A Shape Median Based on Symmetric Area Differences

Benjamin Berkels, Gina Linkmann, Martin Rumpf

Institut für Numerische Simulation, Universität Bonn
Email: {benjamin.berkels, gina.linkmann, martin.rumpf}@ins.uni-bonn.de

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 is based on the area of the symmetric difference of shapes and takes into account different invariance classes. These classes are generated by classical transformation groups such as translation, rotation, anisotropic scaling, and shear. As in the finite dimensional case, non-uniqueness of the median is observed. The numerical approximation of shape medians is based on a level set approach for the description of 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 behavior of the method.

1   Introduction

In this paper we develop a notion of a median of shapes in 2D. Median averaging is renown as a particularly powerful and robust concept on a set of vectorial data in finite dimensions. It can be phrased in terms of an optimization problem. Given $x_1, ..., x_n \in \mathbb{R}^d$, the Euclidean median $m^* = m^*[x_1, ..., x_n]$ is defined as

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

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.

Chen and Parent [6] investigated averages of 2D contours already in 1989. Jiang et al. [9] 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. [8]. Furthermore, the computation of shape distances naturally appears when matching shapes in images. Yezzi and Soatto [16] have investigated shape averages in image structure reconstruction and joint registration. Beg et al. [3] introduced a geodesic
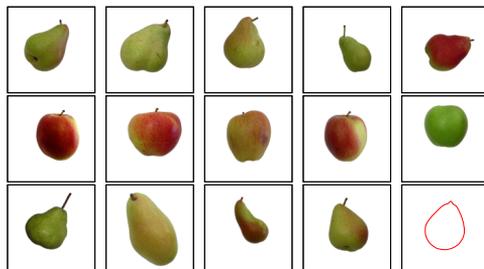
Figure 1: Given nine pears and five apples as input shapes, the median shape (lower right) is computed by the proposed method. The underlying definition of the median includes the invariance with respect to rigid body motions.

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 among others for in-

stance by Joshi et al. [10] and Avants and Gee [2] in comparative medical anatomy. In the simplified case of planar curves Mumford and Michor [12] developed a Riemannian calculus. By definition, the overall problem falls into the class of shape optimization. We refer to the book by Sokołowski and Zolésio [14] 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 consider a joint approach based on a Mumford Shah type variational formulation [13] for the average as a free discontinuity set. For the implementation we pick up a robust and efficient approximation of the Mumford–Shah functional presented by Chan and Vese [5] in the context of piecewise constant image segmentation. As a distance measure between shapes we consider the area of the symmetric difference between sets and regard shapes as 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. A general outline of infinite dimensional diffeomorphic group actions on shapes is discussed by Dupuis et al. [7]. In the underlying numerical approximation the average shape is represented as the zero level set of a level set function. When averaging n shapes the degrees of freedom 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. We give various examples for the qualitative behavior and the overall performance of the resulting algorithm.

2   An area based distance measure

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

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

One easily verifies that for given sets A, B, C and $x \in A \triangle B$ either $x \in A \triangle C$ or $x \in C \triangle B$.

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 $d(A, B) \geq 0$ ($d(A, B) = 0 \Leftrightarrow A = B$) it follows that $d(\cdot, \cdot)$ is a metric on subsets of $\mathbb{R}^2$. The actual comparison of two shapes 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 $\phi : \mathbb{R}^2 \to \mathbb{R}^2$. Examples to be considered later are translations, rigid body motions, or area preserving scalings. Given a particular transformation group T we finally define a distance between shapes represented by sets A, $B \subset \mathbb{R}^2$ with respect to T by

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

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 this metric on the space of 2D shapes can be represented in integral form

$$d_T(A, B) = \inf_{\phi \in T} \int_{\mathbb{R}^2} (\chi_{\phi \circ A} - \chi_B)^2 \, \mathrm{d}x \,,$$

where χ_C denotes the characteristic function of a set C.

3   Transformation groups

In this paper we confine to subgroups of the group of affine transformations ϕ 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_{trans} with $A = \mathbb{1}$. As usual the group of rotations

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

is denoted by $SO(2)$ so that we obtain the group of rigid body motions T_{rigid} via $\phi(x) = Ax + b$ with $A \in SO(2)$. The largest area preserving group in the set of affine transformations is described by matrices $A \in SL(2)$ with $\det A = 1$. Based on the polar decomposition formula they can be parametrized as follows

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

with $Q \in SO(2)$. Shearing in x direction is a one dimensional subgroup of $SL(2)$ given by matrices $A = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$. Anisotropic area preserving scaling is given by matrices $A = Q^T \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} Q$ with $\alpha \in \mathbb{R}^+$ and $Q \in SO(2)$. With respect to the variational model for a shape median to be developed in this paper, we will take into account different groups of transformations and relax the definition, restricting ourselves not to only subgroups but to just subsets, such as the set of concatenations of rotation and shear, cf. Section 7. 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 an 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 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.

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. All of our invariance classes T will contain translations, since translation invariance is fundamental in all applications.

Figure 2: Averaging of two shapes, "1" and "7", with different choices of the transformation group T: The first and second image show the two input images, the third to fifth one the outline of the average shape in case of the transformation class of simple translations (third), rigid body motions combined with isotropic scaling (fourth), and a combination of translation, rotation, isotropic scaling, and shearing (fifth), respectively.
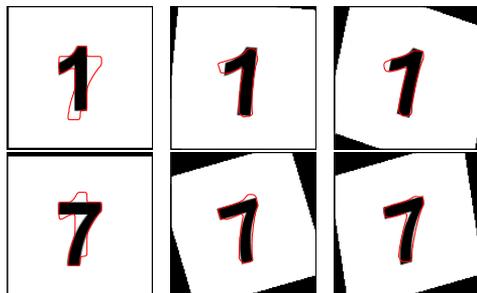
Figure 3: $\phi_i(A_i)$ overlaid with the median shape contour with T allowing translation (first column), translation, rotation, and isotropic scaling (second column), and translation, rotation, isotropic scaling, and shearing (third column).

4 A variational definition of the shape median

Recalling the definition of the median in the context of vectors in $\mathbb{R}^n$, a first definition of a median M^* of shapes $A_1, \cdots, A_n \subset \mathbb{R}^2$ and a set T of invariant transformations is given by

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

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 $A_1 = [0, a]^2$ and a rectangle $A_2 = [0, a] \times [0, b]$ with $a < b$. Obviously, every rectangle $M = [0, a] \times [0, c]$ with $a \leq c \leq b$ renders the sum of distances $\sum_{i=1}^2 d_T(M, A_i)$ to be $(b - a)a$, which is equal to $d_T(A_1, A_2)$ and thus the minimal value. Unfortunately, the set of min-
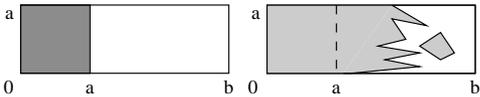
Figure 4: Without the boundary length penalization shape medians turn out to be fairly irregular.

imizers is even larger. We can split the rectangle $[0, a] \times [a, b]$ into any two disjoint measurable subsets $\tilde{A}_1$, $\tilde{A}_2$ and $M = A_1 \cup \tilde{A}_1$ 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 $\mathcal{M}[A_1, \cdots, A_n]$ 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 least perimeter (boundary length) and define this set as the shape median, i.e.

$$M^* = \operatorname*{argmin}_{M \in \mathcal{M}[A_1, \cdots, A_n]} \operatorname{Per}(M),$$

where $\operatorname{Per}(A)$ denotes the length of the boundary ∂A of A if ∂A is rectifiable and is set to ∞ else. For
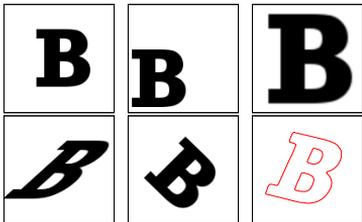
Figure 5: On the bottom right the average of five "B" with T allowing rotation, translation, isotropic scaling, and shearing.

the actual numerical computation we simplify the definition and consider a γ–median shape, defined as the minimizing set M^* of the energy

$$E_\gamma[M] = \frac{1}{n} \sum_{i=1}^n d_T(M, A_i) + \gamma \operatorname{Per}(M),$$

where γ is a small constant. For the applications considered in this paper we have choosen $\gamma = 0.0005$. Figure 5 depicts the resulting shape average in case of a letter represented in different fonts. Due to the 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 $\phi_1, \cdots, \phi_n \in T$ appearing in $d(M, \phi_i(A_i))$. Thus, we are led to a joint minimization problem for the functional

$$E_\gamma[\phi_1, \cdots, \phi_n, M] = \frac{1}{n} \sum_{i=1}^n d(M, \phi_i(A_i))$$
$$+ \gamma \operatorname{Per}(M).$$

As already mentioned in the introduction, in practical applications input data usually are not shapes $A_1, \cdots, A_n$, but images $u_1, \cdots, u_n : \Omega \to \mathbb{R}$, each of them encoding a shape. Throughout this paper, the domain Ω is considered to be the unit square $[0, 1]^2$. To phrase the shape median in terms of these images, we consider a simultaneous segmentation of the input images based on a piecewise constant Mumford–Shah model. Hence, we extend our energy to

$$E_\gamma[(\phi_i)_i, M, (c_i^1)_i, (c_i^2)_i] = \gamma \operatorname{Per}(M)$$
$$+ \frac{1}{n} \sum_{i=1}^n \left(\int_{(\phi_i)^{-1}(M)} (u_i - c_i^1)^2 \, \mathrm{d}x \right. \tag{1}$$
$$\left. + \int_{(\phi_i)^{-1}(\Omega \setminus M)} (u_i - c_i^2)^2 \, \mathrm{d}x \right).$$

Here, a set of parameters $(\phi_i)_i, M, (c_i^1)_i, (c_i^2)_i$ segments the image u_i into a piecewise constant function with intensity c_i^1 on the pull back $(\phi_i)^{-1}(M)$ of the shape M and c_i^2 on the pull back $(\phi_i)^{-1}(\Omega \setminus M)$ of the complement of the shape M under the transformation ϕ_i. Hence, the values c_i^1, c_i^2 are additional degrees of freedom in our model, which compensate for variations in the constrast of the shapes described in the images u_i for $i = 1, \cdots, n$. Figure 6 shows the intensity values c_i^1 and c_i^2 that minimize the energy $E_\gamma[\cdot]$ in the introductory example from Figure 1. Figure 7 demonstrates how constrast modulation in the given shapes might effect the resulting shape average.

5 Level set formulation and numerical approximation

In what follows, we will derive a numerical algorithm for the minimization of the functional given in (1) along the lines of the approach by Chan and Vese [5]: We represent the unknown shape median M via a level set function ζ, i.e.

$$M = \{x \in \Omega \,|\, \zeta(x) < 0\},$$
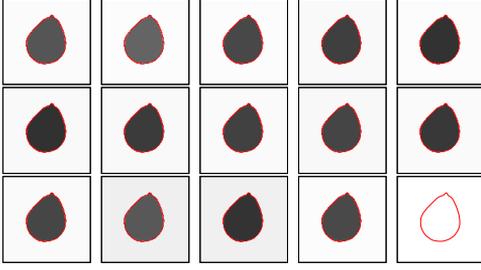
Figure 6: Median shape contour (red) and c_i^1, c_i^2 for the averaging problem given in Figure 1.

Figure 7: Strong variations of contrast in a single shape might lead to wrong matching between the average shape and this shape. One of the input images from Figure 1 (first) is contrast modulated (second). The corresponding shape deformations overlayed with the different resulting average shapes are depicted on the right (third and fourth) based on the same transformation group as in Figure 11.

where Ω is the actual computational domain. 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 ϕ as $\chi_{(\phi_i)^{-1}(M)} = 1 - H(\zeta \circ \phi_i)$. Thus, the integrals in (1) can be expressed in terms of the Heaviside function:

$$\int_{\phi_i^{-1}(M)} (u_i - c_i^1)^2 \mathrm{d}x = \int_{\Omega} (1 - H(\zeta(\phi_i)))(u_i - c_i^1)^2 \mathrm{d}x$$

$$\int_{\phi_i^{-1}(\Omega \setminus M)} (u_i - c_i^2)^2 \mathrm{d}x = \int_{\Omega} H(\zeta(\phi_i))(u_i - c_i^2)^2 \mathrm{d}x$$

Furthermore, the perimeter of the median shape is equal to the total variation of $H(\zeta)$, i. e. $\mathrm{Per}\,(M) = \int_{\Omega} |\nabla(H(\zeta))|\,\mathrm{d}x$ [1]. Hence, we can reformulate the Mumford–Shah functional in (1) as a BV type functional involving the Heaviside function. With respect to a numerical relaxation of this rewritten functional via a gradient descent method, we have to regularize the Heaviside function. In fact, we consider $H_\delta(z) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{z}{\delta}\right)$ 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. [4]). We end up with the following approximation of energy E_γ:

$$E_\gamma^\delta \left[(c_i^1, c_i^2)_i, (\phi_i)_i, \zeta \right] =$$

$$\gamma \int_{\Omega} |\nabla H_\delta(\zeta)|_\rho \; \mathrm{d}x + \frac{1}{n} \sum_{i=1}^{n} E_i^\delta \left[c_i^1, c_i^2, \zeta \circ \phi_i \right]$$

where

$$E_i^\delta [c_i^1, c_i^2, \xi] = \int_{\Omega} (1 - H_\delta(\xi))(u_i - c_i^1)^2$$
$$+ H_\delta(\xi)(u_i - c_i^2)^2 \; \mathrm{d}x.$$

Note that we also applied a standard regularization of the absolute value used in the total variation term, i.e. $|x|_\rho = \sqrt{x^2 + \rho^2}$, to ensure classical differentiability of the resulting functional. Throughout this paper we have chosen $\rho = 0.1$. For the spatial discretization of the level set function ζ we consider bilinear Finite Elements on a regular quadrilateral mesh. Furthermore, each pixel of the images u_i corresponds to a node value of the Finite Element mesh used for ζ. As already discussed, we take into account a parameter vector q_i for every transformation ϕ_i. Thus, we are finally led to a finite dimensional variational problem whose numerical minimization will be discussed in the next section.

6 Algorithm

Following [5] we propose an alternating minimization algorithm for the involved unknowns, namely the constrast values c_i^1, c_i^2 for the joint segmentation problem on the images u_i for $i = 1, \cdots, n$, the transformation parameters q_i parametrizing the deformations ϕ_i, and the vector of nodal values of the level set function ζ describing the averaged shape. Let us first discuss how to minimize E_γ^δ with respect to the different unknowns separately.

Updating the contrast values. We observe that E_γ^δ is quadratic in c_i^j (cf. [5]). Therefore we can directly compute the minimizing contrast values by

the update formulae

$$\left(c_i^1\right)^{k+1} = \frac{\int_\Omega (1 - H(\zeta^k \circ \phi_i^k)) u_i \, dx}{\int_\Omega (1 - H(\zeta^k \circ \phi_i^k)) \, dx}, \quad (2)$$

$$\left(c_i^2\right)^{k+1} = \frac{\int_\Omega H(\zeta^k \circ \phi_i^k) u_i \, dx}{\int_\Omega H(\zeta^k \circ \phi_i^k) \, dx} \quad (3)$$

for a given level set function ζ^k and deformations $\phi_1^k, \cdots, \phi_n^k$.

Usually, $\phi \in T$ does not map from Ω to Ω. Hence, we have to extend functions outside Ω. For the sake of simplicity, we trivially extend all functions by 0.

Gradient descent. The Euler-Lagrange equations for ζ and q_i, respectively, are nonlinear. 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 [11]. The first variation with respect to the level set function ζ in a direction ϑ turns out to be:

$$\partial_\zeta E_\gamma^\rho[(c_i^1, c_i^2)_i, (q_i)_i, \zeta](\vartheta)$$
$$= \frac{1}{n} \sum_{i=1}^n \int_\Omega \Big\{ \left| \det D\phi_i^{-1} \right| H_\delta'(\zeta)\vartheta$$
$$\left[(u_i(\phi_i^{-1}) - c_i^2)^2 - (u_i(\phi_i^{-1}) - c_i^1)^2 \right] \Big\} \, dx$$
$$+ \int_\Omega \nabla(H_\delta'(\zeta)\vartheta) \cdot \frac{\nabla\zeta}{|\nabla\zeta|_\rho} \, dx.$$

Here $D\phi_i$ is the Jacobian of the deformation ϕ_i. The first variation of the energy with respect to one of the parametrization vectors q_j is given by:

$$\partial_{q_j} E_\gamma^\rho[(c_i^1, c_i^2)_i, (q_i)_i, \zeta]$$
$$= \frac{1}{n} \int_\Omega \Big\{ H_\delta'(\zeta \circ \phi_j)\nabla\zeta(\phi_j) \cdot D_{q_j}\phi_j$$
$$\left((u_j - c_j^2)^2 - (u_j - c_j^1)^2 \right) \Big\} \, dx.$$

We refer to Section 7 for explicit parametrizations of different classes of transformations.

Inspired by the Sobolev active contour approach [15], the descent step in ζ 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 ϑ_1, ϑ_2 of the level set function, where σ 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. The initialization of the process is 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 ζ 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 the c_i^j is obtained by applying (2) and (3) using the initial values for ζ and q_i.

Constraints. By definition the averaged shape is described up to the set of transformations T. Explicitly, if $[(c_i^1, c_i^2)_i, (\phi_i)_i, \zeta]$ is a minimizer, then $[(c_i^1, c_i^2)_i, (\psi \circ \phi_i)_i, \zeta \circ \psi^{-1}]$ also is a minimizer for all $\psi \in T$. To get rid of this ambiguity with respect to transformations from the class T, we have to select a suitable representative average shape. Indeed, we constrain the center of mass of the median shape to a fixed center point x_Ω of Ω. In other words, we impose the contraint

$$\int_\Omega (1 - H_\delta(\zeta(x))) x \, dx = x_\Omega. \quad (4)$$

If this is omitted, the numerical algorithm tends to translate the zero line of ζ out of our image domain, essentially resulting in $E = 0$. The ambiguities with respect to the other possibly allowed affine mappings are taken care of via the following 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 σ in the definition of the metric $g(\cdot, \cdot)$. Initially, we choose a fairly large value for σ, i.e. $\sigma = 1.0$. Then, during the gradient descent we successively refine σ. Furthermore, we proceed similarly with the regularization parameter δ 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.
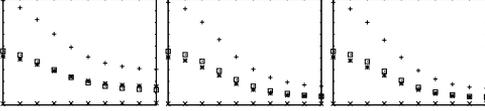
Figure 8: The development of the different energy contributions is plotted over the descent steps of the minimization algorithm. The "+"s represent the total energy, whereas the "□"s and "∗"s indicate the shape distance from the first and the second shape, respectively. Finally, the "×"s show the scaled boundary length of the computed average shape. From left to right the diagrams correspond to the transformation classes depicted already in Figure 2 ranging from pure translation to rigid body motion and shear.

Figure 8 shows the decay of the different energy contributions in the algorithm for the example in Figure 2.

Shapes with inner contours. To reliably find a median shape that contains inner contours, we first use an L_2 gradient flow for ζ instead of the regularized gradient flow (setting $\sigma = 0$) for coarse values of the parameter δ. Thereby one obtains a better initialization of the image segmentation and the transformation parameters. Given this initialization we restart the algorithm, now successively decreasing δ. Still, it appears to be appropriate to consider small values for σ in this case. Figure 9 illustrates the effect of this strategy.
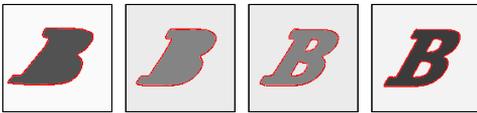
Figure 9: Different iterates of the average shape contour (outlined in red) together with a grey shading according to c_1^j are displayed for the averaging of images 1 and 4 from Figure 5. On the left the computed initial contour is depicted, where the outer shape is already a good approximation of the final average shape contour. The still missing inner contour then appears in a second run of the algorithm involving a successive refinement of the edge sharpness parameter δ. The second and third image show intermediate configurations in the second stage relaxation, whereas the rightmost image represents the final average of the two initial shapes.

7 Parametrized transformations

In what follows we discuss the parametrization of the transformation classes considered in this paper. At first let us define the following matrices,

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

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

with parameters α and β, τ, respectively. In case of rigid body motions we choose $q_i = (\alpha_i, b_i^1, b_i^2)$, where

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

with $b_i = (b_i^1, b_i^2) \in \mathbb{R}^2$. If we enlarge the transformation class by isotropic scaling (skipping the area preservation assumption) we consider $q_i = (s_i, \alpha_i, b_i^1, b_i^2)$ with $\phi_i(q_i, x) = s_i Q(\alpha_i)x + b_i$ and $s_i \in \mathbb{R}$. Anisotropic scaling as introduced in Section 3 has not been taken into account in our computational examples. If we take into account an axially aligned shear in the reference configuration of all images, we apply the parameter vector $q_i = (\beta_i, \gamma_i, \alpha_i, b_i^1, b_i^2)$ describing the transformation

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

Finally, adding again a scaling by the factor s_i as above we obtain a combination of shear, rigid body motion, and isotropic scaling for $\phi_i(q_i, x) = s_i Q(\alpha_i)S(\beta_i, \gamma_i)x + b_i$ and $q_i = (s_i, \beta_i, \gamma_i, \alpha_i, b_i^1, b_i^2)$.

8 Further results

Figures 10 and 11 show the average of 15 images with nine pears and five apples as input shapes (cf. Figure 1). These figures show the deformations ϕ_i acting on the unit square $\Omega = [0, 1]^2$ and demonstrate the difference between the averaged shape M and the deformed shapes $\phi_i(A_i)$. In Figure 10 the transformation class consists just of rigid body motions and scaling, whereas in Figure 11 also shear is taken into account. A comparison of the two figures underlines the impact of shear on the resulting shapes variance (effectively sheared apples are "almost" of pear shape). The required CPU time increases with increasing degrees of freedom in the
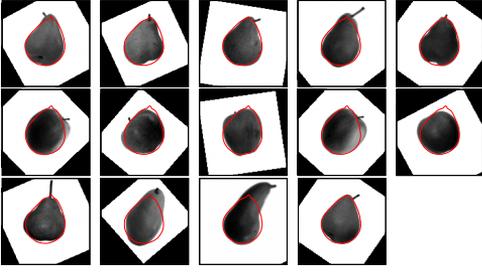
Figure 10: Transformed images $u_i \circ (\phi_i)^{-1}$ overlaid with the zero line of the level set function ζ in case of scaled, rigid body motions as transformation class.
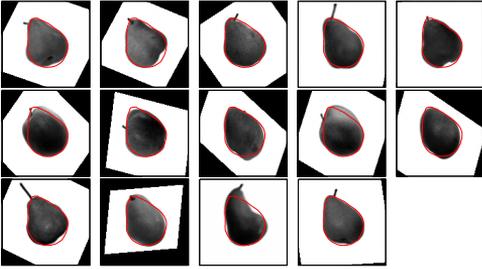
Figure 11: Transformed images $u_i \circ (\phi_i)^{-1}$ overlaid with the zero line of ζ where scaling, rigid body motion, and shear is taken into account.

Figure 12: 19 different wine bottles used as input images.
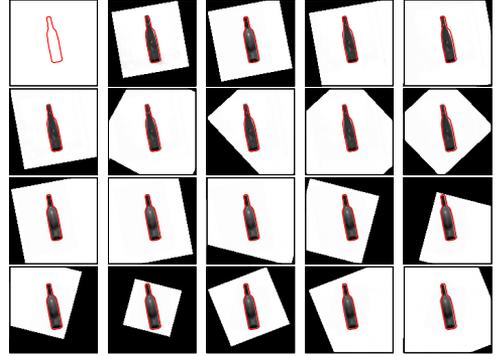
Figure 13: Transformed wine bottle images $u_i \circ (\phi_i)^{-1}$ overlaid with the zero line of ζ.

transformation class. On a single processor Pentium 4 with 3.6 GHz and 2 GByte memory the CPU time of the non performance optimized code ranges from 103 to 121 minutes in this application.

In the next numerical experiment we consider a large set ($n = 19$) of digital photographs of different wine bottles (namely Burgunder and Riesling bottles). Figure 12 shows the 19 images we have used as input, whereas Figure 13 depicts the resulting median shape and the deformed photographs. From this experiment we can see that the proposed model is able to handle a larger number of input images, the only obvious difference to a smaller number of images is a basically linearly increased computing time. Finally, Figure 14 shows the median of two shapes representing the letter "B" and the digit "8", respectively.

References

[1] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, New York, 2000.

[2] B. Avants and J. C. Gee. Geodesic estimation for large deformation anatomical shape averaging and interpolation. *NeuroImage*, 23:S139–S150, 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] V. Caselles, R. Kimmel, and G. Sapiro.

Figure 14: Averaging shapes "B" and "8": The first and second image show the two input images, the third one $\left| u_1 \circ (\phi_1)^{-1} - u_2 \circ (\phi_2)^{-1} \right|$ overlaid with the zero line of ζ, and the fourth one the zero level line of ζ and c_1^j.

Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.

[5] Tony F. Chan and Luminita A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10, no. 2:266 – 277, 2001.

[6] 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.

[7] D. Dupuis, U. Grenander, and M.I. Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*, 56:587–600, 1998.

[8] 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.

[9] 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.

[10] S. Joshi, B. Davis, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, 23:S151–S160, 2004. Supplement 1.

[11] P. Kosmol. *Optimierung und Approximation*. de Gruyter Lehrbuch, 1991.

[12] Peter W. Michor and David Mumford. Riemannian geometries on spaces of plane curves, January 26 2004.

[13] D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Communications on Pure Applied Mathematics*, 42:577–685, 1989.

[14] J. Sokołowski and J-P. Zolésio. *Introduction to shape optimization*. Springer-Verlag, Berlin, 1992. Shape sensitivity analysis.

[15] Ganesh Sundaramoorthi, Jeremy D. Jackson, Anthony Yezzi, and Andrea C. Mennucci. 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.

[16] A. Yezzi and S. Soatto. Deformotion: 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.