]. However, we would like to mention the recent works on Sobolev-type active contours, cf. Charpiat *et al*. [6] and Sundaramoorthi *et al*. [26, 27, 25], which have in common that they either implicitly or explicitly rely on shape metrics, i.e., some notion of distance between shapes, in order to constrain the segmentation. As a consequence, we wish to emphasize that the proposed method assumes that a shape signal has already been obtained and it can thus be combined with *any* method for segmentation-based tracking or video segmentation. Furthermore, even if active-contour-based segmentation methods are augmented by sophisticated shape metrics, the regularity of the obtained segmentation is not comparable to the one obtained by our approach. The most obvious reason for this fact is that segmentation-based tracking is an unidirectional process where previous segmentations cannot be regularized by subsequent ones. We will actually see in Sec. 5.4 that our method nicely complements Sobolev-type active contours by further regularizing their results.

There are of course many more approaches for segmentation-based tracking and video segmentation, especially those which are based on graphical models. The reason why we only consider approaches related to active contours as related is that this area of research is closely related to the one of shape spaces and shape metrics which are an essential prerequisite for the proposed framework.

2.3. Continuous Shape Spaces and Metrics

Shape spaces and shape metrics have a long history in computer vision as well as medical image analysis. Es-

pecially in the area of continuous approaches, e.g. metrics for the space of immersions from S^1 to $\mathbb{R}^2$ modulo all diffeomorphisms of S^1, a tremendous progress has been achieved in the last decade. Noteworthy examples are the paper of Michor and Mumford on the deficiency of the L^2 metric [16], the one of Younes *et al*. on explicit geodesics [30], the one of Srivastava *et al*. on the square root velocity transform [22], and finally the one of Bauer *et al*. [1] on reparametrization invariant metrics. An exhaustive discussion of all works in this area is beyond the scope of this work and we thus refer the interested reader to the excellent recent overviews of Michor and Mumford [17] as well as Bauer *et al*. [2]. However, we would like to emphasize that the proposed framework can be combined with *any* shape space as long as there exists the possibly of computing geodesics between to shapes.

3. Regularization of Shape Signals

In this section we will describe a general framework for regularizing shape signals, which is similar to the one proposed by Rudin, Osher, and Fatemi in [20] as well as the one considered by Weinmann *et al*. [29] and Lellmann *et al*. [14]. It is important to notice that this formulation does not depend on the particular choice of the shape space.

3.1. Notation

Let $\mathcal{M}$ denote a suitable Riemannian shape manifold, which does not need to be specified for the moment. For an element $\mathbf{a} \in \mathcal{M}$ we denote the tangent space at $\mathbf{a}$ by $T_{\mathbf{a}}\mathcal{M}$. We further denote the exponential map at $\mathbf{a}$ by

$$\exp_{\mathbf{a}} : \begin{cases} T_{\mathbf{a}}\mathcal{M} & \to \mathcal{M}, \\ v & \mapsto \exp_{\mathbf{a}}(v) \end{cases} \tag{1}$$

and the inverse exponential map at $\mathbf{a}$ by

$$\log_{\mathbf{a}} : \begin{cases} \mathcal{M} & \to T_{\mathbf{a}}\mathcal{M}, \\ \mathbf{b} & \mapsto v = \log_{\mathbf{a}}(b) \end{cases} \tag{2}$$

where $\log_{\mathbf{a}}(b)$ is the tangent vector in $T_{\mathbf{a}}\mathcal{M}$ such that

$$\exp_{\mathbf{a}}(\log_{\mathbf{a}}(\mathbf{b})) = \mathbf{b}. \tag{3}$$

A natural distance between two points $\mathbf{a}, \mathbf{b} \in \mathcal{M}$ is given by the length of the geodesic joining $\mathbf{a}$ and $\mathbf{b}$. This is precisely the length of the tangent vector $\log_{\mathbf{a}}(\mathbf{b})$:

$$d(\mathbf{a}, \mathbf{b}) = \left\| \log_{\mathbf{a}}(\mathbf{b}) \right\|_{\mathbf{a}}, \tag{4}$$

where $\| \cdot \|_a$ is the length induced by the Riemannian metric in the tangent space of $\mathbf{a}$. The index $\mathbf{a}$ indicates that the metric may depend on $\mathbf{a}$, which is the case for Sobolev-type metrics for instance, cf. [6, 27, 26, 25, 1]. In general, it is hard to find explicit formulas for computing geodesics in case of such metrics. However, we use an embedded representation, cf. Sec. 4, which helps us to circumvent related problems.

3.2. Problem Formulation

A shape(-valued) signal with $k \in \mathbb{N}$ entries is a vector $\mathbf{f} = (\mathbf{f}_1, \dots, \mathbf{f}_k) \in \mathcal{M}^k$. We assume that $\mathbf{f}$ is given, e.g., as a result of a segmentation algorithm, and we wish to find another signal $\mathbf{x}$ such that the functional

$$E(\mathbf{x}) = D(\mathbf{x}, \mathbf{f}) + \alpha R(\mathbf{x}), \quad \alpha > 0, \tag{5}$$

is minimal. While $D(\mathbf{x}, \mathbf{f})$ is a data fidelity term which penalizes the deviation from the original signal $\mathbf{f}$, $R(\mathbf{x})$ is a regularizer penalizing large variations or jumps in $\mathbf{x}$. As a data fidelity term, we consider

$$D(\mathbf{x}, \mathbf{f}) = \sum_{i=1}^{k} (h \circ d)(\mathbf{x}_i, \mathbf{f}_i), \tag{6}$$

where h is one of the following functions: $h(s) = s$ which leads to an ℓ_1-type penalization, $h(s) = s^2/2$ which leads to an ℓ_2-type penalization, and

$$h(s) = \begin{cases} s^2, & s < 1/\sqrt{2}, \\ \sqrt{2}s - 1/2, & \text{otherwise}, \end{cases} \tag{7}$$

which yields the manifold-valued equivalent of the well-known Huber-norm [11] – a differentiable compromise between the ℓ_1-norm and the ℓ_2-norm. Similar to D we consider the regularizer

$$R(\mathbf{x}) = \sum_{i=1}^{k-1} (h \circ d)(\mathbf{x}_i, \mathbf{x}_{i+1}). \tag{8}$$

As $d(\mathbf{x}_i, \mathbf{x}_{i+1})$ can be considered as a manifold-valued forward difference, R can be interpreted as a first order approximation of the classical Tikhonov regularizer, in case of $h(s) = s^2/2$, or the total variation, in case of $h(s) = s$, respectively. In case of (7), R can be regarded as a shape-valued differentiable approximation of the total variation regularizer, which can be used to avoid the staircasing problem associated with total variation denoising, cf. Chambolle and Pock [5].

3.3. Numerical Solution

Minimizing the functional in (5) can be achieved by a cyclic proximal point algorithm, cf. Alg. 1. As demonstrated by Weinmann *et al*. in [29], this algorithm is particularly suited for manifold-valued total variation regularized problems, because all proximal mappings can be implemented via computing points on geodesics, i.e.,

$$[\mathbf{a}, \mathbf{b}]_t = \exp_{\mathbf{a}}(t \log_{\mathbf{a}}(\mathbf{b})), \tag{9}$$

Algorithm 1: Cyclic proximal point algorithm for solving the functional (5).

input : Signal $\mathbf{f}$, parameter α, number of steps l
output: Signal $\mathbf{x}$ (solution of (5))

$\mathbf{x} \leftarrow \mathbf{f}$;
for $j \leftarrow 1$ **to** l **do**
  //compute rel. parameter;
  $\lambda_j \leftarrow$ `CompLambda`(j)
  //proximal mapping of D;
  for $i \leftarrow 1$ **to** k **do**
    $t \leftarrow$ `GeoLengthData`$(\lambda_j, \mathbf{x}_i, \mathbf{f}_i)$
    $\mathbf{x}_i \leftarrow [\mathbf{x}_i, \mathbf{f}_i]_t$;
  end
  //proximal mapping of R;
  for $i \leftarrow 1$ **to** $k-1$ **do**
    $t \leftarrow$ `GeoLengthData`$(\lambda_j, \alpha, \mathbf{x}_i, \mathbf{x}_{i+1})$
    $\hat{\mathbf{x}}_i \leftarrow [\mathbf{x}_i, \mathbf{x}_{i+1}]_t$;
    $\hat{\mathbf{x}}_{i+1} \leftarrow [\mathbf{x}_{i+1}, \mathbf{x}_i]_t$;
    $\mathbf{x}_i \leftarrow \hat{\mathbf{x}}_i$;
    $\mathbf{x}_{i+1} \leftarrow \hat{\mathbf{x}}_{i+1}$;
  end
end

for $\mathbf{a}, \mathbf{b} \in \mathcal{M}$ and an appropriately chosen $t \in [0, 1]$. For a detailed derivation as well as an analysis of this algorithm we refer the reader to [29].

We would like to emphasize that this algorithm can be instantiated for any shape space by providing implementations of the corresponding exponential mapping and its inverse. The functions `GeoLengthData`, and `GeoLengthReg` calculate how far to move on the geodesic under consideration; `GeoLengthData` does so for the data term whereas `GeoLengthReg` does so for the regularizer, cf. Tab. 1. The function `CompLambda` provides a square-summable (but not summable) sequence of relaxation parameters, cf. [29]. In all our experiments we have chosen

$$\lambda_j = 3j^{-(0.95+\frac{1}{2}j^{-0.18})}, \quad j = 1, \dots, l, \tag{10}$$

where $l \in \mathbb{N}$ is the number of iteration steps.

4. Shape Representations

As mentioned before, we decided to use the shape representation introduced by Kendall [13], because the exponential as well as the inverse exponential mapping can be implemented very efficiently. In the following we will briefly fix our notation for this section, review the classical representation proposed by Kendall [13], and derive a non-rotation invariant analogue of it. We will see in Sec. 5 that this novel representation has some advantages over the classical representation – at least for certain applications.

	`GeoLengthData`
ℓ_2	$\lambda/(1+\lambda)$
ℓ_1	$\min(\lambda/d, 1)$
Huber	$\begin{cases} 2\lambda/(1+2\lambda), & d < (1+2\lambda)/\sqrt{2}, \\ \min(\sqrt{2}\lambda/d, 1), & else. \end{cases}$
	`GeoLengthReg`
ℓ_2	$\lambda\alpha/(1+2\alpha\lambda)$
ℓ_1	$\min(\lambda/d, \frac{1}{2})$
Huber	$\begin{cases} 2\lambda/(1+4\lambda), & d < (1+4\lambda)/\sqrt{2}, \\ \min(\sqrt{2}\lambda/d, \frac{1}{2}), & else. \end{cases}$

Table 1: **Geodesic Path Lengths:** Depending on the chosen penalization the functions `GeoLengthData` and `GeoLengthReg` return different geodesic path lengths. The length of the geodesic, cf. (4), is denoted by d.

4.1. Notation

We consider polygonal discretizations of simple planar shapes, i.e., two-dimensional closed curves which do not intersect themselves. The result of such a discretization is a simple n-gon which can be represented by a complex vector

$$z = (z_1, \dots, z_n) \in \mathbb{C}^n, \tag{11}$$

where each entry z_i encodes the coordinates of one boundary point with its real and imaginary part. We will assume that this representation is already normalized w.r.t. translation, i.e.

$$\sum_{i=1}^{n} z_i = 0 \in \mathbb{C}. \tag{12}$$

In order to position the regularized shapes correctly, this mean value (translation vector) has to be stored for all shapes in the signal before the signal can be regularized using the presented framework.

It is also possible to define a real-valued representation of z via identifying $\mathbb{C}^n$ with $\mathbb{R}^{2n}$, i.e.,

$$x = (x_1^1, x_1^2, x_2^1, x_2^2, \dots, x_n^1, x_n^2) \in \mathbb{R}^{2n}, \tag{13}$$

where

$$x_i^1 = \Re(z_i) \quad \text{and} \quad x_i^2 = \Im(z_i). \tag{14}$$

Similar to the complex case, we will assume that x is already normalized w.r.t. translation.

4.2. Classical Kendall Shapes

We will briefly recall the original representation of Kendall [13]. For a concise description of this representation we recommend the recent paper of Fletcher on geodesic regression [8], which also forms the basis for the following derivations.

Kendall employs the complex shape representation in a particularly elegant way. The first step is to notice that by normalizing z w.r.t. to translation we restrict all shapes to the $(n-1)$-dimensional subspace

$$V_{n-1} = \{z \in \mathbb{C}^n : \sum_{i=1}^{n} z_i = 0\} \subset \mathbb{C}^n, \qquad (15)$$

which can itself be identified with $\mathbb{C}^{n-1}$. Roughly speaking, by normalizing w.r.t. translation we are removing *one complex degree of freedom*. Next, we notice that a shape $z \in V_{n-1}$ can be scaled by a factor $s > 0$ and rotated by an angle $\theta \in [0, 2\pi)$ by multiplying all complex components z_i with the complex number $w = s\exp(i\theta) = s\cos(\theta) + is\sin(\theta)$. Consequently, all shapes z which are equivalent w.r.t. translation, rotation, and scaling lie on the complex line

$$L_z = \{w \cdot z : w \in \mathbb{C}\backslash\{0\}\}. \qquad (16)$$

In other words, L_z is the equivalence class of all shapes which are equivalent w.r.t. rigid transformations and scalings. The set of all these equivalence classes can now be identified with the complex projective space $\mathbb{C}P^{n-2}$ or, more intuitively, the complex unit sphere $S_{\mathbb{C}}^{n-2}$ (with antipodal points identified). This means that by enforcing rotation and scale invariance we are removing *another complex degree of freedom*. As a consequence, the exponential mapping and the inverse exponential mapping are given by the respective mappings of $S_{\mathbb{C}}^{n-2}$, i.e.,

$$\exp_z(v) = \cos(\phi) \cdot z + \frac{\|z\|\sin(\phi)}{\phi} \cdot v, \phi = \|v\| \qquad (17)$$

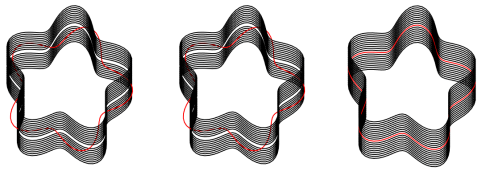
and

$$\log_z(y) = \phi \cdot \frac{y - \Pi_z(y)}{\|y - \Pi_z(y)\|}, \phi = \arccos(\frac{|\langle z, y\rangle|}{\|z\|\,\|y\|}), \qquad (18)$$

where $\Pi_z(y) = z \cdot \langle z, y\rangle / \|z\|^2$ denotes the projection of y onto z. It is important to notice that $\langle \cdot, \cdot \rangle$ denotes the complex scalar product, i.e.,

$$\langle z, y\rangle = \sum_{i=1}^{n} z_i \overline{y_i}, \qquad (19)$$

where $\overline{\cdot}$ denotes the complex conjugation, and $\|\cdot\|$ is the norm induced by the complex scalar product.



(a) (b) (c)

Figure 2: **Pure Rotational Perturbation:** Perturbed shape signal (a), regularized signal using the classical Kendall shape space (b), and regularized signal using the oriented Kendall shape space (c).

4.3. Oriented Kendall Shapes

Based on the considerations from Sec. 4.2, we derive a shape representation which is *not* rotationally invariant. We term this representation *oriented Kendall shapes*. At first, we notice that by normalizing x w.r.t. translation, cf. (13), we are removing *two real degrees of freedom*. Thus, the shape representation is restricted to the real subspace

$$V_{2n-2} = \{x \in \mathbb{R}^{2n} : \sum_{i=1}^{2n} x_i = 0\} \subset \mathbb{R}^{2n}. \qquad (20)$$

Next, we notice that a shape $x \in V_{2n-2}$ can be scaled by multiplying all real components x_i with a real number $s \neq 0$. Consequently, all shapes x which are equivalent w.r.t. translation and scaling lie on the real line

$$L_x = \{s \cdot x : s \in \mathbb{R}\backslash\{0\}\}. \qquad (21)$$

In other words, L_x is the equivalence class of all shapes which are equivalent w.r.t. translations and scalings. The set of all these equivalence classes can now be identified with the real projective space $\mathbb{R}P^{2n-3}$ or, more intuitively, the real unit sphere $S_{\mathbb{R}}^{2n-3}$ (again with antipodal points identified). This means that by enforcing scale invariance we are removing *another real degree of freedom*. As a consequence, the exponential mapping and the inverse exponential mapping are given by the respective mappings of $S_{\mathbb{R}}^{n-3}$, i.e., formulas (17) and (18) but this time with the real-valued scalar product $\langle x, y\rangle$ as well as its induced norm. To put it in a nutshell: By simply exchanging the scalar product for the computation of the exponential and the inverse exponential mappings we can switch between a rotationally invariant and a non-rotationally invariant representation. Moreover, all mappings can be implemented very efficiently as they only require basic linear algebra subroutines (BLAS).

5. Experiments and Discussion

In order to demonstrate the potential of our method we performed experiments using synthetic shape signals as well as shape signals obtained from real imaging data.

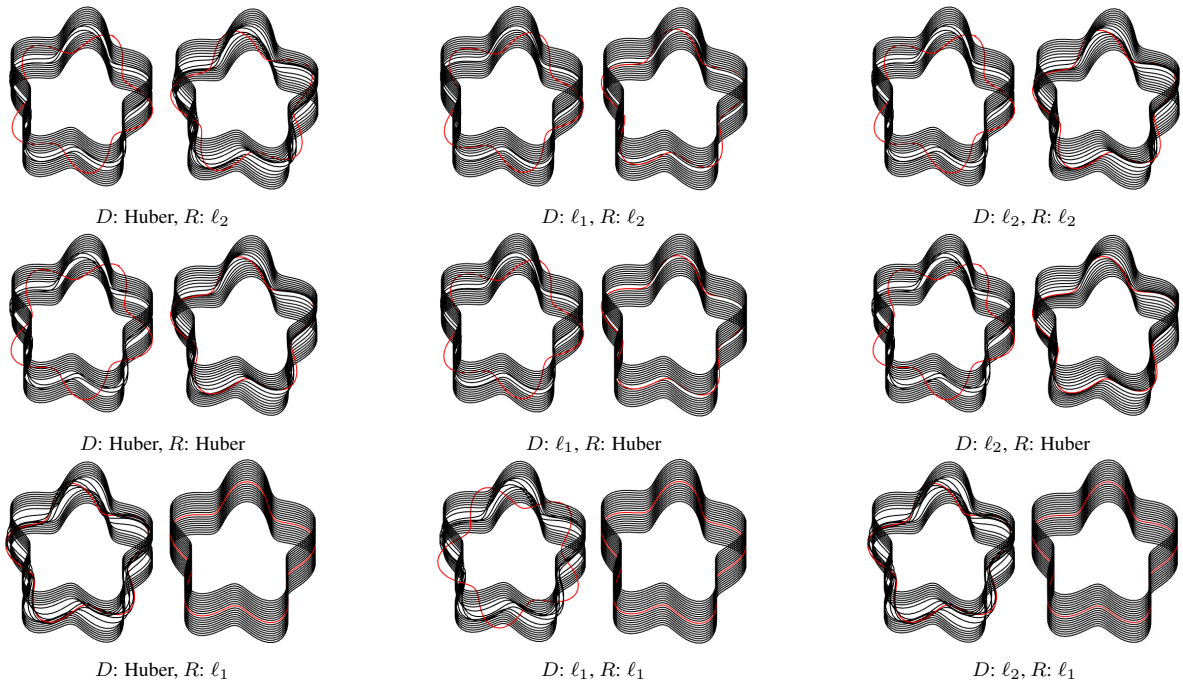


Figure 3: **Comparison of Different Shape Spaces and Penalties:** The red shape was perturbed by a rotation and barely visible deformation of one coordinate. All experiments with the classical Kendall shape space are on the left of each pair, i.e., columns 1, 3, 5, and all experiments with the proposed shape space are shown on the right of each pair, i.e., columns 2, 4, and 6. Note that the proposed shape space in conjunction with an ℓ_1 penalty for the regularizer performs best in reconstructing the original shape signal.

5.1. Experimental Setup and Parameter Choice

All experiments have been performed on a Mac Book Pro Retina (2013) with an Intel Core i7-4850HQ CPU (2,30 GHz), 16GB of RAM, and Windows 7. The algorithm was implemented in C++, compiled with the Visual Studio 2012 compiler, and controlled via MATLAB mex-functions. If not specified differently, we used 1000 iterations per experiment. The maximum processing time for all of the presented experiments was significantly below 2 seconds demonstrating the performance of our method. Thus, our algorithm can easily be used in an interactive scenario. We found that the parameter n was easy to adjust and in general we preferred a conservative strategy of choosing rather too many than too few discretization points – typical values for n where between 100 and 360.

5.2. Synthetic Experiments

We performed a series of synthetic experiments in order to demonstrate the strengths and weaknesses of the classical as well the oriented Kendall shapes in the context of our framework. Therefore, we generated an artificial shape signal consisting of 20 copies of the same shape. We set $\alpha = 5.0$ and discretized all shapes with 100 equally spaced boundary points. Next we perturbed the 10-th shape and regularized the signal with different parameter choices. As expected, the classical Kendall shape space is "blind" w.r.t. a pure rotational perturbation which is shown in Fig. 2. In this experiment we used ℓ_1 penalties, but the classical Kendall shape space would not behave differently in case of another penalty. In contrast to this, the original signal can be recovered by the proposed shape space. Next, we perturbed the 10-th shape not only by a rotation, but also by a tiny and barely visible deformation of less than 2% of the first coordinate. In Fig. 3 all results of this experiment are displayed.

From these results, we draw the following conclusions. Firstly, although the deformational perturbation is barely visible and changes the shape rather insignificantly, the metric of the classical Kendall shape space regards the perturbed shape as being different to the other ones. As a consequence, our algorithm falsely modifies the neighboring shapes. Secondly, also the metric of the proposed shape space regards the perturbed shape as being different. However, the regularized shapes come significantly closer to the original signal in case of the oriented shape space. Thirdly,

Figure 4: **Video Segmentation Example:** We segmented 48 frames of the "monkey" sequence from the SegTrack database of Tsai *et al*. [28]. First row: Five frames with the segmentations obtained with the method of Papazoglou and Ferrari [18]. Second row: The same frames with segmentations regularized with our algorithm. Third row: Shape signal consisting of the original segmentations, where only every third shape is displayed for better visibility. The shapes corresponding to the frames in the upper two rows are highlighted in red. Fourth row: Regularized shape signal with the same shapes highlighted.

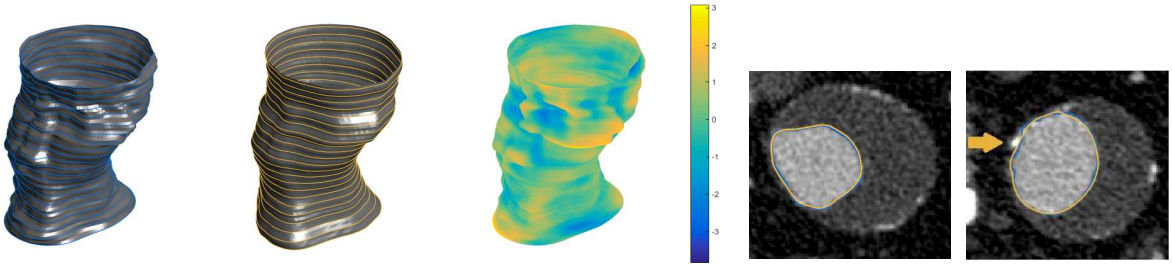
the total variation regularization performs best and almost independently of the data term at reconstructing the original signal. One might for sure argue that in case of a purely rigid object motions the original Kendall shape space might yield better results, because it does not try to correct a rotation by a deformation as it is done by the proposed shape space. However, the experiment in Fig. 3 shows the algorithm for obtaining the shape signal has to be very accurate. Moreover, the rigid body motion has to be perfectly parallel to the image plane and the intrinsic camera calibration needs to be very accurate, too. We strongly believe that these assumptions do hardly apply in practice. As a consequence, we recommend to use the proposed shape space in connection with a total variation regularization for the regularization of shape signals, in particular if they are obtained from projective imaging modalities such as video cameras.

5.3. Video Segmentation

A natural application of our method is the processing of shape signals obtained from video segmentation algorithms, especially in case of low resolution and low quality video data. Thus, we applied the recently proposed video segmentation algorithm of Papazoglou and Ferrari [18] to the "monkey"' sequence, cf. Fig. 4, of the SegTrack database[1] of Tsai *et al*. [28]. We selected this sequence, because it is of low quality, suffers from clearly visible compression artifacts, and the motion of the segmented monkey is very complex. We used the algorithm of Papazoglou and Ferrari with the standard preferences and in combination with the method of Brox and Malik [4] for optical flow estimation as well as the Turbopixels of Levinshtein *et al*. [15] for superpixel generation. Then, we extracted shape contours with 200 equally spaced boundary points per segmented frame and processed this signal with our algorithm. We used an ℓ_1 penalty both for the data term and the regularizer and chose a moderate regularization, i.e., $\alpha = 1.0$. Processing 48 frames of this sequence took 0.53 seconds.

It can be observed that the segmentation boundaries are significantly regularized without deviating too much from the original segmentation. In case of very fast motions the low frame rate of the video is sometimes causing problems, because neighboring shapes can be very different. This leads to a slight over-regularization which can be observed in the second and third frame in the second row of Fig. 4 (frames 19 and 25 in the original video sequence). However, frames four and five in the second of Fig. 4 (frames 40 and 58 in the original video sequence) clearly reveal that our method has in general no problem with concavities of the object boundary. Furthermore, we performed a sensitivity analysis of the proposed method w.r.t. to perturbations of the input shape signals of the 'monkey' sequence, cf. Tab. 2. We performed 1500 runs with Gaussian noise $N(0, \sigma)$ for varying levels of σ on the input shape coordinates of the 'monkey' sequence and evaluated the results w.r.t. the average surface distance (ASD) between the result obtained without noise and the one computed with noise added accumulated for the whole signal. These experiments show that our method a very robust w.r.t. perturbations of the initial shape data. Furthermore, we performed another experiment

[1] cpl.cc.gatech.edu/projects/SegTrack/



(a) original (b) regularized (c) surface distance (voxel) (d) slice 10 (e) slice 57

Figure 5: **Geometry Processing Example:** We applied our algorithm to a segmentation of the lumen of the abdominal part of a human aorta. The model consists of 68 CTA slices segmented with the method of Baust *et al*. [3]. The original model is shown in (a) with every second shape highlighted in blue. The regularized signal is shown in (b) with every second contour highlighted in yellow. Our method successfully removes little cusps and concavities of the original contours, where we colorized the original segmentation with the signed surface distance (in voxel) to the regularized model (c). Two exemplary contours are shown in (d) and (e), where the little arrow indicates that our method is capable of removing small spurious segmentation artifacts due to neighboring calcifications.

using the same sequence and found that α is relatively easy to tune, because the dependency of the result (in terms of ASD) on the choice of α is 0.00, 0.01, 2.09, 8.66, and 15.49 pixel for $\alpha = 0.01$, $\alpha = 0.1$, $\alpha = 1$, $\alpha = 5$, and $\alpha = 10$, respectively.

5.4. Geometry Processing

Besides processing signals obtained from video segmentation our method is also suitable for geometry processing applications. A typical scenario is the slice-wise segmentation of organs, e.g., vasculature acquired with computed tomography angiography (CTA). In order to demonstrate the applicability to such cases, we consider the segmentation of the abdominal part of the aorta from computed tomography angiography, cf. Fig. 5. The contrasted lumen of the aorta, cf. Fig. 5 (d) and (c), was segmented with the method of Baust *et al*. [3]. The segmentation boundaries were discretized with 360 equally spaced points and our algorithm regularized the whole signal consisting of 68 shapes in 1.35 seconds, where we chose $\alpha = 15.0$ as well as ℓ_1 penalties for data term and regularizer. As it depicted in Fig. 5 (a) and (b), our algorithm successfully regularizes the segmentation of the aortic lumen. Thereby, our algorithm is particularly useful in removing little cusps and concavities which is shown by Fig. 5 (c) where we colorized the original segmentation with the signed surface distance between the original and the regularized signal. These cusps correspond to erroneously segmented calcifications in the aortic wall, cf. Fig. 5 (e). Since our algorithm does, however, not alter the segmentation in an unreasonable way, which is shown in Fig. 5 (d) and (e), it is perfectly suited for processing geometric models which shall later be used in biomechanical simulations.

α	$\sigma = 0.5$	$\sigma = 1$	$\sigma = 2$	$\sigma = 5$	$\sigma = 10$
0.01	0.6	1.3	2.5	6.3	12.5
0.1	0.6	1.3	2.5	6.3	12.5
1	0.4	0.8	1.5	3.5	7.3
5	0.2	0.4	1.0	3.1	7.6
10	0.2	0.3	0.7	2.1	5.3

Table 2: **Sensitivity Analysis:** The proposed method is very robust w.r.t. perturbations of the input contours, cf. Sec.5.3. All values (except for α) are in pixel.

6. Conclusion

In this paper we introduced the concept of shape signals, i.e., collections of shapes which enjoy a temporal or spatial ordering. We also presented an algorithm for regularizing such shape signals. Moreover, we derived a non-rotationally invariant analogue of the shape space proposed by Kendall [13] which is better suited for the considered scenario of shape signal regularization. Finally, we demonstrated the advantages of the proposed approach, particularly in the case of total variation regularization, using synthetic and real world data.

Acknowledgments Martin Storath is supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 267439. Andreas Weinmann is supported by the Helmholtz Association (young investigator group VH-NG-526).

References

[1] M. Bauer, M. Bruveris, S. Marsland, and P. W. Michor. Constructing reparameterization invariant metrics on spaces of plane curves. *Differential Geometry and its Applications*, 34(0):139 – 165, 2014. 3

[2] M. Bauer, M. Bruveris, and P. W. Michor. Overview of the geometries of shape spaces and diffeomorphism groups. *Journal of Mathematical Imaging and Vision*, 50(1-2):60–97, 2014. 2, 3

[3] M. Baust, A. Yezzi, G. Unal, and N. Navab. A Sobolev-type metric for polar active contours. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011. 8

[4] T. Brox and J. Malik. Large displacement optical flow: Descriptor matching in variational motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):500–513, 2011. 7

[5] A. Chambolle and T. 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. 3

[6] G. Charpiat, P. Maurel, J.-P. Pons, R. Keriven, and O. Faugeras. Generalized gradients: Priors on minimization flows. *International Journal on Computer Vision*, 73(3):325–344, 2007. 2, 3

[7] D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. *International Journal of Computer Vision*, 72(2):195–215, 2007. 2

[8] T. P. Fletcher. Geodesic regression and the theory of least squares on Riemannian manifolds. *International Journal of Computer Vision*, 105(2):171–185, 2013. 2, 5

[9] 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*, 5(4):483–548, 2006. 2

[10] M. Giaquinta and D. Mucci. Maps of bounded variation with values into a manifold: total variation and relaxed energy. *Pure and Applied Mathematics Quarterly*, 3(2):513–538, 2007. 2

[11] P. Huber. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1):73–101, 1964. 3

[12] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988. 2

[13] D. G. Kendall. Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, 16(2):81–121, 1984. 2, 4, 5, 8

[14] J. Lellmann, E. Strekalovskiy, S. Koetter, and D. Cremers. Total variation regularization for functions with values in a manifold. In *IEEE International Conference on Computer Vision (ICCV)*, pages 2944–2951, 2013. 1, 2, 3

[15] A. Levinshtein, A. Stere, K. Kutulakos, D. Fleet, S. Dickinson, and K. Siddiqi. Turbopixels: Fast superpixels using geometric flows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(12):2290–2297, 2009. 7

[16] P. Michor and D. Mumford. Riemannian geometries on the space of plane curves. *Journal of the European Mathematical Society*, 8:1–48, 2006. 3

[17] P. W. Michor and D. Mumford. An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23(1):74 – 113, 2007. Special Issue on Mathematical Imaging. 2, 3

[18] A. Papazoglou and V. Ferrari. Fast object segmentation in unconstrained video. In *IEEE International Conference on Computer Vision (ICCV)*, pages 1777–1784, 2013. 1, 7

[19] I. Rahman, I. Drori, V. Stodden, D. Donoho, and P. Schröder. Multiscale representations for manifold-valued data. *Multiscale Modeling and Simulation*, 4(4):1201–1232, 2005. 1

[20] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(14):259 – 268, 1992. 3

[21] C. Samir, P.-A. Absil, A. Srivastava, and E. Klassen. A gradient-descent method for curve fitting on Riemannian manifolds. *Foundations of Computational Mathematics*, 12(1):49–73, 2012. 2

[22] A. Srivastava, E. Klassen, S. Joshi, and I. Jermyn. Shape analysis of elastic curves in Euclidean spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(7):1415–1428, 2011. 3

[23] A. Srivastava, P. Turaga, and S. Kurtek. On advances in differential-geometric approaches for 2d and 3d shape analyses and activity recognition. *Image Vision Computing*, 30(6-7):398–416, June 2012. 2

[24] J. Su, I. Dryden, E. Klassen, H. Le, and A. Srivastava. Fitting smoothing splines to time-indexed, noisy points on nonlinear manifolds. *Image and Vision Computing*, 30(67):428 – 442, 2012. 2

[25] G. Sundaramoorthi, A. Mennucci, S. Soatto, and A. Yezzi. Tracking deforming objects by filtering and prediction in the space of curves. In *Conference on Decision and Control*, 2009. 2, 3

[26] G. Sundaramoorthi, A. Yezzi, and A. Mennucci. Coarse-to-fine segmentation and tracking using Sobolev active contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence,*, 30(5):851 –864, 2008. 2, 3

[27] G. Sundaramoorthi, A. Yezzi, A. Mennucci, and G. Sapiro. New possibilities with Sobolev active contours. *International Journal of Computer Vision*, 84(2):113–129, 2009. 2, 3

[28] D. Tsai, M. Flagg, and J. Rehg. Motion coherent tracking with multi-label MRF optimization. In *British Machine Vision Conference*, 2010. 1, 7

[29] A. Weinmann, L. Demaret, and M. Storath. Total variation regularization for manifold-valued data. *SIAM Journal on Imaging Sciences*, 7(4):2226–2257, 2014. 1, 2, 3, 4

[30] L. Younes, P. Michor, J. Shah, and D. Mumford. A metric on shape spaces with explicit geodesics. *Technical report*, 2007. 3