

# An $SL(2)$ Invariant Shape Median

Benjamin Berkels · Gina Linkmann · Martin Rumpf

the date of receipt and acceptance should be inserted later

**Abstract** Median averaging is a powerful averaging concept on sets of vector data in finite dimensions. A generalization of the median for shapes in the plane is introduced. The underlying distance measure for shapes takes into account the area of the symmetric difference of shapes, where shapes are considered to be invariant with respect to different classes of affine transformations. To obtain a well-posed problem the perimeter is introduced as a geometric prior. Based on this model, an existence result can be established in the class of sets of finite perimeter. As alternative invariance classes other classical transformation groups such as pure translation, rotation, scaling, and shear are investigated. The numerical approximation of median shapes uses a level set approach to describe the shape contour. The level set function and the parameter sets of the group action on every given shape are incorporated in a joint variational functional, which is minimized based on step size controlled, regularized gradient descent. Various applications show in detail the qualitative properties of the median.

**AMS Subject Classifications (2000)**  
65M60, 68U10, 49J45

## 1 Introduction

In this paper, we develop a notion of a median of shapes in 2D. Median averaging is regarded as a particularly powerful and robust concept on a set of vectorial data in finite dimensions. It can be phrased in terms of an

---

Institut für Numerische Simulation, Universität Bonn  
Endenicher Allee 60, 53115 Bonn, Germany  
E-mail: {benjamin.berkels, gina.linkmann,  
martin.rumpf}@ins.uni-bonn.de

optimization problem. Given  $x_1, \dots, x_n \in \mathbb{R}^d$ ,

$$m^* = m^*[x_1, \dots, x_n]$$

is called the Euclidean median if it fulfills

$$m^* \in \operatorname{argmin}_{m \in \mathbb{R}^d} \sum_{i=1}^n |m - x_i|.$$

Compared to a minimization of squared differences in case of the classical arithmetic mean, the linear growth ensures robustness with respect to outliers. A well known property of this median is that it is not unique. Indeed, for  $x_1, x_2 \in \mathbb{R}^d$  we have

$$\min_{m \in \mathbb{R}^d} (|m - x_1| + |m - x_2|) = |x_2 - x_1|,$$

and the minimum is attained for each convex combination of  $x_1$  and  $x_2$ . Generalizing this averaging approach to shapes requires first a suitable definition of distances between shapes and then a transfer of the optimality property into the context of shapes. Our intention is to derive a rigorous definition of shape medians and to highlight some of the resulting properties by a set of characteristic examples.

Very basic notions of averaging include the arithmetic mean of landmark positions [12], and the image obtained by the arithmetic mean of the matching deformations [27,6]. Chen and Parent [11] investigated averages of 2D contours already in 1989. Jiang et al. [20] defined median shapes of polygonal curves based on weights for edit operations, which transfer one curve into the other. A generalization of this approach has been presented by Jiang et al. [19]. Furthermore, the computation of shape distances naturally appears when matching shapes in images. Yezzi and Soatto [34] have investigated shape averages in image structure reconstruction and joint registration.

Conceptually, in the last decade correlations of shapes have been studied on the basis of a general framework of a space of shapes and its intrinsic structure. The notion of a shape space was already introduced by Kendall [23] in 1984. Charpiat et al. [9,10] discuss shape averaging and shape statistics based on the notion of the Hausdorff distance and on the  $H^1$  norm of the difference of the signed distance functions of shapes. They study gradient flows for energies defined as functions over these distances for the warping between two shapes. As the underlying metric they use a weighted  $L^2$  metric, which weights translational, rotational, and scale components different than the component in the orthogonal complement of all these transforms. Furthermore, they investigated in this general framework also an PCA-analysis of shapes.

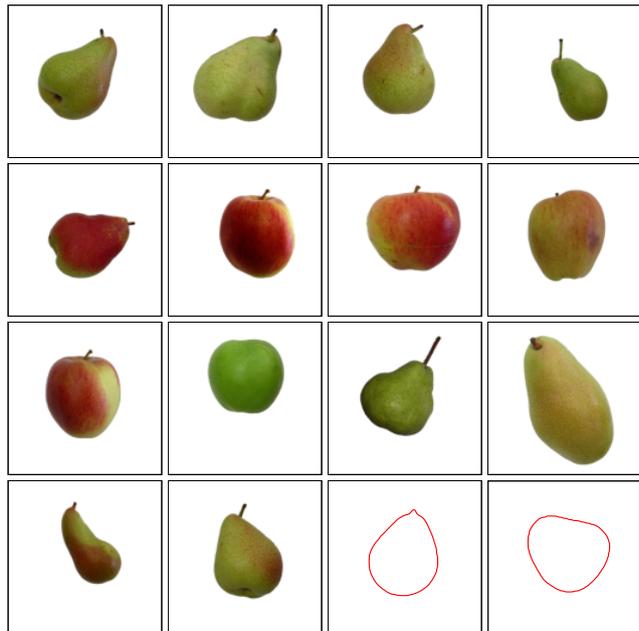
Fletcher et al. [15] studied transformation groups on polygonal medial axis models for shapes in 3D. Here, the average of a set of input shapes is defined as a shape reconstructed from an averaged medial axis. The averaged medial axis minimizes a sum of tangent vectors in the Lie group of the transformation group, which reproduce the set of input medial axes, if the exponential map is applied to them.

Understanding the shape space as an infinite-dimensional Riemannian manifold, Miller et al. [25,24] defined the average shape  $\mathcal{S}$  of samples  $\mathcal{S}_1, \dots, \mathcal{S}_n$  as the minimizer of  $\sum_{i=1}^n d(\mathcal{S}_i, \mathcal{S})^2$  for some metric  $d(\cdot, \cdot)$ , e. g. a geodesic distance in the space of shapes. This generalization of the geometric mean for objects on a Riemannian manifold has originally been proposed by Fréchet already in [17] and further analyzed by Karcher [22]. Beg et al. [3] introduced a geodesic distance between images or shapes based on a regularizing metric on Eulerian transport fields generating Lagrangian deformation between objects. The metric involves  $L^2$  scalar products of higher derivatives. This approach was used in comparative medical anatomy for instance by Joshi et al. [21] and Avants and Gee [2].

In [29,28], the average shape is defined as the shape which minimizes the sum of all elastic energies stored in optimal elastic deformations from the different input shapes onto the selected shape. This approach conceptually differs from the Riemannian approach. Fuchs et al. [18] proposed a viscoelastic notion of the distance between shapes given as boundaries of physical objects. A link between elastic shape matching and a Riemannian geodesic distance is investigated in [33]. Fletcher et al. [16] propose to use a shape median instead of the geometric shape mean. They investigate the median concept for rotations, tensors and for planar, polygonal shape contours.

By definition, the overall shape averaging problem falls into the class of shape optimization. We refer to the book by Sokołowski and Zolésio [31] for a comprehensive overview on this topic.

In this paper, we are investigating shapes which are implicitly described via non-binary images. Hence, shape information has to be extracted from the given data. A first example is depicted in Figure 1. We con-



**Fig. 1** Given nine pears and five apples as input shapes, the median shape is computed by the proposed method. The underlying definition of the median includes the invariance with respect to rigid body motions (second last contour on the bottom right) and with respect to area (and orientation) preserving affine transformations (right most contour on the bottom line).

sider a joint approach based on a Mumford–Shah type variational formulation [26] that handles the median as a free discontinuity set. For the implementation we pick up the robust and effective approximation of the Mumford–Shah functional proposed by Chan and Vese [8]. As a distance measure between shapes we consider the area of the symmetric difference between sets. Thereby, shapes are equivalence classes of sets with respect to certain finite dimensional groups of transformations in the plane. Examples for such transformation groups studied here are translations, rotations, scaling, and shear. For a general outline of infinite dimensional diffeomorphic group actions on shapes we refer to Dupuis et al. [13]. Computationally, the involved transformation group requires a simultaneous minimization over the actual median shape and a set of transformations. If the transformation group is finite the second problem is a labeling problem. In our case the

transformations groups are continuous and smoothly parametrized over a small set of real valued parameters, which enables an effective continuous minimization approach. In the underlying numerical approximation, the median shape is represented as the zero level set of a level set function. When averaging  $n$  shapes the nodal values of a piecewise multilinear finite element representation of this level set function and the parameter sets of the  $n$  transformations of the individual shapes are considered as degrees of freedom for the variational problem. In addition, contrast values related to the Mumford–Shah segmentation model are taken into account as degrees of freedom, which will be investigated below. We give various examples for the qualitative behavior and the overall performance of the resulting algorithm. We note that a preliminary version of part of the work reported in this article has appeared in the proceedings paper [4].

The paper is organized as follows: In Section 2, we define the underlying shape distance and discuss the invariance with respect to different transformation classes, whereas Section 3 gives a variational definition of the shape median. In Section 4, existence of regularized shape medians in the class of set of finite perimeter is established. The numerical approximation based on a level set representation of shapes in a multilinear finite element space and the resulting algorithm are discussed in Section 5.

## 2 An area based distance measure

To derive the variational model for a shape median we will first investigate the underlying distance measure between two sets in  $\mathbb{R}^2$ . The latter is based on the area of the symmetric difference and takes into account a particular invariance class of affine transformations. Even though our particular focus is on  $SL(2)$  invariance, we will discuss other transformation classes as well along with suitable parametrizations. Given this distance measure we then define the actual shape median and investigate the underlying problem of well-posedness.

### 2.1 A transformation invariant symmetric distance between sets

A standard tool for the comparison of sets is the symmetric difference. Given two measurable sets  $A, B \subset \mathbb{R}^2$  one defines  $A \Delta B = A \setminus B \cup B \setminus A$ . Hence, a corresponding measure for the difference of the two sets is the area of the symmetric difference

$$d(A, B) := |A \Delta B|.$$

One easily verifies that for given sets  $A, B, C$  and  $x \in A \Delta B$  either  $x \in A \Delta C$  or  $x \in C \Delta B$ . This implies the triangle inequality  $d(A, B) \leq d(A, C) + d(C, B)$ , and together with the obvious symmetry  $d(A, B) = d(B, A)$  and the positive definiteness  $d(A, B) \geq 0$  ( $d(A, B) = 0 \Leftrightarrow A = B$  (up to a set of measure zero)) it follows that  $d(\cdot, \cdot)$  is a metric on subsets of  $\mathbb{R}^2$ . For the actual comparison of two shapes one has to take into account the invariance with respect to certain underlying group actions. Let us suppose that  $T$  is a group of area preserving deformations  $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ . Examples to be considered later are translations, rigid body motions, and in particular area preserving affine mappings. Given a particular transformation group  $T$  of area preserving transformations we define a distance between shapes represented by sets  $A, B \subset \mathbb{R}^2$  with respect to  $T$  by

$$d_T(A, B) := \inf_{\phi \in T} d(A, \phi(B)). \quad (1)$$

One easily verifies that  $d_T$  defines a metric on equivalence classes of subsets of  $\mathbb{R}^2$  with respect to the transformation group  $T$ . Indeed, from the fact that  $\phi \in T$  is area preserving we deduce

$$d(A, \phi(B)) = d(\phi^{-1}(A), \phi^{-1}(\phi(B))) = d(B, \phi^{-1}(A))$$

and hence

$$d_T(A, B) = \inf_{\phi \in T} d(B, \phi^{-1}(A)) = d_T(B, A)$$

follows from the group property of  $T$ . Finally, with  $\psi = \operatorname{argmin}_{\phi \in T} d(A, \phi(C))$  the triangle inequality follows from the estimate

$$\begin{aligned} \inf_{\phi \in T} d(A, \phi(B)) &\leq d(A, \psi(C)) + \inf_{\phi \in T} d(\psi(C), \phi(B)) \\ &= d_T(A, C) + \inf_{\phi \in T} d(C, \psi^{-1} \circ \phi(B)) \\ &= d_T(A, C) + d_T(C, B). \end{aligned}$$

For later reference let us remark that the metric (1) on the space of 2D shapes can be represented in integral form

$$d_T(A, B) = \inf_{\phi \in T} \int_{\mathbb{R}^2} (1 - \chi_A) \chi_{\phi(B)} + \chi_A (1 - \chi_{\phi(B)}) dx, \quad (2)$$

where  $\chi_C$  denotes the characteristic function of a set  $C$ .

### 2.2 Different classes of transformations

In this paper, we confine to subclasses of the group of affine transformations  $\phi$  with  $\phi(x) = Ax + b$ , where  $A \in GL(2)$  and  $b \in \mathbb{R}^2$ . A particularly simple group is

the translation group  $T_{\text{trans}}$  with  $A = \mathbb{I}$  (the identity matrix). As usual the group of rotations

$$Q = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

is denoted by  $\text{SO}(2)$  so that we obtain the group of rigid body motions  $T_{\text{SO}(2)}$  via  $\phi(x) = Ax + b$  with  $A \in \text{SO}(2)$ . The largest area (and orientation) preserving group in the set of affine transformations is described by matrices  $A \in \text{SL}(2) = \{A \in \text{GL}(2) \mid \det A = 1\}$ , here denoted by  $T_{\text{SL}(2)}$ . Obviously,  $T_{\text{SO}(2)}$  is a subgroup of  $T_{\text{SL}(2)}$ . In many applications the shapes to be averaged are of significantly different enclosed area, which renders the  $\text{SL}(2)$  invariant median computation questionable. Hence, throughout this paper in case of a median computation with respect to the transformation group  $T_{\text{SL}(2)}$  we a priori scale all input shapes (after standard binary Mumford–Shah segmentation) to ensure that they are all of equal enclosed area. Based on the polar decomposition formula they can be parametrized as follows

$$A = Q \begin{pmatrix} \beta & \gamma \\ \gamma & \frac{1+\gamma^2}{\beta} \end{pmatrix}$$

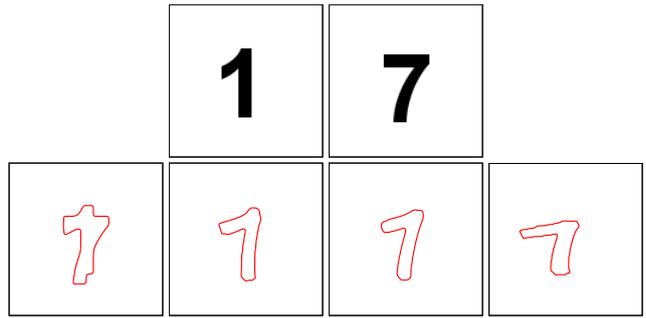
with  $Q \in \text{SO}(2)$ . For example shearing in  $x$  direction is a one dimensional subgroup of  $\text{SL}(2)$  given by matrices

$$A = \begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix}$$

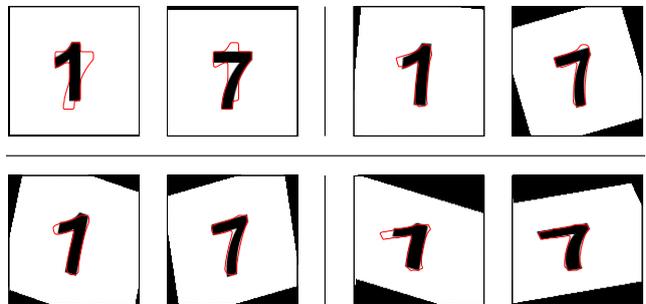
and anisotropic area preserving scaling, another subgroup, is given by matrices

$$A = Q^T \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} Q$$

with  $\alpha \in \mathbb{R}^+$  and  $Q \in \text{SO}(2)$ . With respect to the variational model for a shape median to be developed in this paper, we will take into account different groups of area preserving transformations restricting ourselves not to only subgroups but to just subsets, such as the set of concatenations of rotation and shear. In many applications invariance with respect to scaling is a desirable property. To incorporate isotropic scaling in our model the distance measure has to be modified. This goes beyond the scope of the current paper. Thus, we confine here to a simple incorporation of scale invariance in the current model, even though the underlying  $d_T(\cdot, \cdot)$  is no longer a metric. Figure 2 shows the impact of the transformation group on the average shape - to be defined in the next paragraph - of two different shapes. Figure 3 complements the previous figure by depicting  $\phi_1(A_1)$  and  $\phi_2(A_2)$  overlaid with the median for the same experiment. Throughout this paper, shapes are described via image contours. To show the underlying transformation, we render in



**Fig. 2** Averaging of two shapes, “1” and “7”, with different choices of the transformation group  $T$ : The first row shows the two input images, the second row shows the average shape in case of the transformation class of simple translations  $T_{\text{trans}}$  (first), rigid body motions combined with isotropic scaling  $T_{T,\text{SO}(2)}$  (second), the concatenation of rigid body motion, shearing and isotropic scaling  $T_{T,\text{shear}}$  (third), and  $T_{\text{SL}(2)}$  (forth), respectively.



**Fig. 3** The transformed input shapes are overlaid with the median shape contour for the group of translations  $T_{\text{trans}}$  (top left block), concatenation of rigid body motion and isotropic scaling from  $T_{T,\text{SO}(2)}$  (top right block), concatenation of rigid body motion, shearing and isotropic scaling  $T_{T,\text{shear}}$  (bottom left block), and the transformation group  $T_{\text{SL}(2)}$  (bottom right block).

figures also the transformed image domain  $\phi^{-1}([0, 1]^2)$  on a black background. With respect to the numerical algorithm - to be discussed in Section 5 - let us discuss the parametrization of the aforementioned transformation classes. In general, we will consider mappings  $\phi(q, x) = A(q)x + b(q)$ , where  $A(q) \in \text{GL}(2)$ ,  $b(q) \in \mathbb{R}^2$  and  $q$  denotes a suitable parametrization of the degrees of freedom of the affine transformations we allow for. In fact, we have one parameter set  $q_i$  for every input shape  $A_i$  with  $i = 1, \dots, n$ . All of our invariance classes  $T$  will contain translations, since translation invariance is fundamental in all applications. First, we introduce the matrices

$$Q(\theta) := \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix},$$

$$S(\beta, \gamma) := \begin{pmatrix} 1 & \tan(\beta) \\ \tan(\gamma) & 1 \end{pmatrix}$$

with parameters  $\theta$ ,  $\beta$ , and  $\gamma$ , respectively. In case of rigid body motions from the class  $T_{\text{SO}(2)}$  we choose  $q_i =$

$(\theta_i, b_i^1, b_i^2)$ , where

$$\phi_i(q_i, x) = Q(\theta_i)x + b_i$$

with  $b_i = (b_i^1, b_i^2)$ . If we enlarge the transformation class by isotropic scaling we obtain the class  $T_{\tau, \text{SO}(2)}$  (in particular skipping the area preservation assumption) and consider  $q_i = (\tau_i, \theta_i, b_i^1, b_i^2)$  with  $\phi_i(q_i, x) = \tau_i Q(\theta_i)x + b_i$  and  $\tau_i > 0$ . Taking into account an axially aligned shear in the reference configuration of all images, we obtain the parameter vector  $q_i = (\beta_i, \gamma_i, \theta_i, b_i^1, b_i^2)$  describing the transformation

$$\phi_i(q_i, x) = Q(\theta_i)S(\beta_i, \gamma_i)x + b_i.$$

from the class  $T_{\text{shear}}$ . Furthermore, multiplication with a isotropic scaling factor  $\tau_i$  as above we obtain the class  $T_{\tau, \text{shear}}$  of concatenations of isotropic scaling, rigid body motion, and shearing and set  $q_i = (\tau_i, \beta_i, \gamma_i, \theta_i, b_i^1, b_i^2)$ . Finally, we obtain a parametrization of the transformation group  $T_{\text{SL}(2)}$  with

$$\phi_i(q_i, x) = Q(\theta_i)B(\beta_i, \gamma_i)x + b_i \text{ where}$$

$$B(\beta, \gamma) := \begin{pmatrix} \beta & \gamma \\ \gamma & \frac{1+\gamma^2}{\beta} \end{pmatrix}$$

and the parameter vector  $q_i = (\beta_i, \gamma_i, \theta_i, b_i^1, b_i^2)$ .

### 3 A variational description of the shape median

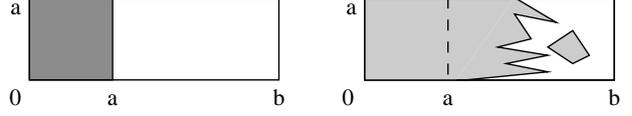
In this Section, we will see that in the context of sets a suitable regularization process has been taken into account to derive a proper definition of a median of a set of input shapes. Here, we will derive a suitable variational model for which we will prove existence in Section 4. The numerical discretization in Section 5 picks up this model and discretizes it via a level set approach.

#### 3.1 Imposing regularity of the median shape

Recalling the definition of the median in the context of vectors in  $\mathbb{R}^n$ , a preliminary definition of a median  $M^*$  of shapes  $S_1, \dots, S_n \subset \mathbb{R}^2$  and a set  $T$  of invariant transformations is given by

$$M^* \in \operatorname{argmin}_{M \subset \mathbb{R}^2} \sum_{i=1}^n d_T(M, S_i).$$

As in the case of the median of numbers (cf. Section 1) the median is not unique. As an example let us consider a square  $S_1 = [0, a]^2$  and a rectangle  $S_2 = [0, a] \times [0, b]$  with  $a < b$  in combination with  $T = T_{\text{trans}}$ . Obviously, every rectangle  $M = [0, a] \times [0, c]$  with  $a \leq c \leq b$  renders the sum of distances  $\sum_{i=1}^2 d_T(M, S_i)$  to be  $(b-a)a$ , which is equal to  $d_T(S_1, S_2)$  and thus the minimal value.



**Fig. 4** Without the boundary length penalization shape medians turn out to be fairly irregular.

Unfortunately, the set of minimizers is even larger. We can split the rectangle  $[0, a] \times [0, b]$  into any two disjoint measurable subsets  $\tilde{S}_1, \tilde{S}_2$  and  $M = S_1 \cup \tilde{S}_1$  will be a minimizer as well (cf. Figure 4). To select from the set of minimizers a proper candidate for the median we have to filter out irregular “median” shapes. Thus, let  $\mathcal{M}[S_1, \dots, S_n]$  be the set of all sets  $M \subset \mathbb{R}^2$  which minimize the sum of shape distances. Then, select from this set the shape  $M^*$  with the smallest perimeter (boundary length) and define this set as the shape median, i. e.

$$M^* \in \operatorname{argmin}_{M \in \mathcal{M}[S_1, \dots, S_n]} \operatorname{Per}(M), \quad (3)$$

where  $\operatorname{Per}(A)$  denotes the length of the boundary  $\partial A$  of  $A$  if  $\partial A$  is rectifiable and else is set to  $\infty$ . Still the shape median might not be unique, as pointed out by the following example. Let

$$S_1 = B_1(0) \setminus A \left( \frac{1}{2}, \frac{3}{4}, -\frac{\pi}{2}, \frac{\pi}{2} \right) \cup A \left( 1, 2, -\frac{\pi}{16}, \frac{\pi}{16} \right)$$

$$S_2 = B_1(0) \setminus A \left( \frac{1}{2}, \frac{3}{4}, \frac{\pi}{4}, \frac{3\pi}{4} \right) \setminus A \left( \frac{1}{2}, \frac{3}{4}, -\frac{3\pi}{4}, -\frac{\pi}{4} \right) \cup A \left( 1, 2, -\frac{\pi}{8}, \frac{\pi}{8} \right)$$

where

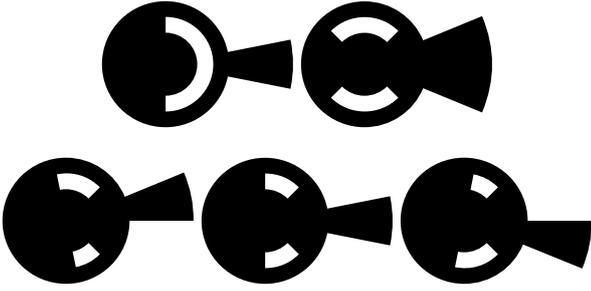
$$A(r_1, r_2, \alpha_1, \alpha_2) = \{x = (r \sin(\alpha), r \cos(\alpha)) \mid r \in [r_1, r_2], \alpha \in [\alpha_1, \alpha_2]\}.$$

We observe that

$$M(\beta) = B_1(0) \setminus \left[ A \left( \frac{1}{2}, \frac{3}{4}, \frac{\pi}{4}, \frac{\pi}{2} - \beta \right) \cup A \left( \frac{1}{2}, \frac{3}{4}, -\frac{\pi}{2} - \beta, -\frac{\pi}{4} \right) \right] \cup A \left( 1, 2, -\frac{\pi}{16} - \beta, \frac{\pi}{16} - \beta \right)$$

for every  $\beta \in [-\frac{\pi}{16}, \frac{\pi}{16}]$  is a rigid body motion invariant median shape, with respect to the definition (3) (cf. Figure 5). But in general variations of the median shape appear to be strongly limited (in this case there is only a one parameter family). Let us point out the conceptual similarity to the non uniqueness of the median of a point sets in  $\mathbb{R}^d$ .

The above definition based on a selection principle from a candidate set based on the perimeter is neither analytically nor computationally very handsome.



**Fig. 5** For the two shapes  $S_1$  and  $S_2$  on the top, three different median shapes  $M(-\frac{\pi}{16})$ ,  $M(0)$ , and  $M(\frac{\pi}{16})$  from a one parameter family of median shapes are depicted in the bottom row.

Hence, in what follows we consider a regularized median shape as the minimizer of a regularized functional. We introduce a  $\gamma$ -median shape  $M_\gamma^*$  as a minimizing set of the energy  $\frac{1}{n} \sum_{i=1}^n d_T(M, S_i) + \gamma \text{Per}(M)$ , where  $\gamma$  is a small, positive constant. For the applications considered in this paper we have chosen  $\gamma = 0.0005$ , where the images from which the shapes are extracted via segmentation are defined on the image domain  $[0, 1]^2$  (see below). Figure 6 depicts the resulting shape median in case of a letter represented in different fonts. Due to the



**Fig. 6** Median of four “B” (right most), with respect to the transformation class  $T_{\text{SL}(2)}$ .

built-in transformation invariance of the distance measure  $d_T(\cdot, \cdot)$  we have to minimize not only over the set of shapes  $M$  but at the same time over the set of transformations  $\phi_1, \dots, \phi_n \in T$  appearing in  $d(M, \phi_i(A_i))$ . Thus, we are led to a joint minimization problem for the functional

$$E_\gamma[M, \Phi] = \frac{1}{n} \sum_{i=1}^n d(M, \phi_i(S_i)) + \gamma \text{Per}(M). \quad (4)$$

where  $\Phi = (\phi_1, \dots, \phi_n)$  is defined as the  $n$ -tuple of transformations (in the implementation the finite dimensional vector of all transformation parameters). In the next section we will generalize the problem in the context of shapes being described via image contours and comment on the embedding of the above set based problem.

### 3.2 A shape median for shapes encoded in images

As already mentioned in the introduction, in practical applications the input data usually is not a number

of sets  $S_1, \dots, S_n$ , but images with intensity functions  $u_1, \dots, u_n : \Omega \rightarrow \mathbb{R}$ , each of them encoding a shape. Throughout the applications considered in this paper, the domain  $\Omega$  is considered to be the unit square  $[0, 1]^2$ . Furthermore, we replace the set  $M$  by a characteristic function  $m$  with  $M = \{x \in \Omega \mid m^+(x) = 1\}$  and consider  $BV(\Omega)$  as the natural space for these functions. Here,  $m^+$  is the approximate lim sup of  $m$ . If the image functions  $u_i$  are characteristic functions, i.e.  $u_i = \chi_{S_i}$ , and  $m \in BV$  then based on the observation in (2) we obtain for the above energy

$$\mathcal{E}_\gamma[m, \Phi] = \frac{1}{n} \sum_{i=1}^n \int_{\Omega} (1 - m \circ \phi_i) u_i^2 + m \circ \phi_i (1 - u_i)^2 dx + \gamma |Dm|(\Omega), \quad (5)$$

where  $|Dm|(\Omega)$  denotes the total variation of  $m$  on  $\Omega$ . We have chosen an exponent 2 in the above functional to stress the relation to a more general image model, which will be introduced in what follows. Obviously, one can rephrase this minimization problem considering the functional

$$\tilde{\mathcal{E}}_\gamma[m, \Phi] = \frac{1}{\gamma n} \left( \sum_{i=1}^n \int_{\Omega} (1 - m \circ \phi_i) \chi_{S_i} + m \circ \phi_i (1 - \chi_{S_i}) dx - D \right) + |Dm|(\Omega),$$

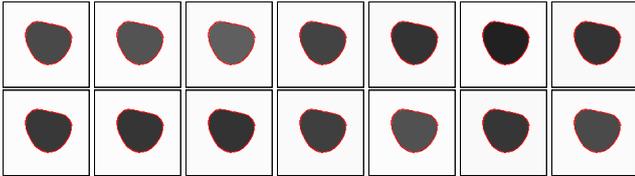
where  $D = \inf_{M \subset \mathbb{R}^2} \sum_{i=1}^n d_T(M, S_i)$ . Thus, the open problem - which is beyond the scope of this paper - is if  $\tilde{\mathcal{E}}_\gamma$   $\Gamma$ -converges to the functional  $\tilde{\mathcal{E}}$  with respect to the weak  $\star$  topology in  $BV$ , where  $\tilde{\mathcal{E}}[m, \Phi] = |Dm|(\Omega)$  for  $\sum_{i=1}^n \int_{\Omega} (1 - m \circ \phi_i) \chi_{S_i} + m \circ \phi_i (1 - \chi_{S_i}) dx = D$  and  $\infty$  else. For a comprehensive introduction to  $BV$  we refer to the text book by Ambrosio et al. [1]. To allow the evaluation of  $m \circ \phi$  at points  $x \in \Omega$  with  $\phi(x)$  outside  $\Omega$  we introduce the convention that  $m = 0$  outside  $\Omega$ . In case of general images, shapes are frequently extracted via segmentation. Here, we pick up the piecewise constant Mumford–Shah model as the underlying segmentation model and suppose that a shape  $S_i$  in image  $u_i$  corresponds to an approximate gray value  $c_i^{\text{in}}$ , whereas the complement  $\Omega \setminus S_i$  is represented by an approximate gray value  $c_i^{\text{ext}}$ . To phrase the shape median in terms of these images, we formulate a simultaneous segmentation model based on the energy

$$\mathcal{E}_\gamma[m, \Phi, \mathcal{C}] = \frac{1}{n} \sum_{i=1}^n \int_{\Omega} (1 - m \circ \phi_i) (u_i - c_i^{\text{ext}})^2 + m \circ \phi_i (u_i - c_i^{\text{in}})^2 dx + \gamma |Dm|(\Omega). \quad (6)$$

Here,  $\mathcal{C} = (c_i^{\text{in}}, c_i^{\text{ext}})_{i=1, \dots, n}$  denotes the vector of all gray values. Alternatively, one could consider a  $L^1 - TV$

approach to improve robustness with respect to noise. We confine here to an  $L^2 - TV$  to limit the computational complexity. A perfect match of the median shape and the input shape encoded in image  $u_i$  is characterized by  $u_i = c_i^{\text{in}}$  on the pull back of the median shape  $M$  described by  $m$  under the deformation  $\phi_i$  and  $u_i = c_i^{\text{ext}}$  on the pull back of the complement of  $M$ . In (5)  $c_i^{\text{ext}} = 0$  and  $c_i^{\text{in}} = 1$  by assumption and if we in addition describe sets  $M$  and  $S_i$  by their characteristic functions  $m$  and  $u_i$ , respectively, we get back to the initial formulation for sets (4). Indeed,  $E_\gamma[M, \Phi] = \mathcal{E}_\gamma[\chi_M, \Phi, ((1, 0), \dots, (1, 0))]$ .

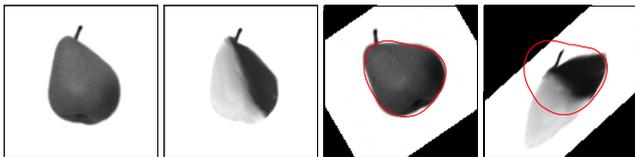
In this generalized model the values  $c_i^{\text{in}}$ ,  $c_i^{\text{ext}}$  for  $i = 1, \dots, n$  are additional degrees of freedom in our model, typically compensating contrast variations in the underlying images. Figure 7 shows the minimizing intensity values  $c_i^{\text{in}}$  and  $c_i^{\text{ext}}$  in the introductory example from Figure 1. Figure 8 demonstrates how contrast modulation in the given shapes might effect the resulting shape average.



**Fig. 7** Foreground and background of the  $T_{\text{SL}(2)}$  invariant median shape from the application depicted in Figure 1 are shaded with the intensity values  $c_i^{\text{in}}$  and  $c_i^{\text{ext}}$ , respectively, for every input image  $u_i$  with  $i = 1, \dots, 14$ .

#### 4 Existence of a shape median

In this section, we state and prove existence of a  $\gamma$ -median shape of a given set of input shapes under a suitable non degeneracy assumption. Here, we take into account the variational approach on a set of input images which we actually consider in the numerical ap-



**Fig. 8** Strong variations of contrast in a single shape might lead to wrong matching between the median shape and this shape. One of the input images from Figure 1 (first) is contrast modulated (second). The corresponding shape deformations overlaid with the different resulting average shapes are depicted on the right (third and fourth) based on the transformation group  $T_{\text{SL}(2)}$  (cf. Figure 13).

plications in this paper. For the sake of simplicity, we assume the images to be characteristic functions. Generalizations are straightforward for general binary image data with a sufficient contrast between foreground and background. In explicit, we obtain

**Theorem 1 (Existence of a minimizer for the binary Mumford–Shah model (6))** *Suppose that a set of input shapes  $S_1, \dots, S_n \subset \Omega$ , which are represented by image functions  $u_i$  with  $u_i = \chi_{S_i}$  on a bounded image domain  $\Omega$  with Lipschitz boundary, is given. These shapes are assumed to be non degenerate in the sense that there exist fixed constants  $\mu, \delta > 0$  and a set of points  $y_1, \dots, y_n$  such that  $B_\delta(y_i) \subset S_i$  and  $(|\Omega| - |S_i|) \geq \mu|\Omega|$  for  $i = 1, \dots, n$ . Furthermore, assume  $2\gamma < \delta\mu^2$ . Then, there exists a minimizer  $(m^*, \Phi^*, \mathcal{C}^*)$  of the energy  $\mathcal{E}_\gamma$  defined in (6) in the class of functions  $\{m \in BV(\Omega) | m(x) \in \{0, 1\} \text{ a. e.}\}$ , vectors of area preserving affine transformations  $\Phi = (\phi_i)_{i=1, \dots, n}$ , and contrast values  $\mathcal{C} \in \mathbb{R}^{2n}$ . The corresponding median shape  $\mathcal{M}^* := \{x \in \Omega | (m^*)^+(x) = 1\}$  is non degenerate, in the sense that there exist  $\epsilon > 0$  and  $y \in \Omega$  with  $B_\epsilon(y) \subset \mathcal{M}^*$ .*

*Proof* We prove the theorem applying the direct method in the calculus of variations, making use of the weak lower semi continuity of the total variation on the space  $BV$  and controlling the set of transformations in a uniform way. We proceed in several steps.

*Step 1.* For fixed  $m, \Phi$  the contrast values  $c_i^{\text{ext}}$  and  $c_i^{\text{in}}$  are functions of  $m$  and  $\phi_i$  for  $i = 1, \dots, n$ . In fact, from the Euler Lagrange equations for  $c_i^{\text{ext}}$  and  $c_i^{\text{in}}$  we deduce

$$\begin{aligned} c_i^{\text{ext}}[m, \phi_i] &= \frac{\int_\Omega (1 - m \circ \phi_i) u_i \, dx}{\int_\Omega (1 - m \circ \phi_i) \, dx} \\ c_i^{\text{in}}[m, \phi_i] &= \frac{\int_\Omega m \circ \phi_i u_i \, dx}{\int_\Omega m \circ \phi_i \, dx} \end{aligned} \quad (7)$$

and write  $\mathcal{C}[m, \Phi] = (c_i^{\text{in}}[m, \phi_i], c_i^{\text{ext}}[m, \phi_i])_{i=1, \dots, n}$ . As a convention, in case of a vanishing denominator in the above expressions we define the corresponding gray value as zero. Indeed this choice does not have any impact on the energy because the resulting value is evaluated on a set of measure zero in the functional.

*Step 2.* Let us consider a minimizing sequence

$$(m^k, \Phi^k, \mathcal{C}[m^k, \Phi^k])_{k=1, \dots}$$

and denote by  $\mathcal{E}$  the infimum and by  $\bar{\mathcal{E}}$  the maximal value of  $\mathcal{E}_\gamma[m^k, \Phi^k, \mathcal{C}[m^k, \Phi^k]]$  on this sequence. From  $|Dm^k|(\Omega) \leq \frac{1}{\gamma} \bar{\mathcal{E}}$  and  $m_k(x) \in \{0, 1\}$  a. e. on the bounded domain  $\Omega$  we deduce that  $(m^k)_{k=1, \dots}$  is bounded in  $BV(\Omega)$ . Thus, there exists a subsequence, for the sake of simplicity again denoted by  $(m_k)_{k=1, \dots}$ , such that  $m_k$  converges weak-\* to some

$m^*$  in  $BV$  (cf. Thm 3.23 in [1]). In particular  $m_k$  converges strongly in  $L^1$ , which implies that  $m^* = \chi_{M^*}$  with  $M^* := \{x \in \Omega \mid (m^*)^+(x) = 1\}$ . Furthermore, by the weak lower semicontinuity of the total variation we obtain that  $|Dm^*|(\Omega) \leq \liminf_{k=1, \dots} |Dm^k|(\Omega)$ .

*Step 3.* Now, we show that  $M^*$  itself is also non degenerate, i. e. it contains a ball of radius  $\delta > 0$ . If  $M^* = \partial M^*$  then we deduce from the boundedness of the total variation of  $m^*$  and the definition of 1- and 2-dimensional Hausdorff measures that  $|M^*| = 0$ . Hence, we can replace  $M^*$  by the empty set and  $m^*$  by 0 without increasing the limit energy, i. e. we obtain that

$$\begin{aligned} \mathcal{E} &= \frac{1}{n} \sum_{i=1}^n \int_{\Omega} \left( \chi_{S_i} - \frac{|S_i|}{|\Omega|} \right)^2 dx \\ &= \frac{1}{n} \sum_{i=1}^n \int_{\Omega} \chi_{S_i}^2 - 2 \frac{|S_i|}{|\Omega|} \chi_{S_i} + \frac{|S_i|^2}{|\Omega|^2} dx \\ &= \frac{1}{n} \sum_{i=1}^n |S_i| \left( 1 - \frac{|S_i|}{|\Omega|} \right). \end{aligned}$$

Now, we replace  $M^*$  by  $B_\delta(y)$  for some  $y \in \Omega$  with  $B_\delta(y) \subset \Omega$  and correspondingly replace  $m^*$  by  $\chi_{B_\delta(y)}$ . Furthermore, we define  $\tilde{\Psi}$  by  $\tilde{\phi}_i(x) = x - y_i + y$  and choose  $\tilde{C}$  with  $\tilde{c}_i^{\text{in}} = 1$  and  $\tilde{c}_i^{\text{ext}} = \frac{|S_i|}{|\Omega|}$ . Then, we can estimate

$$\begin{aligned} \mathcal{E}_\gamma[\chi_{B_\delta(y)}, \tilde{\Psi}, \tilde{C}] &= \frac{1}{n} \sum_{i=1}^n \int_{\Omega} (1 - \chi_{B_\delta(y)}(x - y_i + y)) \left( \chi_{S_i} - \frac{|S_i|}{|\Omega|} \right)^2 \\ &\quad + \chi_{B_\delta(y)}(x - y_i + y) (\chi_{S_i} - 1)^2 dx \\ &\quad + 2\pi\delta\gamma \\ &\leq \underline{\mathcal{E}} + 2\pi\delta\gamma - \pi\delta^2\mu^2. \end{aligned}$$

From this estimate and our assumption that  $2\gamma < \delta\mu^2$  we obtain the inequality  $\mathcal{E}_\gamma[\chi_{B_\delta(y)}, \tilde{\Psi}, \tilde{C}] < \underline{\mathcal{E}}$  which contradicts the definition of  $\underline{\mathcal{E}}$ . Hence, there exist a  $\epsilon > 0$  and a  $y \in \Omega$  such that  $B_\epsilon(y) \subset M^*$ .

*Step 4.* Now, we investigate the convergence of the vector of affine area preserving transformations  $\Phi^k = (\phi_i^k)_{i=1, \dots, n}$ . Each transformation  $\phi_i^k$  can be rewritten as  $\phi_i^k(x) = A_i^k x + b_i^k$ , where  $A_i^k \in \text{SL}(2)$  and  $b_i^k \in \mathbb{R}^2$ . Applying the polar decomposition, we can rewrite

$$A_i^k = Q_i^k B_i^k = Q_i^k R_i^k \text{diag}(\lambda_i^k, (\lambda_i^k)^{-1}) (R_i^k)^T,$$

where  $Q_i^k$  and  $R_i^k$  are rotation matrices in  $\mathbb{R}^2$  corresponding to rotation angles  $\alpha_i^k$  and  $\beta_i^k$ , respectively, and  $B_i^k$  is a symmetric matrix which is in the second step diagonalized. Rotation angles can be constrained to the interval  $[0, 2\pi)$ . Hence, to prove boundedness of the transformation  $\phi_i^k$  we have to show that there exists  $C, \underline{\lambda}, \bar{\lambda}$ , such that  $|b_i^k| \leq C$  and  $\underline{\lambda} \leq \lambda_i^k \leq \bar{\lambda}$  for

all  $i = 1, \dots, n$  and all  $k = 1, \dots$ . We prove this by contradiction. Thus, let us assume that on a subsequence of transformations, again denoted by  $(\phi_i^k)_{k=1, \dots}$  the translation vectors  $b_i^k$  or  $\lambda_i^k$  diverge for  $k \rightarrow \infty$ . Here, we remark that if  $\lambda_i^k \rightarrow \infty$  then  $(\lambda_i^k)^{-1} \rightarrow 0$ . Then  $m^k \circ \phi_i^k \rightarrow 0$  a. e. on  $\Omega$ . Thus, as above  $c_i^{\text{ext}, k} \rightarrow \frac{|S_i|}{|\Omega|}$  and by the Lebesgue convergence theorem

$$\begin{aligned} &\int_{\Omega} (1 - m^k \circ \phi_i^k)(u_i - c_i^{\text{ext}, k})^2 + m^k \circ \phi_i^k (u_i - c_i^{\text{in}, k})^2 dx \\ &\rightarrow \int_{\Omega} \left( \chi_{S_i} - \frac{|S_i|}{|\Omega|} \right)^2 dx = |S_i| \left( 1 - \frac{|S_i|}{|\Omega|} \right). \end{aligned}$$

Now, we replace for all  $k = 1, \dots$  the transformations  $\phi_i^k$ , and the contrast values  $c_i^{\text{in}, k}$ , and  $c_i^{\text{ext}, k}$  by  $\tilde{\phi}_i : x \mapsto \frac{\text{diam}(\Omega)}{2\delta}(x - y_i) + x_\Omega$  (where  $x_\Omega$  is chosen such that  $\Omega \subset B_{\frac{\text{diam}(\Omega)}{2}}(x_\Omega)$ ),  $\tilde{c}_i^{\text{in}} = 1$ , and  $\tilde{c}_i^{\text{ext}} = \frac{|S_i|}{|\Omega|}$ , respectively. Obviously,

$$\begin{aligned} &B_{\frac{2\delta\epsilon}{\text{diam}(\Omega)}} \left( y_i + \frac{2\delta}{\text{diam}(\Omega)}(y - x_\Omega) \right) \\ &= \tilde{\phi}_i^{-1}(B_\epsilon(y)) \subset \tilde{\phi}_i^{-1}(M^*) \subset B_\delta(y_i) \subset S_i. \end{aligned}$$

Therefore, we can estimate

$$\begin{aligned} &\int_{\Omega} (1 - m^* \circ \tilde{\phi}_i)(u_i - \tilde{c}_i^{\text{ext}})^2 + m^* \circ \tilde{\phi}_i (u_i - \tilde{c}_i^{\text{in}})^2 dx \\ &\leq |S_i| \left( 1 - \frac{|S_i|}{|\Omega|} \right) - \frac{4\pi\delta^2\epsilon^2}{\text{diam}(\Omega)^2} \left( 1 - \frac{|S_i|}{|\Omega|} \right)^2 \end{aligned}$$

and finally observe that replacing on the minimizing sequence  $\phi_i^k$  by  $\tilde{\phi}_i$  the infimum of the energy  $\underline{\mathcal{E}}$  can be further reduced, which is a contradiction. Hence, the above boundedness of the sequence of transformation vectors  $(\Phi^k)$  is established.

*Step 5.* Finally, we are able to prove lower semicontinuity of the total energy  $\mathcal{E}_\gamma$ . Therefore, given the boundedness of the transformations we select another subsequence, as above again denoted  $(m^k, \Phi^k)_{k=1, \dots}$ , such that the  $\phi_i^k$  converge to some  $\phi_i^*$  for all  $i = 1, \dots, n$ . The contrast values  $c_i^{\text{in}}$  and  $c_i^{\text{ext}}$  are continuous functionals with respect to weak-\* convergence of  $m$  in  $BV$  and convergence of  $\phi_i$  as long as  $\int_{\mathbb{R}^2} m(\chi_\Omega \circ (\phi_i^*)^{-1}) dx > 0$  and  $\int_{\mathbb{R}^2} (1 - m)(\chi_\Omega \circ (\phi_i^*)^{-1}) dx > 0$ . This can easily be seen extending the integrals in (7) from  $\Omega$  to  $\mathbb{R}^2$  and applying an integral transform to achieve

$$\begin{aligned} c_i^{\text{ext}}[m, \phi_i] &= \frac{\int_{\mathbb{R}^2} (1 - m)(u_i \chi_\Omega) \circ \phi_i^{-1} dx}{\int_{\mathbb{R}^2} (1 - m)(\chi_\Omega \circ \phi_i^{-1}) dx}, \\ c_i^{\text{in}}[m, \phi_i] &= \frac{\int_{\mathbb{R}^2} m(u_i \chi_\Omega) \circ \phi_i^{-1} dx}{\int_{\mathbb{R}^2} m(\chi_\Omega \circ \phi_i^{-1}) dx}. \end{aligned}$$

Here we have taken into account that  $\det D\phi_i = 1$  and  $\phi_i(\mathbb{R}^2) = \mathbb{R}^2$ . Hence,

$$\begin{aligned} & \int_{\Omega} (1 - m^k \circ \phi_i^k)(u_i - c_i^{\text{ext},k})^2 + m^k \circ \phi_i^k(u_i - c_i^{\text{in},k})^2 \\ & \rightarrow \int_{\Omega} (1 - m^* \circ \phi_i^*)(u_i - c_i^{\text{ext},*})^2 + m^* \circ \phi_i^*(u_i - c_i^{\text{in},*})^2 \end{aligned}$$

To verify this we again applying the convergence of  $\phi_i$ ,  $c_i^{\text{in}}[m, \phi_i]$ , and  $c_i^{\text{ext}}[m, \phi_i]$  for  $i = 1, \dots, n$  and the strong convergence of  $m$  in  $L^1(\Omega)$ . Let us remark, that from Step 4 we deduce  $\int_{\mathbb{R}^2} m(\chi_{\Omega} \circ (\phi_i^*)^{-1}) dx > 0$ . In case  $\int_{\mathbb{R}^2} (1 - m)(\chi_{\Omega} \circ (\phi_i^*)^{-1}) dx = 0$  continuity of  $c_i^{\text{ext}}[m, \phi_i]$  is not required to prove convergence of the above part of the energy, because  $\int_{\Omega} (1 - m^k \circ \phi_i^k)(u_i - c_i^{\text{ext},k})^2 \rightarrow 0$  for  $k \rightarrow \infty$ . Together with the lower semi continuity of the total variation from Step 2, we finally obtain, that  $\mathcal{E}_{\gamma}[m^*, \Phi^*, \mathcal{C}[m^*, \Phi^*]] \leq \mathcal{E}$ , which proves the minimizing property of  $(m^*, \Phi^*, \mathcal{C}[m^*, \Phi^*])$ .  $\square$

A straightforward modification of the arguments in the above proof allows us to establish an existence result also in the simpler situation of the variational formulation in (5). Here the domain  $\Omega$  no longer plays an essential role, as long as it contains a sufficient large neighborhood of all input shapes. Indeed, for  $S$  compactly contained in  $B_{R_0} := B_{R_0}(0)$  and  $R > R_0$  we achieve  $\text{Per}(S, B_R) := |D\chi_S|(B_R) = \text{Per}(S, B_{R_0})$ . Thus, we also do not need the second and technical non degeneracy assumption on the input shapes. Hence, we obtain the following theorem.

**Theorem 2 (Existence of a minimizer for the BV model (5))** *Suppose that a set of input shapes  $S_1, \dots, S_n \subset\subset B_{R_0} \subset B_R = \Omega$  with finite perimeter is given. These shapes are assumed to be non degenerate in the sense that there exist a fixed constant  $\delta > 0$  and a set of points  $y_1, \dots, y_n$  such that  $B_{\delta}(y_i) \subset S_i$  for  $i = 1, \dots, n$ . Then, for  $R_0$  fixed and  $R$  sufficiently large there exists a non degenerate median shape  $M^* = \{x \in \mathbb{R}^2 \mid (m^*)^+(x) = 1\}$  for  $m^* \in BV(B_R)$  and a vector of affine transformations  $\Phi^*$  such that  $(m^*, \Phi^*)$  is a minimizer of the energy  $\mathcal{E}_{\gamma}$  defined in (5) in the class of functions  $\{m \in BV(B_R) \mid m(x) \in \{0, 1\} \text{ a. e.}\}$  and vectors of area preserving affine transformations  $\Phi = (\phi_i)_{i=1, \dots, n}$ . For sufficiently large radius  $R$  the median shape  $M^*$  is independent of this radius.*

## 5 Numerical Discretization

In what follows, we will derive a numerical algorithm for the minimization of the functional given in (6) along the lines of the approach by Chan and Vese [8]. In fact, we replace in the variational approach (6) the

BV function  $m$  by a level set function  $\zeta$  assuming that  $m(x) = 1$  for  $\zeta < 0$ , else  $m(x) = 0$ . Hence, the corresponding median shape  $M$  is given as the sub level set, i. e.  $M = \{x \in \Omega \mid \zeta(x) < 0\}$  on the actual computational domain  $\Omega$ . Utilizing the Heaviside function  $H$ , defined as  $H(z) = 1$  for  $z > 0$  and 0 elsewhere, one can rewrite the characteristic function of the pull back of the domain  $M$  under the transformation  $\phi$  as  $\chi_{(\phi_i)^{-1}(M)} = 1 - H(\zeta \circ \phi_i)$ . Furthermore, the perimeter of the median shape is equal to the total variation of  $H(\zeta)$ , i. e.  $\text{Per}(M) = |DH \circ \zeta|(\Omega)$ .

### 5.1 A regularized functional

With respect to a numerical relaxation of the correspondingly rewritten functional we have to regularize the Heaviside function as usual. In fact, we consider  $H_{\delta}(z) = \frac{1}{2} + \frac{1}{\pi} \arctan(\frac{z}{\delta})$  for a given small scale parameter  $\delta > 0$ . Let us emphasize that the desired guidance of the initial zero level line to the boundary of the median shape relies on the nonlocal support of this regularized Heaviside function (cf. [7]). Furthermore, we regularize the absolute value used in the total variation term by  $|x|_{\rho} = \sqrt{x^2 + \rho^2}$ , to ensure classical differentiability of the resulting functional. Throughout this paper we have chosen  $\rho = 0.1$ . We end up with the following approximation of the energy  $\mathcal{E}_{\gamma}$ :

$$\begin{aligned} \mathcal{E}_{\gamma}^{\delta, \rho} [(c_i^{\text{in}}, c_i^{\text{ext}})_i, (\phi_i)_i, \zeta] &= \gamma \int_{\Omega} |\nabla H_{\delta}(\zeta)|_{\rho} dx \\ &+ \frac{1}{n} \sum_{i=1}^n \mathcal{E}_i^{\delta} [c_i^{\text{in}}, c_i^{\text{ext}}, \zeta \circ \phi_i] \end{aligned}$$

where

$$\begin{aligned} \mathcal{E}_i^{\delta} [c_i^{\text{in}}, c_i^{\text{ext}}, \xi] &= \int_{\Omega} (1 - H_{\delta}(\xi))(u_i - c_i^{\text{in}})^2 \\ &+ H_{\delta}(\xi)(u_i - c_i^{\text{ext}})^2 dx. \end{aligned}$$

For the spatial discretization of the level set function  $\zeta$  we consider bilinear Finite Elements on a regular quadrilateral mesh. Furthermore, each pixel value of the images  $u_i$  corresponds to a node value of the Finite Element representation of the level set function  $\zeta$ . Thus, we are finally led to a finite dimensional variational problem whose numerical minimization will be discussed in the next paragraph.

### 5.2 The relaxation algorithm

Following [8] we propose an alternating minimization algorithm for the involved unknowns, namely the contrast values  $c_i^1, c_i^2$  for the joint segmentation problem

on the images  $u_i$  for  $i = 1, \dots, n$ , the transformation parameters  $q_i$  parametrizing the deformations  $\phi_i$ , and the vector of nodal values of the level set function  $\zeta$  describing the median shape. Let us first discuss how to minimize  $\mathcal{E}_\gamma^{\delta, \rho}$  with respect to the different degrees of freedom separately.

*Updating the contrast values* The energy  $\mathcal{E}_\gamma^{\delta, \rho}$  is quadratic in  $c_i^{\text{ext}}$  and  $c_i^{\text{in}}$ . Thus, as above in Section 4 we can directly compute the minimizing contrast values by the update formulae

$$\begin{aligned} (c_i^{\text{ext}})^{k+1} &= \frac{\int_\Omega H_\delta(\zeta^k \circ \phi_i^k) u_i \, dx}{\int_\Omega H_\delta(\zeta^k \circ \phi_i^k) \, dx}, \\ (c_i^{\text{in}})^{k+1} &= \frac{\int_\Omega (1 - H_\delta(\zeta^k \circ \phi_i^k)) u_i \, dx}{\int_\Omega (1 - H_\delta(\zeta^k \circ \phi_i^k)) \, dx} \end{aligned} \quad (8)$$

for a given level set function  $\zeta^k$  and deformations  $\phi_1^k, \dots, \phi_n^k$ . Thereby, we trivially extend  $\zeta^k$  to some positive value outside  $\Omega$ .

*Gradient descent* The Euler-Lagrange equations for  $\zeta$  and  $q_i$ , respectively, are nonlinear. Thus, we consider a step size controlled gradient descent in these degrees of freedom. Thus, we consider a step size controlled gradient descent in these degrees of freedom. As step size control we apply Armijo's rule with widening [5]: Given an energy  $E$ , a position  $x$  and a descent direction  $d$ , we choose  $\tau$  such that

$$E(x) - E(x + \tau d) \geq -\frac{1}{2} \tau \nabla E(x) \cdot d.$$

This ensures that at least half of the energy decay predicted by  $\nabla E$  is realized by the chosen step size.

The first variation with respect to the level set function  $\zeta$  in a direction  $\vartheta$  turns out to be:

$$\begin{aligned} \partial_\zeta \mathcal{E}_\gamma^{\delta, \rho} [(c_i^{\text{in}}, c_i^{\text{ext}})_i, (q_i)_i, \zeta](\vartheta) &= \int_\Omega \nabla(H'_\delta(\zeta)\vartheta) \cdot \frac{\nabla \zeta}{|\nabla \zeta|_\rho} \, dx + \\ &\frac{1}{n} \sum_{i=1}^n \int_\Omega \left\{ [(u_i(\phi_i^{-1}) - c_i^{\text{ext}})^2 - (u_i(\phi_i^{-1}) - c_i^{\text{in}})^2] \right. \\ &\quad \left. \chi_{\Omega \circ \phi_i^{-1}} |\det D\phi_i^{-1}| H'_\delta(\zeta)\vartheta \right\} \, dx. \end{aligned}$$

Here  $D\phi_i$  is the Jacobian of the deformation  $\phi_i$ . Let us remark, that using integration by parts the first term can be rewritten as  $-\int_\Omega H'_\delta(\zeta) \operatorname{div} \left( \frac{\nabla \zeta}{|\nabla \zeta|_\rho} \right) \vartheta \, dx$  considered to be the  $L^2$  product of a weighted, regularized mean curvature and  $\vartheta$ . In the finite element implementation this can be encoded in terms of a nodal vector for this weighted, regularized mean curvature. The

first variation of the energy with respect to one of the parametrization vectors  $q_j$  is given by:

$$\begin{aligned} \partial_{q_j} \mathcal{E}_\gamma^{\delta, \rho} [(c_i^{\text{in}}, c_i^{\text{ext}})_i, (q_i)_i, \zeta] &= \frac{1}{n} \int_\Omega \left\{ H'_\delta(\zeta \circ \phi_j) \nabla \zeta(\phi_j) \cdot D_{q_j} \phi_j \right. \\ &\quad \left. ((u_j - c_j^{\text{ext}})^2 - (u_j - c_j^{\text{in}})^2) \right\} \, dx. \end{aligned}$$

We refer to Section 2 for explicit parametrizations of different classes of transformations.

Inspired by the Sobolev active contour approach [32], the descent step in  $\zeta$  is based on a regularizing metric

$$g(\vartheta_1, \vartheta_2) = \int_\Omega \vartheta_1 \vartheta_2 + \frac{\sigma^2}{2} \nabla \vartheta_1 \cdot \nabla \vartheta_2 \, dx$$

on variations  $\vartheta_1, \vartheta_2$  of the level set function, where  $\sigma$  represents a filter width of the corresponding time discrete and implicit heat equation filter kernel. Let us emphasize that the resulting regularized descent does not affect the energy landscape itself, but solely the descent path towards the set of minimizers.

*Initialization* In our numerical examples the initialization of the process turned out to be usually not that important. An exception is the case of shapes with holes where the overall procedure needs some care with respect to the choice of the initial level set function (cf. the discussion below). Usually we choose  $\zeta$  as the signed distance function of a circle. The  $q_i$  are initialized such that  $\phi_i = \phi(q_i, \cdot)$  is the identity, and finally the initial value for  $(c_i^{\text{in}}, c_i^{\text{ext}})_i$  is obtained by applying (8) using the initial values for  $\zeta$  and  $q_i$ .

*Constraints* By definition the averaged shape is described up to a transformation from the corresponding transformation class  $T$ . Explicitly, if  $[(c_i^{\text{in}}, c_i^{\text{ext}})_i, (\phi_i)_i, \zeta]$  is a minimizer, then due to the invariance property  $[(c_i^{\text{in}}, c_i^{\text{ext}})_i, (\psi \circ \phi_i)_i, \zeta \circ \psi^{-1}]$  also is a minimizer for all  $\psi \in T$ . To get rid of this ambiguity in the descent algorithm, we have to select a suitable representative transformation. Thus, we constrain the center of mass of the median shape to the center point  $x_\Omega$  of  $\Omega$  by imposing the constraint

$$\int_\Omega (1 - H_\delta(\zeta(x))) x \, dx = x_\Omega.$$

The constraint is imposed by translating  $\zeta$  such that the constraint is fulfilled and adjusting  $(\phi_i)_i$  accordingly after each gradient descent step. There are further ambiguities with respect to the different classes of transformations. We take care of them via the following additional linear constraints:

- the sum of the scaling parameters has to equal the number of images,
- the sum of the shearing values has to be 0,
- the sum of the rotation angles has to be 0.

*Multiscale minimization* The energy landscape is fairly complicated in basically all non trivial applications. In order not to get stuck in local minima we propose to apply a multiscale strategy. As scale parameter we consider the regularization parameter  $\sigma$  in the definition of the metric  $g(\cdot, \cdot)$ . Initially, we choose a fairly large value for  $\sigma$ , i.e.  $\sigma = 1.0$ . Then, during the gradient descent we successively refine  $\sigma$ . Furthermore, we proceed similarly with the regularization parameter  $\delta$  of the Heaviside function. This implies that in early stages of the algorithm far reaching contour interaction takes place, whereas in later stages the perimeter functional is effectively approximated.

*Shapes with inner contours* Due to the lack of convexity of  $\mathcal{E}_{\gamma}^{\delta, \rho}$  with respect to the level set function  $\zeta$ , the reliable computation of median shapes that contain inner contours is difficult and it may be necessary to take into account a two pass multi scale relaxation approach in order to capture the inner contours. For details on this we refer to [4]. Furthermore, let us remark that the generation of inner contours in shape optimization is associated with the notion of a topological derivative [30].

Here, we confine to showing an example for the computation of a median shape with inner contours that in fact did not need any special energy relaxation strategy, cf. Figure 9.

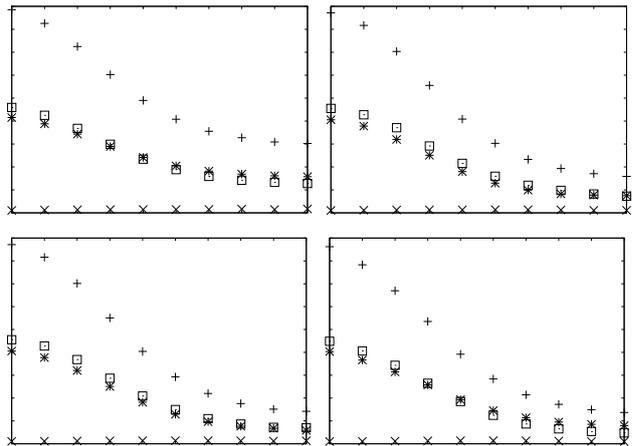


**Fig. 9** Averaging shapes “B” and “8”: The first and second image are the input images, the third one renders  $|u_1 \circ (\phi_1)^{-1} - u_2 \circ (\phi_2)^{-1}|$  overlaid with the zero line of  $\zeta$ , and the fourth depicts the actual median shape with respect to the transformation group  $T_{SL(2)}$ .

Figure 10 shows the decay of the different energy contributions in the algorithm for the example in Figure 2.

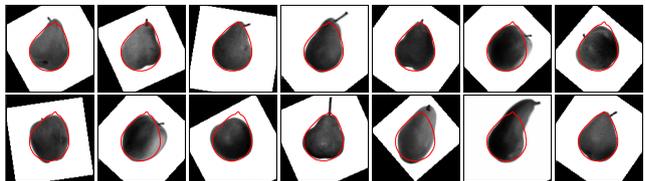
### 5.3 Further numerical results

Figures 11 and 12 render the average of 15 images with nine pears and five apples as input shapes (cf. Figure 1).



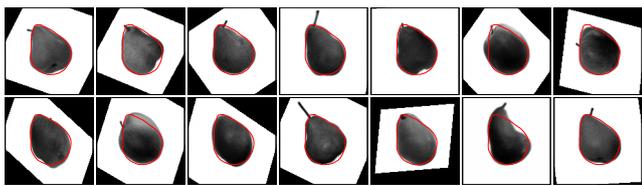
**Fig. 10** The development of the different energy contributions is plotted over the descent steps of the minimization algorithm. The “+”s represent the total energy, whereas the “□”s and “\*”s indicate the shape distance from the first and the second shape, respectively. Finally, the “x”s show the scaled boundary length of the computed average shape. From left to right and from top to bottom the diagrams correspond to the group of translations  $T_{trans}$  (first), concatenation of rigid body motions and isotropic scaling  $T_{T,SO(2)}$  (second), concatenations of rigid body motions, shearing and isotropic scaling  $T_{T,shear}$  (third), and the transformation group  $T_{SL(2)}$  (fourth).

In the first figure, we consider the transformation group  $T_{T,SO(2)}$  consisting of rigid body motion and isotropic scaling. The images show the deformations  $\phi_i$  acting on the domain  $\Omega = [0, 1]^2$  and visualize the difference between the median shape  $M$  and the deformed shapes  $\phi_i(S_i)$ . In Figure 11 the transformation class consists just of rigid body motions and scaling, whereas in Figure 12 also shear is taken into account. A comparison of the two figures underlines the impact of shear on the symmetric difference between the deformed shapes and the median shape - effectively sheared apples are “almost” of pear shape. Furthermore in Figure 13 the corresponding result for the invariance group  $T_{SL(2)}$  is depicted. The required CPU time increases with increas-

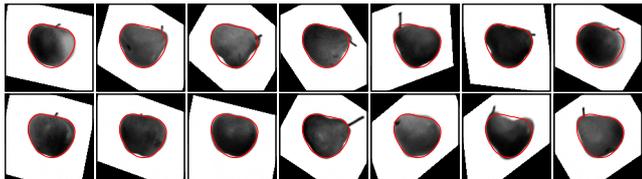


**Fig. 11** Transformed images  $u_i \circ (\phi_i)^{-1}$  overlaid with the zero line of the level set function  $\zeta$  in case of the transformation class  $T_{T,SO(2)}$  consisting of the concatenation of scaling and rigid body motion.

ing degrees of freedom in the transformation class. On a single processor Pentium 4 with 3.6 GHz and 2 GByte memory the CPU time of the non performance opti-



**Fig. 12** Transformed images  $u_i \circ (\phi_i)^{-1}$  overlaid with the zero line of  $\zeta$  where the transformation class  $T_{\tau, \text{shear}}$  based on a concatenation of scaling, rigid body motion, and shearing is taken into account.



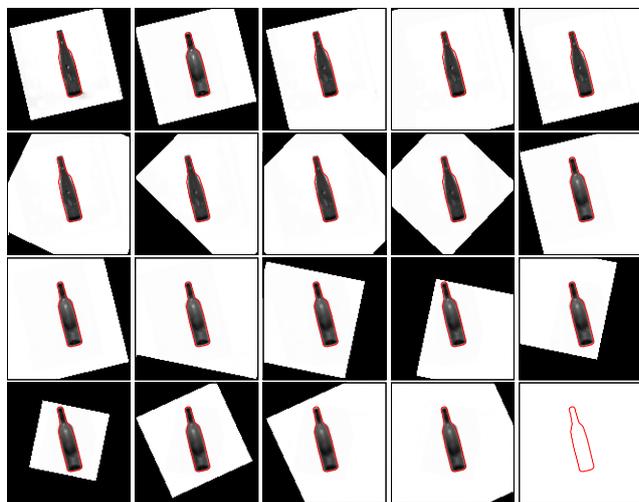
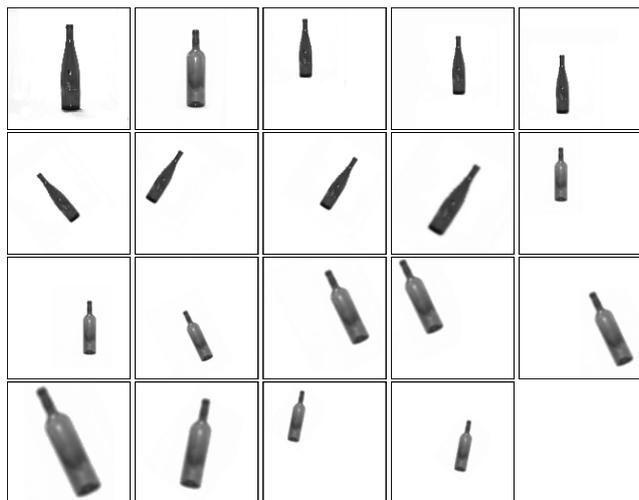
**Fig. 13** Transformed images  $u_i \circ (\phi_i)^{-1}$  overlaid with the zero line of  $\zeta$  where the underlying transformation group is  $T_{\text{SL}(2)}$ .

mized code ranges from 103 to 121 minutes in this application. The computational complexity basically scales linear in the number of input shapes.

Finally, we consider a large set ( $n = 19$ ) of digital photographs of different wine bottles (namely Burgunder and Riesling bottles). Figure 14 renders the 19 images we have used as input and also depicts the resulting median shape and the correspondingly transformed photographs. Here the transformation group is  $T_{\tau, \text{SO}(2)}$  consisting of rigid body motions and isotropic scaling.

## 6 Conclusions

We have presented a shape averaging method for two dimensional shapes based on the symmetric area difference and a notion of a corresponding median shape which factors out different classes of affine transformations. To avoid arbitrarily irregular median contours a suitable regularization is investigated. A combination with a binary Mumford–Shah type segmentation model allows the application on shapes described via images. In this context, existence of a median shapes is proven for  $T_{\text{SL}(2)}$  as the underlying group of transformations. The developed approach has been implemented via a level set method and various qualitative effects of the median approach are studied for different sets of input shapes and different transformation invariance classes. The focus of this paper is surely on the conceptual discussion of the model and its mathematical foundation. The combination of our averaging model with recent global optimization approaches in image segmentation [14] would probably significantly speed up the algorithm. Furthermore, a rigorous investigation of the shape median as minimizer of (4) in the limit



**Fig. 14** A set of input images of wine bottles is rendered in the first four rows. The transformed images  $u_i \circ (\phi_i)^{-1}$  overlaid with the zero line of  $\zeta$  are shown underneath. The bottom right image represents the median shape. The underlying transformation group used here is  $T_{\tau, \text{SO}(2)}$ .

$\gamma \rightarrow 0$  is still open.

## Acknowledgements.

The authors acknowledge support by the Hausdorff Center for Mathematics at Bonn University. Furthermore, Benjamin Berkels was supported by the German Science foundation, in particular via the priority program 1114.

## References

1. L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathe-

- mathematical Monographs. Oxford University Press, New York, 2000.
2. B. Avants and J. C. Gee. Geodesic estimation for large deformation anatomical shape averaging and interpolation. *NeuroImage*, 23:139–150, 2004. Supplement 1.
  3. M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, February 2005.
  4. B. Berkels, G. Linkmann, and M. Rumpf. A shape median based on symmetric area differences. In O. Deussen, D. Keim, and D. Saupe, editors, *Vision, Modeling, and Visualization Proceedings*, pages 399–407, 2008.
  5. D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 2nd edition, 1999.
  6. K. K. Bhatia, J. V. Hajnal, B. K. Puri, A. D. Edwards, and D. Rueckert. Consistent groupwise non-rigid registration for atlas construction. In *IEEE International Symposium on Biomedical Imaging: Nano to Macro*, volume 1, pages 908–911, 2004.
  7. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.
  8. T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.
  9. G. Charpiat, O. Faugeras, and R. Keriven. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Foundations of Computational Mathematics*, 5(1):1–58, 2005.
  10. G. Charpiat, O. Faugeras, R. Keriven, and P. Maurel. Distance-based shape statistics. In *Acoustics, Speech and Signal Processing, 2006 (ICASSP 2006)*, volume 5, 2006.
  11. S. E. Chen and R. E. Parent. Shape averaging and its applications to industrial design. *IEEE Computer Graphics and Applications*, 9(1):47–54, 1989.
  12. T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models—their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.
  13. D. Dupuis, U. Grenander, and M. Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*, 56:587–600, 1998.
  14. P. F. Felzenszwalb. Representation and detection of deformable shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(2):208–220, February 2005.
  15. P. T. Fletcher, C. Lu, and S. Joshi. Statistics of shape via principal geodesic analysis on Lie groups. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR*, volume 1, pages 95–101, 2003.
  16. T. Fletcher, S. Venkatasubramanian, and S. Joshi. Robust statistics on Riemannian manifolds via the geometric median. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.
  17. M. Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. H. Poincaré*, 10:215–310, 1948.
  18. M. Fuchs, B. Jüttler, O. Scherzer, and H. Yang. Shape metrics based on elastic deformations. *J. Math. Imaging Vis.*, 35(1):86–102, 2009.
  19. X. Jiang, K. Abegglen, H. Bunke, and J. Csirik. Dynamic computation of generalised median strings. *Pattern Analysis and Applications*, 6(3):185–193, December 2003.
  20. X. Jiang, L. Schiffmann, and H. Bunke. Computation of median shapes. In *Proceedings of the Fourth IEEE Asian Conference on Computer Vision (ACCV'00)*, pages 300–305, 2000.
  21. S. Joshi, B. Davis, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, 23:151–160, 2004. Supplement 1.
  22. H. Karcher. Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, 30(5):509–541, 1977.
  23. D. G. Kendall. Shape manifolds, procrustean metrics, and complex projective spaces. *Bull. London Math. Soc.*, 16:81–121, 1984.
  24. M. Miller, A. Trouvé, and L. Younes. On the metrics and Euler-Lagrange equations of computational anatomy. *Annual Review of Biomedical Engineering*, 4:375–405, 2002.
  25. M. I. Miller and L. Younes. Group actions, homeomorphisms, and matching: a general framework. *International Journal of Computer Vision*, 41(1–2):61–84, 2001.
  26. D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Communications on Pure Applied Mathematics*, 42:577–685, 1989.
  27. D. Rueckert, A. F. Frangi, and J. A. Schnabel. Automatic construction of 3D statistical deformation models using non-rigid registration. In W. Niessen and M. Viergever, editors, *Medical Image Computing and Computer-Assisted Intervention, MICCAI*, volume 2208 of *LNCS*, pages 77–84, 2001.
  28. M. Rumpf and B. Wirth. An elasticity approach to principal modes of shape variation. In *Proceedings of the Second International Conference on Scale Space Methods and Variational Methods in Computer Vision (SSVM 2009)*, volume 5567 of *Lecture Notes in Computer Science*, pages 709–720, 2009.
  29. M. Rumpf and B. Wirth. A nonlinear elastic shape averaging approach. *SIAM Journal on Imaging Sciences*, 2(3):800–833, 2009.
  30. J. Sokolowski and A. Zochowski. Optimality conditions for simultaneous topology and shape optimization. *SIAM J. Control Optim.*, 42:1198–1221, 2003.
  31. J. Sokolowski and J.-P. Zolésio. *Introduction to shape optimization*. Springer-Verlag, Berlin, 1992. Shape sensitivity analysis.
  32. G. Sundaramoorthi, J. D. Jackson, A. Yezzi, and A. C. Menon. Tracking with Sobolev active contours. In *CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 674–680, Washington, DC, USA, 2006. IEEE Computer Society.
  33. B. Wirth, L. Bar, M. Rumpf, and G. Sapiro. Geodesics in shape space via variational time discretization. In *EMM-CVPR'09, to appear*, 2009.
  34. A. Yezzi and S. Soatto. Deformation: Deforming motion, shape average and the joint registration and approximation of structures in images. *International Journal of Computer Vision*, 53(2):153–167, July 2003.