# Total Variation Regularization of Shape Signals

Maximilian Baust[1], Laurent Demaret[2], Martin Storath[3], Nassir Navab[14], 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.

## 1 Introduction

In this paper, we introduce 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 latter can be seen as generalization of the Rudin-Osher-Fatemi (ROF) functional to shape-valued data, cf. also [4]. Our framework is generic in the sense that it can be combined with any shape space. Concretely, we derive a variant of the classical finite-dimensional representation of Kendall which allows for the explicit computation of geodesics. So, it facilitates the efficient numerical treatment of the variational formulation by means of a cyclic proximal point algorithm. Similar to the ROF-functional, we demonstrate experimentally that $\ell_1$-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.

## 2 The Model

A shape(-valued) signal of length $k$ is a vector $\mathbf{f} = (\mathbf{f}_1, \dots, \mathbf{f}_k) \in \mathcal{M}^k$, with each member being an element of the shape space $\mathcal{M}$. We assume that $\mathbf{f}$ is given, e.g., as a result of a segmentation algorithm; we wish to find a regularized version $\mathbf{x}^*$ which is given as a minimizer of the functional

$$E(\mathbf{x}) = D(\mathbf{x}, \mathbf{f}) + \alpha R(\mathbf{x}), \quad \alpha > 0. \tag{1}$$

While $D(\mathbf{x}, \mathbf{f})$ is a data fidelity term which penalizes the deviation from $\mathbf{f}$, $R(\mathbf{x})$ is a regularizer penalizing large variation in $\mathbf{x}$. We consider data terms

$$D(\mathbf{x}, \mathbf{f}) = \sum_{i=1}^{k} (h \circ d)(\mathbf{x}_i, \mathbf{f}_i), \tag{2}$$

where $h$ is one of the following functions: $h(s) = s$ which leads to an $\ell_1$-type penalization, $h(s) = s^2/2$ which leads to an $\ell_2$-type penalization, and

$$h(s) = \begin{cases} s^2, & s < 1/\sqrt{2}, \\ \sqrt{2}s - 1/2, & \text{otherwise}, \end{cases} \tag{3}$$

which yields the manifold-valued equivalent of the well-known Huber-norm – a differentiable compromise between the $\ell_1$-norm and the $\ell_2$-norm. Furtherore, we consider regularizers of the form

$$R(\mathbf{x}) = \sum_{i=1}^{k-1} (h \circ d)(\mathbf{x}_i, \mathbf{x}_{i+1}). \tag{4}$$

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 $h(s) = s^2/2$. Further, it can be seen as the total variation regularizer if $h(s) = s$. In case of (3), $R$ is a shape-valued differentiable approximation of the total variation regularizer (sometimes called Huber-ROF), which can be used to avoid the staircasing problem associated with total variation denoising, cf. Chambolle and Pock [1].

## 3 The Shape Space

We derive a shape representation which is *not* rotationally invariant which we call *oriented Kendall shapes*. By normalizing $x$ w.r.t. translation, we
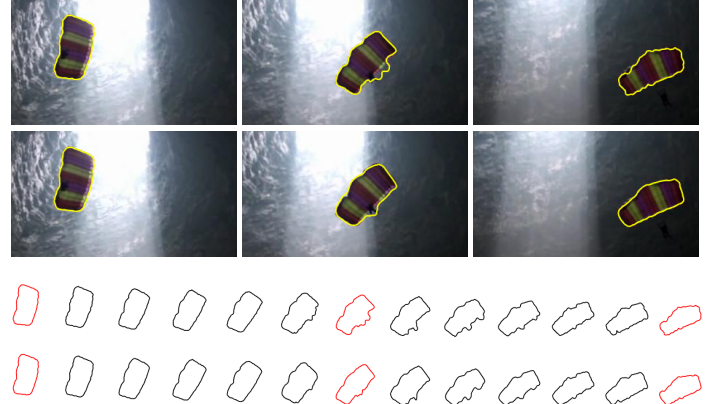


Figure 1: **Regularization of Shape Signals:** Objects segmented from video data enjoy a natural temporal ordering and thus form a *shape signal*. First row: Three frames of the "parachute" sequence from [3] segmented with the method proposed by [2]. Second row: Regularized shapes obtained with our method. 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.

remove *two real degrees of freedom* and obtain as the representation space

$$V_{2n-2} = \left\{ x \in \mathbb{R}^{2n} : \sum_{i=1}^{2n} x_i = 0 \right\} \subset \mathbb{R}^{2n}. \tag{5}$$

Next, we notice that a shape $x \in V_{2n-2}$ can be scaled by multiplying all 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

$$L_x = \{s \cdot x : s \in \mathbb{R} \setminus \{0\}\}. \tag{6}$$

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, synonymously, with the real unit sphere $S_{\mathbb{R}}^{2n-3}$ (with antipodal points identified). This means that by enforcing scale invariance we are removing *another degree of freedom*. As a consequence, the exponential mapping and the inverse exponential mapping are given by the respective mappings of $S_{\mathbb{R}}^{n-3}$. In contrast to the rotation invariant version of the Kendall shape space, we employ the real-valued scalar product $\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$ as well as its induced norm on $S_{\mathbb{R}}^{2n-3}$.

[1] 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.

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

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

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