

A NONLINEAR ELASTIC SHAPE AVERAGING APPROACH

MARTIN RUMPF[†] AND BENEDIKT WIRTH^{†‡}

NOVEMBER 6, 2009

Abstract. A physically motivated approach is presented to compute a shape average of a given number of shapes. An elastic deformation is assigned to each shape. The shape average is then described as the common image under all elastic deformations of the given shapes, which minimizes the total elastic energy stored in these deformations. The underlying nonlinear elastic energy measures the local change of length, area, and volume. It is invariant under rigid body motions, and isometries are local minimizers. The model is relaxed involving a further energy which measures how well the elastic deformation image of a particular shape matches the average shape, and a suitable shape prior can be considered for the shape average. Shapes are represented via their edge sets, which also allows for an application to averaging image morphologies described via ensembles of edge sets. To make the approach computationally tractable, sharp edges are approximated via phase fields, and a corresponding variational phase field model is derived. Finite elements are applied for the spatial discretization, and a multi-scale alternating minimization approach allows the efficient computation of shape averages in 2D and 3D. Various applications, e.g. averaging the shape of feet or human organs, underline the qualitative properties of the presented approach.

Key words. shape averaging, non-rigid registration, nonlinear elasticity, Mumford–Shah approach, phase field approximation, finite element discretization

AMS subject classifications. 49J45, 65N55, 65M60, 74B20, 68U10

1. Introduction. This paper is concerned with a physically motivated notion of shape averages. As shapes we consider object contours or image morphologies, both encoded as edge sets in images. We propose to employ concepts from nonlinear elasticity which reflect first principles and to define shape averages which incorporate natural local measures of the underlying deformation and its dissimilarity from an isometry. Furthermore, we will introduce a robust numerical approximation based on a diffusive phase field description of shapes.

Averaging is a fundamental task for the quantitative analysis of ensembles of shapes and has already been extensively studied in the literature. Very basic notions of averaging include the arithmetic mean of landmark positions [22], and the image obtained by the arithmetic mean of the matching deformations [57, 7]. For general images there are various qualitatively different notions of averaging. The intention is a fusion or blending to simultaneously represent complementary information of different but related images [39, 36]. It appears natural to study relations between shapes or more general image structures via deformations which transform one shape or image onto another [22, 57, 17].

During the past decade, the problem of finding the average of a number of shapes or images has attracted much attention in neuroanatomy research, where anatomical atlases of the brain are constructed from data sets of different subjects [57, 7, 36, 60, 42, 6]. Such an atlas considered as a deformable template can be mapped onto images from an individual patient to detect pathologic abnormalities or to obtain information about the location and extent of certain brain regions [18, 45, 19, 55]. Similar techniques are also employed in object recognition methods such as the active shape model in [22]. Another important application of shape averaging concerns ready-made clothing, where it would be favorable to know the shape of the average human body to design clothes which sufficiently fit as many people as possible. In this paper we will discuss both medical and manufacturing applications in 3D.

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 introduced by Kendall [40] already in 1984. Charpiat et al. [14, 13] discuss shape averaging and shape statistics

[†]Institute for Numerical Simulation, Bonn University, 53113 Bonn, Germany
(`{martin.rumpf,benedikt.wirth}@ins.uni-bonn.de`)

[‡]Benedikt Wirth has been supported by the Bonn International Graduate School for Mathematics (BIGS) and the Hausdorff Center for Mathematics.

based on the notion of the Hausdorff distance of sets. They propose to use smooth approximations of the Hausdorff distance based on a comparison of the signed distance functions of shapes and investigate the correlation of different shapes via gradient descent type morphing from one shape onto the other. In their work, a particular type of splitting the gradient of the shape distance functional and a separate weighting of the different components mimic frame indifference. Based on the singular value decomposition of the covariance matrix in a suitable scalar product, they have additionally analyzed the dominant modes in the variation of the averaged shape.

Fletcher et al. [29] 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. Among other applications they also considered the averaging of 3D kidneys, which are given in a medial axis representation.

An abstract distance measure between objects in (different) metric spaces is the Gromov–Hausdorff distance, which allows to compute an isometrically invariant distance measure between objects $\mathcal{S}$ and $\tilde{\mathcal{S}}$. One definition for the distance comes along with a pair of maps $\phi : \mathcal{S} \rightarrow \tilde{\mathcal{S}}$ and $\psi : \tilde{\mathcal{S}} \rightarrow \mathcal{S}$ which minimize the supremum $\frac{1}{2} \sup_{y_i=\phi(x_i), \psi(y_i)=x_i} |d(x_1, x_2) - d(y_1, y_2)|$ over all maps matching point pairs (x_1, x_2) in $\mathcal{S}$ with pairs (y_1, y_2) in $\tilde{\mathcal{S}}$, respectively. Hence, the Gromov–Hausdorff distance allows to measure—globally and based on an L^∞ type functional—the lack of isometry between two different shapes. Mémoli and Sapiro [43] introduced this concept into the shape analysis community [43] and have proposed this distance measure for shapes described by point clouds, and they discussed efficient numerical algorithms to compute shape distances based on a robust notion of intrinsic distances $d(\cdot, \cdot)$ on the shapes. Bronstein et al. incorporate the Gromov–Hausdorff distance concept in various classification and modeling approaches in geometry processing [9]. Mémoli [44] investigated L^p type variants of the Gromov–Hausdorff distance. In contrast to these global measures of the defect from an isometry, the nonlinear elastic energy functional involved in our approach measures this defect locally, and locally isometric deformations indeed minimize the corresponding local functional.

Understanding shape space as an infinite-dimensional Riemannian manifold, Miller et al. [46, 47] defined the average $\mathcal{S}$ of samples $\mathcal{S}_i$, $i = 1, \dots, n$, as the minimizer of

$$\mathcal{E}[\mathcal{S}] = \sum_{i=1}^n d(\mathcal{S}_i, \mathcal{S})^2$$

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 [31] and further analyzed by Karcher [38]. To ensure that the maps between different shapes, which are generated via integration of the motion fields along geodesics, are diffeomorphisms, suitable regularized metrics have been investigated, either based on regularizing elliptic operators on the surrounding space [61] or on the geometry itself [70]. Fuchs et al. [32] proposed a viscoelastic notion of the distance between shapes $\mathcal{S}$ given as boundaries of physical objects $\mathcal{O}$. They define a metric on infinitesimal variations of a shape $\mathcal{S}$ via the dissipation induced by the corresponding normal displacement on the object boundary. The dissipation functional $\int_{\mathcal{O}(t)} \lambda \text{tr}(\mathcal{D}u + \mathcal{D}u^T)^2 + \mu |\mathcal{D}u + \mathcal{D}u^T|^2 dx$ measures the infinitesimal change of area weighted by λ and length weighted by μ on objects $\mathcal{O}(t)$ on some path in shape space as in linearized elasticity. The method is applied to compute geodesics between two-dimensional shapes using finite elements on a given triangulation of one of the two objects bounded by the corresponding shape contour. Physically, they suppose that due to the viscous dissipation, the object material relaxes instantaneously and remains in a stress-free state. This is a major difference to our approach, where the objects in the deformed configurations bear elastic stresses and the averaged shape $\mathcal{S}$ is characterized by a true balance of these stresses on the shape contour. Furthermore, frame indifference only holds in the limit for decreasing time step size, whereas in our model invariance with respect to rigid body motions

is incorporated in the energy functionals.

The elasticity paradigm for shape analysis, on which our approach is founded, differs significantly from the metric approach in shape space. Given two shapes $\mathcal{S}_1$ and $\mathcal{S}_2$ and an elastic deformation ϕ from $\mathcal{S}_1$ to $\mathcal{S}_2$, we have to distinguish the usually stress-free reference configuration $\mathcal{S}_1$ from the deformed configuration $\mathcal{S}_2 = \phi(\mathcal{S}_1)$, which is under stress and not in equilibrium. Due to the axiom of elasticity, the energy at the deformed configuration $\mathcal{S}_2$ is independent of the path which connects $\mathcal{S}_1$ with $\mathcal{S}_2$. Hence, there is no notion of shortest paths if we consider a purely elastic shape model. As outlined above, if the shapes are assumed to be made of viscous or visco-plastic materials, the underlying dissipation allows to measure the length of connecting paths as long as the final configuration is again stress-free.

Fletcher et al. [30] propose to use a shape median instead of the geometric shape mean. The median is defined as the minimizer of the functional $\mathcal{E}[\mathcal{S}] = \sum_{i=1}^n d(\mathcal{S}_i, \mathcal{S})$. Furthermore, they apply a step size controlled gradient descent for the numerical implementation of the median and demonstrate its robustness in the case of averaging planar curves.

Shape averaging is closely linked to joint registration. In fact, the ensemble of given shapes is jointly registered to the a priori unknown average shape via the underlying deformations. In our averaging method we apply a variational approach for registration of edge based image morphology. Image registration is always based on a particular notion of image similarity. There are landmark-based similarity measures [22, 17], basic measures comparing directly image intensities [39, 16, 34], or measures for the joint entropy or mutual information and related concepts [7, 60, 67, 68, 56]. We refer to Faugeras and Hermosillo [28] for an analytical discussion and comparison of different approaches in the context of mutual information based registration. The morphological similarity based on a comparison of the ensemble of level sets of images was proposed in [26, 27]. To prevent arbitrarily irregular registration deformations, the number of degrees of freedom can be restricted by only allowing rigid [67, 69] or affine deformations [68], B-splines [7, 59] or clamped-plate splines [42]. If an infinite-dimensional space of deformations is considered, regularity has to be ensured via the variational set-up. Hence, either a sufficiently regular metric has to be used for the definition of path lengths in shape space [46, 47], or structural conditions have to be imposed on the functional which measures the energy of a deformation. In fact, this is related to the general regularity theory of linearized elasticity [39, 45, 19], (nonlinear) hyperelasticity [55, 26, 27], or to viscous fluid regularization [17, 18, 16, 8]. For an overview on the latter field we refer to the chapters by Thompson in the handbook [64, 63].

Furthermore, since shapes are often encoded in images or volume data, the averaging procedure is inherently associated with the detection or segmentation of structures. An enormous body of literature on segmentation refers to the seminal paper by Mumford and Shah [51] on a variational approach which states the segmentation problem as a free discontinuity problem for the edge set of an image. The theoretic examination of the Mumford–Shah minimization problem with existence results is quite elaborate [49, 25, 24], and a number of numerically tractable model approximations have been formulated. A very successful approach due to Chan and Vese describes the segments by level sets [11, 12]. Another widely used approach due to Ambrosio and Tortorelli [1] encodes the segment boundaries as a diffused edge set represented by a so-called phase field function. For the actual implementation of our approach we pick up this phase field model. In many applications, the underlying shape description is blurry or corrupted by noise. Here, a joint approach, which couples the segmentation of individual shapes with the overall averaging, allows to restore incomplete shape description via a pullback of the averaged shape onto the image describing the particular given shape. Such joint approaches which combine different image processing tasks in a single variational problem have proven very powerful. For instance, Yezzi, Zöllei and Kapur [37] and Unal et al. [65] have combined segmentation and registration by applying geodesic active contours described by level sets in the given images. Vemuri et al. [66] have also used a level set technique for a reference segmentation in an atlas.

The paper is organized as follows. Section 2 will suggest a variational definition of the shape average, which is based on a hyperelastic energy of the deformed input shapes. We will discuss the physical

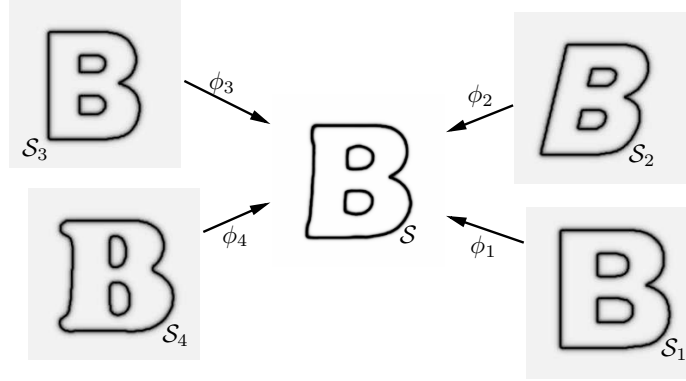


FIG. 2.1. *Sketch of the elastic shape averaging. Input shapes $\mathcal{S}_i$ ($i = 1, \dots, 4$) extracted from input images u_i are mapped onto a shape $\mathcal{S}$ via elastic deformations ϕ_i . The shape $\mathcal{S}$ which minimizes primarily the elastic deformation energy plus some shape regularization to be discussed later is called the shape average of the given input shapes. (Displayed are actual numerical results obtained by the algorithm to be developed in this paper. The resolution of the underlying grid is 257×257 , and the values for the involved parameters are $\gamma = 10^7$, $\mu = 1$, $(a_1, a_2, a_3) = (10^6, 0, 10^6)$.)*

intuition of the model and finally introduce a relaxed formulation for it. Then, we will combine the averaging model with the process of image segmentation to obtain a joint segmentation and averaging model in § 3. The model is stated in terms of phase fields in § 4, proposing adequate approximations to the functionals defined in the previous sections. Sections 5 and 6 are devoted to the resulting Euler Lagrange equations and existence results for minimizers of the resulting phase field functionals. Furthermore, in § 7 the numerical implementation is described based on a suitable finite element discretization and an efficient multi-scale optimization. Finally, some numerical examples of 2D and 3D shape averaging as well as image morphology averaging are presented in § 8, before we conclude in § 9.

2. A nonlinear elastic spring model. In this paper we consider shapes encoded in images. In the simplest case, such an image $u : \Omega \rightarrow \mathbb{R}$ is a characteristic function $u = \chi_{\mathcal{O}}$ representing an object $\mathcal{O}$ as an open set on some domain $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$, and we define the object shape $\mathcal{S} := \partial\mathcal{O}$. More generally we are interested in a shape $\mathcal{S}$ defined as the morphology of an image and represented via the image edge set.

Let us assume that n images $u_i : \Omega \rightarrow \mathbb{R}$ for $i = 1, \dots, n$ are given with a sufficiently regular $(d - 1)$ -dimensional edge sets $\mathcal{S}_i$. We are interested in an average shape which reflects the geometric characteristics of the n given shapes in a physically intuitive manner. Suppose $\mathcal{S} \subset \mathbb{R}^d$ denotes a candidate for this unknown set. Now, we take into account elastic deformations $\phi_i : \Omega \rightarrow \mathbb{R}^d$ with $\phi_i(\mathcal{S}_i) = \mathcal{S}$. Assigned to each of these deformations is an elastic energy $\mathcal{W}[\mathcal{O}_i, \phi_i]$, and we ask for a shape $\mathcal{S}$ such that the total energy given as the sum over all the energies $\mathcal{W}[\mathcal{O}_i, \phi_i]$ for $i = 1, \dots, n$ is minimal. As described below we will have to consider in addition a further energy contribution which acts as a prior on the shape $\mathcal{S}$ in this variational approach. This shape is then called the shape average of the given n shapes (cf. Figure 2.1).

Obviously, this is a constrained variational problem. We simultaneously have to minimize over n deformations ϕ_i and the unknown shape $\mathcal{S}$ given n constraints $\phi_i(\mathcal{S}_i) = \mathcal{S}$.

The model is related to the physical interpretation of the arithmetic mean of n points $x_1, \dots, x_n$ in $\mathbb{R}^d$. Indeed, the arithmetic mean $x \in \mathbb{R}^d$ minimizes $\sum_{i=1}^n \alpha d(x, x_i)^2$, where $d(x, x_i)$ is the distance between x and x_i , and due to Hooke's law the stored energy $\alpha d(x, x_i)^2$ in the spring connecting x_i and x is proportional to the squared distance.

In what follows, we will derive the variational shape averaging model, motivate the structure of the involved energy terms, and discuss different variants.

Nonlinear elastic deformations. We aim for an elastic energy which is not restricted to small displacements, and it should be consistent with first principles. At first, let us briefly recall some background from elasticity. For details we refer to the comprehensive introductions in the books by Ciarlet [20] and Marsden & Hughes [41]. We assume the deformation ϕ_i to be defined on an open object domain $\mathcal{O}_i$, which is supposed to be a subset of $\bar{\Omega}$. If there is no such object, the whole image domain Ω is considered to be elastic, and we set $\mathcal{O}_i = \bar{\Omega}$. In the actual implementation, we will for practical reasons consider a very soft elastic material outside the object on $\Omega \setminus \mathcal{O}_i$, whose material coefficients are several orders of magnitude smaller than those on the object $\mathcal{O}_i$. This ensures a proper extension of the elastic deformation ϕ_i on the whole of Ω with very small interference with the actual object deformation. Furthermore, we suppose $\partial\mathcal{O}_i \setminus \partial\Omega \subset \mathcal{S}_i$ and $\phi_i = \mathbb{1}$ to represent the stress-free reference deformation of the object $\mathcal{O}_i$. It is well-known that the norm of the Jacobian of the deformation $|\mathcal{D}\phi_i|_2$ controls the isotropically averaged change of length under the deformation, where $|A|_2 := (A : A)^{\frac{1}{2}}$ for $A \in \mathbb{R}^{d,d}$ and $A : B = \text{tr}(A^T B)$ for $A, B \in \mathbb{R}^{d,d}$. Secondly, the local volume transformation under a deformation ϕ_i is represented by $\det \mathcal{D}\phi_i$. If $\det \mathcal{D}\phi_i$ changes sign, local self-penetration may be observed. Furthermore, for $d = 3$, $|\text{cof} \mathcal{D}\phi_i|_2$ is a proper measure for the averaged change of area. Here, $\text{cof} A \in \mathbb{R}^{d,d}$ denotes the cofactor matrix for a matrix $A \in \mathbb{R}^{d,d}$, for $A \in GL(d)$ given by $\text{cof} A = (\det A)A^{-T}$. If the local elastic energy W is assumed to depend only on $\mathcal{D}\phi_i$, i. e. $W : \mathbb{R}_+^{d,d} \rightarrow \mathbb{R}$, $A \mapsto W(A)$, such materials are called hyperelastic. Here, $\mathbb{R}_+^{d,d}$ is the space of all $d \times d$ matrices with positive determinant. If we assume an isotropic elastic material in $\mathcal{O}_i$ and the frame indifference principle holds, then W is effectively a function $\bar{W}$ of only the d invariants of $\mathcal{D}\phi_i$. For $d = 3$ these are $(I_1, I_2, I_3) := (|\mathcal{D}\phi_i|_2^2, |\text{cof}(\mathcal{D}\phi_i)|_2^2, \det(\mathcal{D}\phi_i))$. In what follows, we restrict to the three-dimensional case to simplify the exposition. The first Piola–Kirchhoff stress tensor is then recovered as $\sigma[\phi_i] = W_{,A}(\mathcal{D}\phi_i)$, where $W_{,A}(A) = \frac{\partial W(A)}{\partial A}$. Here, we consider a special class of so-called polyconvex energy functionals [23]

$$\mathcal{W}[\mathcal{O}_i, \phi_i] = \int_{\mathcal{O}_i} \hat{W}(\mathcal{D}\phi_i, \text{cof} \mathcal{D}\phi_i, \det \mathcal{D}\phi_i) dx = \int_{\mathcal{O}_i} \bar{W}(|\mathcal{D}\phi_i|_2^2, |\text{cof} \mathcal{D}\phi_i|_2^2, \det \mathcal{D}\phi_i) dx,$$

where $\hat{W}$ is supposed to be convex. Typical energy densities in this class are of the form

$$\bar{W}(I_1, I_2, I_3) = a_1(I_1 - 3)^{\frac{p}{2}} + a_2(I_2 - 3)^{\frac{q}{2}} + a_3 \left(I_3^{-s} + \frac{s}{r} I_3^r - \frac{r+s}{r} \right) \quad (2.1)$$

with $a_1, a_2, a_3 > 0$. In nonlinear elasticity such material laws have been proposed by Ogden [52], and for $p = q = 2$ we obtain the Mooney–Rivlin model [20]. The built-in penalization of volume shrinkage, i. e. $\bar{W}(I_1, I_2, I_3) \xrightarrow{I_3 \rightarrow 0} \infty$, enables us to control local injectivity (cf. [5]). Furthermore, a deformation which is locally isometric, i. e. $\mathcal{D}\phi_i^T(x) \mathcal{D}\phi_i(x) = \mathbb{1}$, is a local minimizer of the energy. We actually consider $p = q = 2$ and $r = s = 1$.

Incorporation of a nonlinear elastic energy for the deformations ϕ_i onto a prospective shape average is indispensable for our approach:

- It allows to incorporate large deformations with strong material and geometric nonlinearities, which cannot be treated by a linear approach.
- The dependency of the energy density $\bar{W}$ follows from first principles and measures the physical effects of length, area, and volume distortion, which reflect the local distance from an isometry.
- Finally, it balances in an intrinsic way expansion and collapse of the elastic objects and hence frees us to impose artificial conditions, e. g. boundary conditions, on the expected average shape as the image of the objects under the set of deformations.

Hong et al. [35] incorporated the stored energy from linearized elasticity as a distance measure between shapes and used this to compute shape averages and principal modes of shape variation. Their energy functional is invariant with respect to rigid body motions only in an infinitesimal sense, and it measures deformations not only within the objects, but rather on the whole ambient space. In Pennec et al. [54, 53],

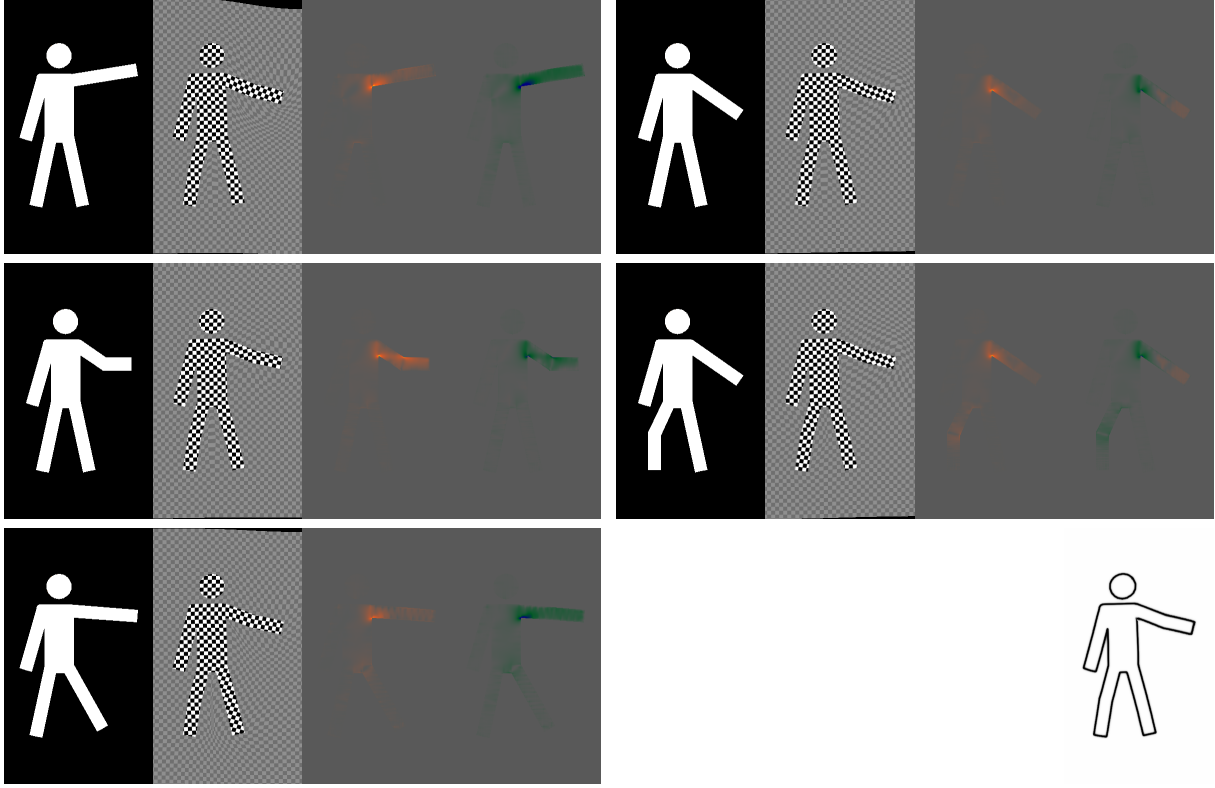


FIG. 2.2. Given five silhouettes of a person as input objects, a shape average (bottom right) is computed based on our elastic averaging approach. The original images are depicted along with their deformations ϕ_i (continued to the region outside the object and displayed acting on a checkerboard on the image domain Ω) and the distribution of local change of length $|\mathcal{D}\phi_i|_2$ and local change of area $\det(\mathcal{D}\phi_i)$ (from left to right). These local densities ranging over $[0.97\sqrt{2}, 1.03\sqrt{2}]$ and $[0.97, 1.03]$, respectively, are color-coded using the color map . The underlying image resolution is 513×513 , and the energy parameters are $\gamma = 10^7$, $\mu = 10^{-2}$, $(a_1, a_2, a_3) = (10^{10}, 0, 10^{10})$.

the regularizing prior in nonlinear registration of images is based on a symmetric distance measure between images, which picks up a related concept from continuum mechanics. In fact, a nonlinear elastic energy is defined as the integral over the ambient space of a St. Venant–Kirchhoff type energy density depending on the logarithm of the Cauchy–Green strain tensor $\mathcal{D}\phi^T \mathcal{D}\phi$. This energy also measures distance from an isometry or an average deformation and acts as a penalty to avoid material interpenetration. However, it is in general not quasi-convex, which renders the existence theory of minimizers difficult. Furthermore, even though the ansatz deals with large deformations, the resulting distance measure is designed for a local comparison of deformations near an averaged deformation field (cf. the discussion on Riemannian distance and elastic energies in § 1).

Variational definition of the shape average. As already outlined above, for a fixed shape $\mathcal{S}$ we restrict the set of admissible deformations for each object $\mathcal{O}_i$ imposing the constraint $\phi_i(\mathcal{S}_i) = \mathcal{S}$ to deduce a suitable energy on shapes $\mathcal{S}$ being candidates for the shape average and sets of deformations $(\phi_i)_{i=1,\dots,n}$ matching given shapes $\mathcal{S}_i$ with $\mathcal{S}$:

$$\hat{\mathcal{E}}[\mathcal{S}, (\phi_i)_{i=1,\dots,n}] = \begin{cases} \frac{1}{n} \sum_{i=1,\dots,n} \mathcal{W}[\mathcal{O}_i, \phi_i] & ; \phi_i(\mathcal{S}_i) = \mathcal{S} \text{ for } i = 1, \dots, n \\ \infty & ; \text{else} . \end{cases}$$

Finally, we define the shape average $\mathcal{S}$ as the corresponding minimizer over a suitable set of admissible

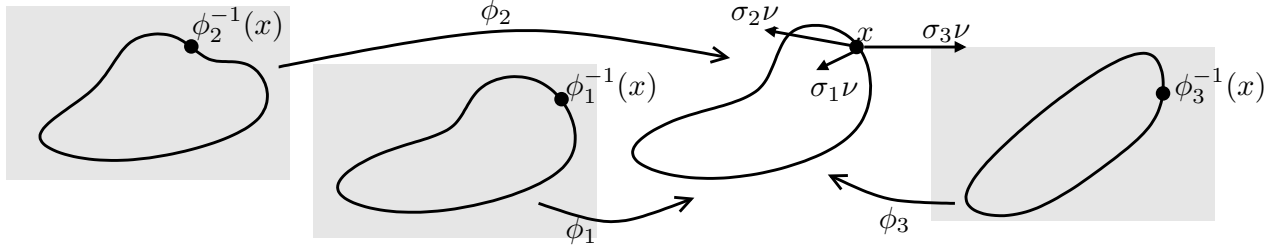


FIG. 2.3. Sketch of the stress balance relation on the averaged shape. $\sigma_i\nu$ is short for $\sigma^{def}[\phi_i]\nu[\mathcal{S}]$.

shapes $\mathcal{A}_{\mathcal{S}}$, i. e.

$$\mathcal{S} = \arg \min_{\tilde{\mathcal{S}} \in \mathcal{A}_{\mathcal{S}}, \phi_i: \mathcal{O}_i \rightarrow \mathbb{R}^d} \mathcal{E}[\tilde{\mathcal{S}}, (\phi_i)_{i=1, \dots, n}] = \arg \min_{\tilde{\mathcal{S}} \in \mathcal{A}_{\mathcal{S}}} \left(\arg \min_{\phi_i: \mathcal{O}_i \rightarrow \mathbb{R}^d} \hat{\mathcal{E}}[\tilde{\mathcal{S}}, (\phi_i)_{i=1, \dots, n}] + \mu \mathcal{L}[\tilde{\mathcal{S}}] \right).$$

Here $\mathcal{L}$ denotes an energy functional on shapes which acts as a regularization to restrict the space of admissible shapes.

In fact, we interpret the elastic energy $\mathcal{W}[\mathcal{O}_i, \phi_i]$ associated with each deformation ϕ_i which maps one of the shapes $\mathcal{S}_i$ onto the shape $\mathcal{S}$ as a nonlinear counterpart of the energy stored in a spring in the above classical interpretation of an averaged position. It measures in a physically rigorous way locally the lack of isometry as already mentioned above (cf. Figure 2.2).

Existence of an average shape strongly depends on the regularity of the input shapes and the prior $\mathcal{L}$. Indeed, if we suppose the initial shapes $\mathcal{S}_i$ and a given shape $\mathcal{S}$ to consist of sufficiently smooth unions of hypersurfaces on bounded domains $\mathcal{O}_i$ with sufficiently smooth boundary $\partial\mathcal{O}_i$ for $i = 1, \dots, n$, then by the classical existence theory in nonlinear elasticity [20] there exists a set of minimizers $(\phi_i)_{i=1, \dots, n}$ which allow us to compute $\arg \min_{\phi_i: \mathcal{O}_i \rightarrow \mathbb{R}^d} \hat{\mathcal{E}}[\mathcal{S}, (\phi_i)_{i=1, \dots, n}]$. These minimizers are weakly differentiable and locally injective. The integrability exponent for the derivatives depends on the exponents p, q, s, r . By Sobolev embedding we know that the minimizing deformations are Hölder continuous if $p > d$. Let us remark that the regularity theory allows to prove Lipschitz continuity only under certain strong conditions [15]. Hence, there is not sufficient control on the images $\phi_i(\mathcal{S}_i)$ as long as we do not restrict $\mathcal{S}$ to some compact set of sufficiently smooth shapes. This compactness can be ensured by considering only a finite dimensional space of shapes $\mathcal{A}_{\mathcal{S}}$, e. g. some spline space. In this case, a classical Arzela–Ascoli argument can be applied to prove that the minimum of the total energy without any prior $\mathcal{L}[\cdot]$ is attained. Alternatively, one can consider a regularizing prior $\mathcal{L}$ as proposed here. This prior is needed to control the regularity of the shapes $\mathcal{S}$ on a minimizing sequence for the total energy. To ensure existence of a minimizing average shape the prior should enforce the shapes $\mathcal{S}$ on a minimizing sequence to consist of at least Lipschitz hypersurfaces. At the same time the prior should be weakly lower semi continuous on the space $\mathcal{A}_{\mathcal{S}}$ of admissible shapes. These requirements render the existence problem an open problem in case of general classes of admissible shapes.

We refer to § 6 for a rigorous proof of an existence result in case of the phase field approximation we have actually implemented. In particular, with regard to an effective implementation of this approach we confine to a prior $\mathcal{L}[\mathcal{S}]$ which measures the $\mathcal{H}^{d-1}$ -measure of $\mathcal{S}$, i. e.

$$\mathcal{L}[\mathcal{S}] = \int_{\mathcal{S}} da.$$

A necessary condition for a set of minimizing deformations are the corresponding Euler Lagrange equations. As usual, inner variations of one of the deformations lead to the classical system of partial differential equations $\operatorname{div} W_{,A}(\mathcal{D}\phi_i) = 0$ for every deformation ϕ_i on $\mathcal{O}_i \setminus \mathcal{S}_i$. Due to the set of constraints

$(\phi_i(\mathcal{S}_i) = \mathcal{S})_{i=1,\dots,n}$, the conditions on ϕ_i are interlinked. For simplicity let us assume that $\mathcal{S}_i = \partial\mathcal{O}_i$ for all $i = 1, \dots, n$. Then we can deduce a balance relation between deformation stresses on the averaged shape $\mathcal{S}$. Namely, for $x \in \mathcal{S}$ we obtain

$$\sum_{i=1,\dots,n} \sigma[\phi_i](\phi_i^{-1}(x)) \nu[\mathcal{S}_i](\phi_i^{-1}(x)) \, da[\mathcal{S}_i] = 0, \quad (2.2)$$

where $\nu[\mathcal{S}_i]$ is the outer normal on $\mathcal{S}_i$, $da[\mathcal{S}_i]$ the surface element on $\mathcal{S}_i$, and $\phi_i^{-1}(x)$ the pre-image of x under the deformation ϕ_i . From elasticity theory we know that the forces $\sigma[\phi_i](\phi_i^{-1}(x)) \nu[\mathcal{S}_i](\phi_i^{-1}(x)) \, da[\mathcal{S}_i]$ in the reference configuration equal the corresponding forces $\sigma^{\text{def}}[\phi_i](x) \nu[\mathcal{S}](x) \, da[\mathcal{S}]$ in the deformed configuration so that (2.2) directly implies a balance of all normal stresses on the average shape $\mathcal{S}$,

$$0 = \sum_{i=1,\dots,n} \sigma^{\text{def}}[\phi_i](x) \nu[\mathcal{S}](x),$$

where $\nu[\mathcal{S}](x)$ is the outer normal on $\mathcal{S}$ and $\sigma^{\text{def}}[\phi_i] = (\sigma[\phi_i](\det \mathcal{D}\phi_i)^{-1} \mathcal{D}\phi_i^T) \circ \phi_i^{-1}$ the usual Cauchy stress tensor corresponding to the deformation ϕ_i in the deformed configuration (cf. Figure 2.3).

This gives a refined physical interpretation of the shape average as the stable shape on which all surface forces implied by the elastic deformations are balanced (cf. Figure 2.4). Obviously, there is a straightforward generalization involving jumps of normal stresses on interior interfaces in case of components of $\mathcal{S}_i$ which are interior edges in $\mathcal{O}_i$.

To verify (2.2) let us consider a set of consistent variations $(\mathbb{1} + \epsilon u) \circ \phi_i$, which lead to a variation $(\mathbb{1} + \epsilon u)(\mathcal{S})$ of the average shape. Due to the optimality, we obtain

$$\frac{d}{d\epsilon} \sum_{i=1,\dots,n} \mathcal{W}[\mathcal{O}_i, (\mathbb{1} + \epsilon u) \circ \phi_i] \Big|_{\epsilon=0} = 0.$$

Thus, differentiation and an integration by parts leads to

$$\begin{aligned} 0 &= \sum_{i=1,\dots,n} \int_{\mathcal{O}_i} W_{,A}(\mathcal{D}\phi_i) : \mathcal{D}(u \circ \phi_i) \, dx \\ &= 0 + \sum_{i=1,\dots,n} \int_{\mathcal{S}_i} W_{,A}(\mathcal{D}\phi_i) : (u \circ \phi_i) \otimes \nu[\mathcal{S}_i] \, da[\mathcal{S}_i], \end{aligned}$$

where $\nu[\mathcal{S}_i]$ is the outer normal in $\mathcal{S}_i$ and “ $\otimes$ ” denotes the rank-1 product $v \otimes w = v w^T$ for two vectors $v, w \in \mathbb{R}^d$. Here, we have made use of the above PDE systems of nonlinear elasticity on $\mathcal{O}_i \setminus \mathcal{S}_i$. Now, we consider displacements u with local support and let this support collapse at some point x on $\mathcal{S}$.

On some similar alternative models. In our definition of an elastic shape average we have considered elastic deformations ϕ_i which map input shapes $\mathcal{S}_i$ on the shape average candidate $\mathcal{S}$. Alternatively, one might want to consider deformations from the shape average onto the input shapes. Then we would have to take into account the corresponding inverse maps $\psi_i = \phi_i^{-1}$ and their elastic energy. We readily verify that $\det \mathcal{D}\psi_i = \frac{1}{\det \mathcal{D}\phi_i}$, $\mathcal{D}\psi_i = \frac{\text{cof} \mathcal{D}\phi_i^T}{\det \mathcal{D}\phi_i}$, and $\text{cof} \mathcal{D}\psi_i = \frac{\mathcal{D}\phi_i^T}{\det \mathcal{D}\phi_i}$. Hence, the elastic energies associated with the inverse deformations ψ_i on the stress-free reference configuration $\mathcal{S}$ can be phrased as a functional on the input domains $\mathcal{S}_i$, and we obtain

$$\mathcal{W}^{\text{inv}}[\psi_i] = \mathcal{W}^{\text{inv}}[\mathcal{O}_i, \phi_i] := \int_{\mathcal{O}_i} |\det \mathcal{D}\phi_i| \bar{W} \left(\frac{|\text{cof} \mathcal{D}\phi_i|_2^2}{(\det \mathcal{D}\phi_i)^2}, \frac{|\mathcal{D}\phi_i|_2^2}{(\det \mathcal{D}\phi_i)^2}, \frac{1}{\det \mathcal{D}\phi_i} \right) \, dx.$$

Alternatively, one might consider a symmetric version in terms of the sum of the pair of energies $\mathcal{W}^{\text{inv}}[\mathcal{O}_i, \phi_i] + \mathcal{W}[\mathcal{O}_i, \phi_i]$. In this paper, we restrict ourselves to the initial model derived above.

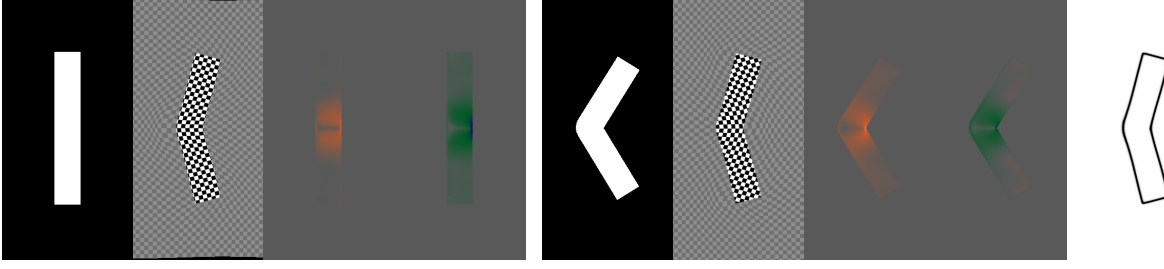


FIG. 2.4. A straight and a folded bar as a test case. As above, the input images are depicted along with their deformations ϕ_i and the distribution of $|\mathcal{D}\phi_i|_2$ and $\det(\mathcal{D}\phi_i)$ with ranges of $[0.97\sqrt{2}, 1.03\sqrt{2}]$ and $[0.97, 1.03]$ color-coded as [blue, red] . Apparently, isometries are preserved distant from the folding point, and the region of higher deformation energies restricts to the area around the fold. The original bars describe an angle of 180° and 118° , while the average approximately has an angle of 150° . The image resolution is 513×513 , and parameters are as in Figure 2.2.

Our approach can be considered as a PDE-constrained shape optimization method [58], where we optimize the shape $\mathcal{S}$ with respect to the cost functional $\sum_{i=1}^n \mathcal{W}[\mathcal{O}_i, \phi_i^*]$ under the constraint that ϕ_i^* is a minimizer of the functional $\phi_i \mapsto \mathcal{W}[\mathcal{O}_i, \phi_i]$ with $\phi_i(\mathcal{S}_i) = \mathcal{S}$. A phase field model in shape optimization related to the one to be developed in this paper is discussed in [10]. For some applications it might be useful to decouple the cost functional on $\mathcal{S}$ from the constraint on ϕ^* , e. g. the actual cost functional could depend on stresses on $\mathcal{S}$.

A relaxed formulation. The hard constraint $\phi_i(\mathcal{S}_i) = \mathcal{S}$ is often inadequate in applications. Due to local shape fluctuations, for example, or noise in the shape acquisition there are frequently local details like spurious edges in the input shapes which should not be encountered in the shape average. Hence, we relax the constraint and introduce a penalty functional which measures the symmetric difference of the input shapes $\mathcal{S}_i$ and the pullback $\phi_i^{-1}(\mathcal{S})$ of the shape $\mathcal{S}$ and is given by

$$\mathcal{F}[\mathcal{S}_i, \phi_i, \mathcal{S}] = \mathcal{H}^{d-1}(\mathcal{S}_i \Delta \phi_i^{-1}(\mathcal{S})),$$

where $A \Delta B = A \setminus B \cup B \setminus A$. Finally, to sum up, our shape averaging model is based on the energy

$$\mathcal{E}^\gamma[\mathcal{S}, (\phi_i)_{i=1, \dots, n}] = \frac{1}{n} \sum_{i=1}^n (\mathcal{W}[\mathcal{O}_i, \phi_i] + \gamma \mathcal{F}[\mathcal{S}_i, \phi_i, \mathcal{S}]) + \mu \mathcal{L}[\mathcal{S}]. \quad (2.3)$$

This approach is related to groupwise registration and segmentation results [68, 69]. Here, $\mathcal{F}$ acts as a fidelity term measuring the quality of the registration of the shapes $\mathcal{S}_i$ with a given shape $\mathcal{S}$ under the deformations ϕ_i . In the next section we will extend the approach to joint edge segmentation and averaging.

3. A joint segmentation and averaging method. In the derivation of our shape averaging model we have assumed that the shapes $\mathcal{S}_i$ to be averaged can be robustly extracted from a set of images u_i with $i = 1, \dots, n$ and are a priori given. However, if we consider shapes defined as the morphology of images represented by edges, some of these edges will be characterized by significant noise or low contrast and hence will be difficult to extract. Here, it might help to take into account the corresponding edges in the other images, which all refer to the same edge of the average shape, in particular, we will pursue a joint approach of shape segmentation and registration with an averaged shape: We do no longer assume the shapes $\mathcal{S}_i$ to be given a priori, but we seek for them simultaneously to the averaging process, using one single functional for both tasks. This has several advantages: On the one hand, the quality of shape averaging highly depends on the robustness of the edge detection in the input images. On the other hand, a reliable average shape can be used to improve edge detection in case of poor image quality. For instance, Young and Levy [71] used segmentation results from one image to guide edge detection in consecutive

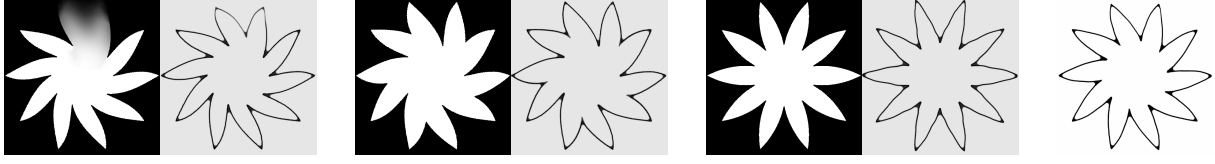


FIG. 3.1. *Blurred edges can be restored based on a joint approach for image segmentation and averaging. The three input images u_i^0 are depicted along with their segmented edge sets (described as phase field v_i) as computed by the joint segmentation and averaging. The computed average shape is also shown (right). Apparently, the strongly blurred edges in the first input image are reconstructed based on the corresponding edges in the other images. (Resolution 513×513 , $\gamma = 10^7$, $\mu = 10^{-2}$, $(a_1, a_2, a_3) = (10^8, 0, 10^8)$, $\alpha = 10^{10}$, $\beta = 10^5$, $\nu = 10^6$.)*

images, and Yezzi, Zöllei and Kapur [37] simultaneously segmented contours in multiple images using an active contour type model. In [27] a joint segmentation and morphology registration method has been proposed.

Joint shape segmentation and averaging can be phrased in terms of a variational approach, where shape segmentation, shape matching with a shape average, and optimization of this shape average are performed simultaneously via the minimization of a joint energy functional. For shape segmentation we pick up the approach presented by Mumford and Shah [51] and consider the functional

$$\mathcal{E}_{\text{MS}}[u, \mathcal{S}, u^0] = \alpha \int_{\Omega} (u - u^0)^2 dx + \beta \int_{\Omega \setminus \mathcal{S}} |\nabla u|^2 dx + \nu \mathcal{H}^{d-1}(\mathcal{S}), \quad (3.1)$$

where $u^0 : \Omega \rightarrow \mathbb{R}$ represents the original image and u a piecewise smooth approximation with edge set $\mathcal{S}$. This functional balances the approximation of a possibly noisy or corrupted input image u^0 , the smoothness of the piecewise smooth approximation u , and the length ($d = 2$) or area ($d = 3$) of the edge interfaces $\mathcal{S}$. Incorporating this approach in the above shape averaging functional (2.3) we end up with the joint functional

$$\mathcal{E}_{\text{joint}}^{\gamma}[\mathcal{S}, (u_i, \mathcal{S}_i, \phi_i)_{i=1, \dots, n}] = \frac{1}{n} \sum_{i=1}^n (\mathcal{E}_{\text{MS}}[u_i, \mathcal{S}_i, u_i^0] + \mathcal{W}[\Omega, \phi_i] + \gamma \mathcal{F}[\mathcal{S}_i, \phi_i, \mathcal{S}] + \mu \mathcal{L}[\mathcal{S}]), \quad (3.2)$$

which has to be relaxed simultaneously in the unknowns u_i , $\mathcal{S}_i$, ϕ_i for $i = 1, \dots, n$ and $\mathcal{S}$ for a fixed given set of input images $(u_i^0)_{i=1, \dots, n}$. Figure 3.1 demonstrates that in a joint approach blurry edges in the input images can be segmented, if sufficiently strong evidence for this edge from other input images is integrated into the averaged shape.

4. Phase field approximation. Since explicit treatment of an edge set is difficult in a variational setting [50], many approximations to the original Mumford–Shah model (3.1) have been proposed. The approach by Ambrosio and Tortorelli [1, 2] consists in encoding an edge set $\mathcal{S}$ by a smooth phase field function $v : \Omega \rightarrow \mathbb{R}$, which is close to zero on edges and one in between. The specific shape of this phase field function, which represents the edge set, is determined by suitable terms in the following functional,

$$\mathcal{E}_{\text{AT}}^{\epsilon}[u, v, u^0] = \alpha \int_{\Omega} (u - u^0)^2 dx + \beta \int_{\Omega} (v^2 + k_{\epsilon}) |\nabla u|^2 dx + \nu \int_{\Omega} \epsilon |\nabla v|^2 + \frac{1}{4\epsilon} (v - 1)^2 dx, \quad (4.1)$$

where ϵ can be interpreted as the width of the diffused edge representation in v . $k_{\epsilon} > 0$ is a small parameter, needed for analytical purposes, that converges to zero as $\epsilon \rightarrow 0$ (in our computations, however, we always set $k_{\epsilon} = 0$). As $\epsilon \rightarrow 0$, the last integral in the above functional indeed Γ -converges to $\mathcal{H}^{d-1}(\mathcal{S})$. We pick up this approximation and describe the average shape $\mathcal{S}$ via a phase field function v . For the input shapes $\mathcal{S}_i$ we assume the corresponding phase field description v_i to be given a priori. (Usually, v_i can be computed beforehand, minimizing (4.1) on the input images u_i , or explicitly constructed for a

given edge set using the comparison function from the slicing argument in [2].)

This approximation is favorable due to its quadratic form, which will allow direct computation of the smoothed image and the phase field in an alternating minimization approach. It is numerically easier to handle than for instance the strongly nonlinear level set approach by Vese and Chan in [11] or an approximation of Modica–Mortola type [48]. Also, as opposed to level sets, the Ambrosio–Tortorelli phase field can describe edges in images which are not boundary contours and therefore allows a simple handling of very complex shape topologies. However, this is at the expense of not possessing an inherent distinction between the inside and the outside of a shape, which level sets do possess. Also, level sets have the advantage of a sharp instead of a diffuse interface description, which on the other hand prevents them to be used with different scales. Finally, there is a well-established existence and Γ -convergence theory for the phase field approach.

Given a phase field parameter ϵ and corresponding phase field representations v of $\mathcal{S}$ and v_i of $\mathcal{S}_i$, we define an approximate mismatch penalty

$$\mathcal{F}^\epsilon[v_i, \phi_i, v] = \frac{1}{\epsilon} \int_{\Omega} (v \circ \phi_i)^2 (1 - v_i)^2 + v_i^2 (1 - v \circ \phi_i)^2 \, dx.$$

Here, we suppose v to be extended by 1 outside the computational domain Ω . The first term in the integrand is close to 1 on $\mathcal{S}_i \setminus \phi_i^{-1}(\mathcal{S})$, because $(1 - v_i) \approx 1$ on $\mathcal{S}_i$ and $v \circ \phi_i \approx 1$ apart from the vicinity of $\phi_i^{-1}(\mathcal{S})$. It tends to 0 with increasing distance from this set. Analogously, the second term acts as an approximate indicator function for $\phi_i^{-1}(\mathcal{S}) \setminus \mathcal{S}_i$. Let us emphasize that $\mathcal{F}^\epsilon[v_i, \phi_i, v]$ is expected to be truly proportional to $\mathcal{F}[\mathcal{S}_i, \phi_i, \mathcal{S}]$ only, if ϕ_i is neither distending nor compressive orthogonally to the shape, i. e. $\mathcal{D}\phi_i \nu[\phi_i^{-1}(\mathcal{S})] \cdot \nu[\phi_i^{-1}(\mathcal{S})] = 1$ on $\phi_i^{-1}(\mathcal{S})$. Nevertheless, because we are primarily interested in the limit for $\gamma \rightarrow \infty$, $\mathcal{F}^\epsilon[v_i, \phi_i, v]$ acts as a proper penalty functional.

Next, we have to describe the phase field v , which is not given a priori, in an implicit variational form. We pick up the idea by Ambrosio and Tortorelli and consider the energy

$$\mathcal{L}^\epsilon[v] = \int_{\Omega} \epsilon |\nabla v|^2 + \frac{1}{4\epsilon} (v - 1)^2 \, dx,$$

which additionally acts as a regularization energy measuring an approximation of the $\mathcal{H}^{d-1}$ measure of the shape $\mathcal{S}$ represented by the phase field v .

So far the elastic energy is evaluated on the object domains $\mathcal{O}_i$ only. For practical reasons of the later numerical discretization, we now let the whole computational domain behave elastically with an elasticity several orders of magnitude softer outside the object domains $\mathcal{O}_i$ on the complement set $\Omega \setminus \mathcal{O}_i$. We suppose that, based on a priori segmentation of the images u_i , a smooth approximation $\chi_{\mathcal{O}_i}^\epsilon$ of the characteristic function $\chi_{\mathcal{O}_i}$ is given and define a corresponding approximate elastic energy

$$\mathcal{W}^\epsilon[\mathcal{O}_i, \phi_i] = \int_{\Omega} \left((1 - \delta) \chi_{\mathcal{O}_i}^\epsilon + \delta \right) W(\mathcal{D}\phi_i) \, dx,$$

where in our application $\delta = 10^{-4}$. Also, in the above we implicitly assumed that deformations ϕ_i map the domain Ω onto itself; for numerical implementation we will relax this assumption and perform integrations only in regions where all integrands are defined. Finally, the resulting approximation of the total energy functional to be minimized reads

$$\mathcal{E}^{\gamma, \epsilon}[v, (\phi_i)_{i=1, \dots, n}] = \frac{1}{n} \sum_{i=1}^n (\mathcal{W}^\epsilon[\mathcal{O}_i, \phi_i] + \gamma \mathcal{F}^\epsilon[v_i, \phi_i, v]) + \mu \mathcal{L}^\epsilon[v]. \quad (4.2)$$

Let us remark that we are particularly interested in the case where $\mathcal{F}^\epsilon$ acts as a penalty with $\gamma \gg 1$ and $\mathcal{L}^\epsilon$ ensures a mild regularization of the averaged shape with $\mu \ll 1$.

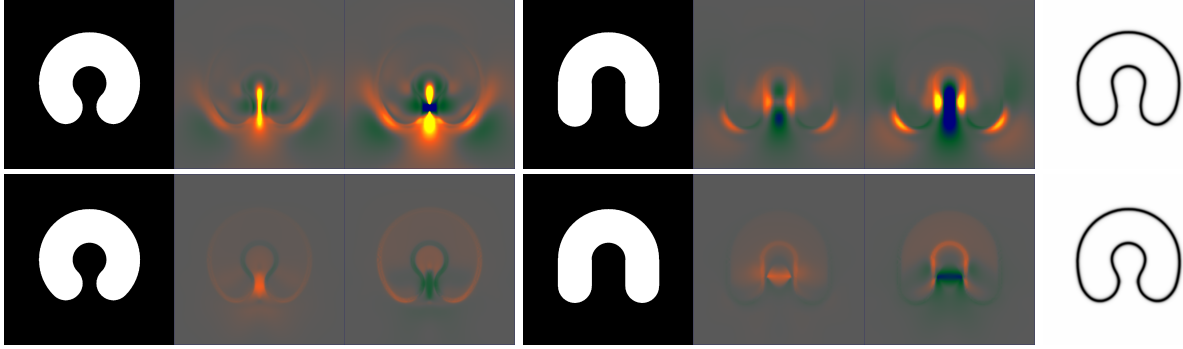


FIG. 4.1. *Input images together with $|\mathcal{D}\phi_i|_2$ and $\det(\mathcal{D}\phi_i)$ (ranges of $[0.6\sqrt{2}, 1.4\sqrt{2}]$ and $[0.6, 1.4]$ color-coded as ) and the average phase field (rightmost). In the top row only the interior of the two shapes is considerably stiff, whereas in the bottom row the whole computational domain is considered to be homogeneously elastic. Obviously, in the upper case far stronger strains are visible in the region of the gap, and in the lower case it is much more expensive to pull the lobes apart in the first shape than to push them together in the second shape. Hence, the resulting average in the second row is characterized by stronger bending of the two lobes than in the first row. (Results are obtained for a grid resolution 1025×1025 and parameter values $\gamma = 10^7$, $\mu = 0.1$, $\epsilon = 6h$, $(a_1, a_2, a_3) = (10^6, 0, 10^6)$.)*

The structure of the penalty functional $\mathcal{F}^\epsilon$ tries to match the shapes of the given phase field functions v_i and the pullback $v \circ \phi_i$ of the phase field v to be determined. This implies a particular stiffness of the deformations ϕ_i on the diffused interface around the shapes $\mathcal{S}_i$. Indeed, the set of deformations $\phi_1 \dots, \phi_n$ tries to minimize stretch or compression normal to the shape contour measured in terms of $\mathcal{D}\phi_i \nu[\mathcal{S}_i] \cdot \nu[\mathcal{S}_i] - 1$ (cf. Figure 8.5). This does not hamper the elastic deformation in the limit for $\gamma \rightarrow \infty$, because the other (tangential) components of the deformation tensor can relax freely.

Figure 4.1 shows the impact of the choice of the elastic domain on the average shape. Here, we once consider the whole computational domain as homogeneously elastic, and alternatively and in many cases physically more sound only the object domain is assumed to be elastic and considerably stiff. The region between both lobes is more severely dilated if the elastic energy is weighted with a small factor outside the shape, which becomes obvious especially in the plots of the deformation invariants.

Finally, let us give the phase field approximation for the joint model (3.2), where we simultaneously segment and average shapes. Here, for fixed input images $(u_i^0)_{i=1, \dots, n}$, we end up with the functional

$$\mathcal{E}_{\text{joint}}^{\gamma, \epsilon}[v, (u_i, v_i, \phi_i)_{i=1, \dots, n}] = \frac{1}{n} \sum_{i=1}^n (\mathcal{E}_{\text{AT}}^\epsilon[u_i, v_i, u_i^0] + \mathcal{W}^\epsilon[\Omega, \phi_i] + \gamma \mathcal{F}^\epsilon[v_i, \phi_i, v]) + \mu \mathcal{L}^\epsilon[v]. \quad (4.3)$$

5. Euler Lagrange equations. In what follows, we will consider the Euler Lagrange equations of the above phase field energies. We thus need to compute the variations of the energy contributions with respect to the phase field v and the involved deformations $\phi_1, \dots, \phi_n$. The variation of an energy $\mathcal{G}$ in direction ζ with respect to a function z will be denoted by $\langle \delta_z \mathcal{G}, \zeta \rangle$. Using straightforward differentiation,

for sufficiently smooth v and ϕ_i we obtain

$$\begin{aligned}\langle \delta_v \mathcal{F}^\epsilon[v_i, \phi_i, v], \vartheta \rangle &= \frac{2}{\epsilon} \int_{\phi_i(\Omega)} (v(1 - v_i \circ \phi_i^{-1})^2 - (v_i \circ \phi_i^{-1})^2(1 - v)) \vartheta |\det(\mathcal{D}\phi_i^{-1})| dx, \\ \langle \delta_{\phi_i} \mathcal{F}^\epsilon[v_i, \phi_i, v], \psi \rangle &= \frac{2}{\epsilon} \int_{\Omega} ((1 - v_i)^2(v \circ \phi_i) - v_i^2(1 - v \circ \phi_i)) (\nabla v \circ \phi_i) \cdot \psi dx, \\ \langle \delta_v \mathcal{L}^\epsilon[v], \vartheta \rangle &= 2 \int_{\Omega} \epsilon \nabla v \cdot \nabla \vartheta + \frac{1}{4\epsilon} (v - 1) \vartheta dx, \\ \langle \delta_{\phi_i} \mathcal{W}^\epsilon[\mathcal{O}_i, \phi_i], \psi \rangle &= \int_{\Omega} \left((1 - \delta) \chi_{\mathcal{O}_i}^\epsilon + \delta \right) W_{,A}(\mathcal{D}\phi_i) : \mathcal{D}\psi dx\end{aligned}$$

for scalar test functions ϑ and vector valued displacement type test functions ψ . Concerning the particular polyconvex energy integrand given in (2.1), we obtain

$$\begin{aligned}W_{,A}(A) : B &= 2 \partial_{I_1} \bar{W}(|A|_2^2, |\text{cof} A|_2^2, \det A) A : B + \\ &\quad 2 \partial_{I_2} \bar{W}(|A|_2^2, |\text{cof} A|_2^2, \det A) \text{cof} A : \partial_A \text{cof}(A)(B) + \\ &\quad \partial_{I_3} \bar{W}(|A|_2^2, |\text{cof} A|_2^2, \det A) \partial_A \det(A)(B),\end{aligned}$$

where

$$\begin{aligned}\partial_A \det(A)(B) &= \det(A) \text{tr}(A^{-1} B), \\ \partial_A \text{cof}(A)(B) &= \det(A) \text{tr}(A^{-1} B) A^{-T} - \det(A) A^{-T} B^T A^{-T}, \\ \partial_{I_1} \bar{W}(I_1, I_2, I_3) &= \frac{p}{2} a_1 (I_1 - 3)^{\frac{p-2}{2}}, \\ \partial_{I_2} \bar{W}(I_1, I_2, I_3) &= \frac{q}{2} a_2 (I_2 - 3)^{\frac{q-2}{2}}, \\ \partial_{I_3} \bar{W}(I_1, I_2, I_3) &= s a_3 (I_3^{r-1} - I_3^{-s-1}).\end{aligned}$$

For ease of implementation we will only consider the case $p, q = 4$ and $r, s = 2$ below. A local minimum $v, \phi_1, \dots, \phi_n$ of (4.2) is characterized by the Euler Lagrange conditions

$$\langle \delta_v \mathcal{E}^{\gamma, \epsilon}[v, (\phi_i)_{i=1, \dots, n}], \vartheta \rangle = 0, \quad \langle \delta_{\phi_i} \mathcal{E}^{\gamma, \epsilon}[v, (\phi_j)_{j=1, \dots, n}], \psi \rangle = 0, \quad (5.1)$$

where $\delta_v \mathcal{E}^{\gamma, \epsilon} = \frac{\gamma}{n} \sum_{i=1}^n \delta_v \mathcal{F}^\epsilon + \mu \delta_v \mathcal{L}^\epsilon$ and $\delta_{\phi_i} \mathcal{E}^{\gamma, \epsilon} = \frac{1}{n} \sum_{i=1}^n (\delta_{\phi_i} \mathcal{W}^\epsilon + \gamma \delta_{\phi_i} \mathcal{F}^\epsilon)$.

Upon integration by parts and use of the fundamental lemma of the calculus of variations, these conditions readily imply a coupled system of nonlinear second order partial differential equations for $v, \phi_1, \dots, \phi_n$ and corresponding boundary conditions. The resulting natural boundary condition for v is $\nabla v \cdot \nu[\partial\Omega] = 0$ on $\partial\Omega$ for the outward normal $\nu[\partial\Omega]$. Effectively, in the limit for $\epsilon \rightarrow 0$ this implies that if the average shape $\mathcal{S}$ meets the the boundary $\partial\Omega$ of the computational domain Ω , then it is perpendicular to this boundary, since $\frac{\nabla v}{|\nabla v|}$ approximates the normal $\nu[\mathcal{S}]$ on $\mathcal{S}$. The natural boundary condition for the deformation ϕ_i is $\sigma[\phi_i] \nu[\partial\Omega] = 0$ on $\partial\Omega$, which reflects the boundary condition at the outer boundary of the soft material phase. The actually physically relevant boundary condition is diffused in the transition layer around $\partial\mathcal{O}_i$, where we obtain the PDE

$$\sigma[\phi_i] \nabla \chi_{\mathcal{O}_i}^\epsilon = - \frac{(1 - \delta) \chi_{\mathcal{O}_i}^\epsilon + \delta}{1 - \delta} \text{div} \sigma[\phi_i] + \frac{1}{\epsilon} \eta^\delta[v_i, \phi_i, v] \nu$$

with $\eta^\delta[v_i, \phi_i, v] = \frac{2\gamma}{1-\delta} ((1 - v_i)^2(v \circ \phi_i) - v_i^2(1 - v \circ \phi_i)) |\nabla v \circ \phi_i|$ and $\nu = \frac{\nabla v \circ \phi_i}{|\nabla v \circ \phi_i|}$, where for smooth shapes $\mathcal{S}_i$ we expect the first summand on the right hand side to be uniformly bounded in ϵ and δ ,

whereas for $\epsilon, \delta \rightarrow 0$ the scaled gradient of the smoothed characteristic function $\epsilon \nabla \chi_{\mathcal{O}_i}^\epsilon$ converges to the normal $\nu[\mathcal{S}_i]$ in the sense of measures, and ν converges to $\nu[\mathcal{S}]$. Thus, in the limit we recover an effective boundary condition $\sigma[\phi_i] \nu[\mathcal{S}_i] = \eta^{\delta \rightarrow 0}[v_i, \phi_i, v] \nu[\mathcal{S}]$ on $\partial \mathcal{O}_i$ for every $i = 1, \dots, n$, which is interlinked with the corresponding boundary conditions for the other deformations via the phase field constraint (cf. (2.2)).

Let us finally also state the Euler Lagrange conditions for the fully joint model (4.3). Besides conditions (5.1), which hold analogously, the necessary conditions for u_i, v_i are

$$\langle \delta_{v_i} \mathcal{E}_{\text{joint}}^{\gamma, \epsilon}, \vartheta \rangle = 0, \quad \langle \delta_{u_i} \mathcal{E}_{\text{joint}}^{\gamma, \epsilon}, \vartheta \rangle = 0$$

with $\delta_{v_i} \mathcal{E}_{\text{joint}}^{\gamma, \epsilon} = \frac{1}{n} \sum_{i=1}^n (\delta_{v_i} \mathcal{E}_{\text{AT}}^\epsilon + \gamma \delta_{v_i} \mathcal{F}^\epsilon)$ and $\delta_{u_i} \mathcal{E}_{\text{joint}}^{\gamma, \epsilon} = \frac{1}{n} \sum_{i=1}^n \delta_{u_i} \mathcal{E}_{\text{AT}}^\epsilon$, where

$$\begin{aligned} \langle \delta_{v_i} \mathcal{E}_{\text{AT}}^\epsilon[u_i, v_i], \vartheta \rangle &= \int_{\Omega} 2\beta v_i \vartheta |\nabla u_i|^2 + 2\nu \left(\epsilon \nabla v_i \cdot \nabla \vartheta + \frac{1}{4\epsilon} (v_i - 1) \vartheta \right) dx, \\ \langle \delta_{v_i} \mathcal{F}^\epsilon[v_i, \phi_i, v], \vartheta \rangle &= \frac{2}{\epsilon} \int_{\Omega} ((v \circ \phi_i)^2 (v_i - 1) + v_i (1 - v \circ \phi_i)^2) \vartheta dx, \\ \langle \delta_{u_i} \mathcal{E}_{\text{AT}}^\epsilon[u_i, v_i], \zeta \rangle &= \int_{\Omega} 2\alpha (u_i - u_i^0) \zeta + 2\beta v_i^2 \nabla u_i \cdot \nabla \zeta dx \end{aligned}$$

for scalar test functions ϑ and ζ .

6. Existence results for the phase field models. In this section we will consider the phase field model (4.2) and prove for fixed $\epsilon > 0$ an existence result for the phase field v describing an average $\mathcal{S}$ and a set of deformations $(\phi_i)_{i=1, \dots, n}$ which match given shapes $\mathcal{S}_i$ represented by given phase fields v_i with the average shape $\mathcal{S}$. As already mentioned in the introduction, we restrict here to the three dimensional case and choose the fixed image domain $\Omega = [0, 1]^3$ for ease of presentation. By $W^{n,p}$ we denote the usual Sobolev space of functions with weak derivatives up the order n in L^p . We obtain the following theorem:

THEOREM 6.1 (Existence of a phase field shape average). *Suppose $d = 3$, $\epsilon, \delta, \gamma, \mu > 0$, and consider a set of admissible deformations*

$$\begin{aligned} \mathcal{A} := \{ \phi : \Omega \rightarrow \Omega \mid \phi \in W^{1,p}(\Omega), \text{cof} \mathcal{D}\phi \in L^q(\Omega), \\ \det \mathcal{D}\phi \in L^r(\Omega), \det \mathcal{D}\phi > 0 \text{ a.e. in } \Omega, \phi = \mathbb{1} \text{ on } \partial\Omega \} \end{aligned}$$

on a uniform image domain $\Omega = [0, 1]^3$, where $p, q > 3$ and $r > 1$. Furthermore, we consider the elasticity model described in § 2, where the integrand $\hat{W}$ appearing in the elastic energy $\mathcal{W}^\epsilon[\mathcal{O}_i, \phi_i]$ is supposed to be convex and there exist constants $\kappa, s \in \mathbb{R}$, $\kappa > 0$, and $s > \frac{2q}{q-3}$ such that $\bar{W}(I_1, I_2, I_3) \geq \kappa(I_1^{\frac{p}{2}} + I_2^{\frac{q}{2}} + I_3^r + I_3^{-s})$ for all $I_1, I_2 \in \mathbb{R}$ and $I_3 \in \mathbb{R}^+$. If the input phase fields $(v_i)_{i=1, \dots, n}$ lie in $W^{1,2}(\Omega)$ with $0 \leq v_i \leq 1$, then the energy

$$\mathcal{E}^{\gamma, \epsilon}[v, (\phi_i)_{i=1, \dots, n}] = \frac{1}{n} \sum_{i=1}^n (\mathcal{W}^\epsilon[\mathcal{O}_i, \phi_i] + \gamma \mathcal{F}^\epsilon[v_i, \phi_i, v]) + \mu \mathcal{L}^\epsilon[v]$$

attains its minimum over phase fields v in $W^{1,2}(\Omega)$ and n -tuples $(\phi_i)_{i=1, \dots, n}$ of deformations in $\mathcal{A}^n$. The minimizing v and ϕ_i for $i = 1, \dots, n$ are embedded in classical function spaces, namely $v \in C^{1,\alpha}(\bar{\Omega})$, $\phi_i \in C^{0,\beta}(\bar{\Omega})$, $v \circ \phi_i \in C^{0,\beta}$ for all $0 < \alpha < 1 - \frac{3}{s+1}$, $0 < \beta < 1 - \frac{3}{p}$. Furthermore, the minimizing deformations are homeomorphisms.

Proof. Apparently, the total energy is bounded below by zero. Also, $v \equiv 0$ and $\phi_i \equiv \mathbb{1}$, $i = 1, \dots, n$, show that there are phase fields and deformations for which the energy is finite.

Let $((\phi_i^k)_{i=1,\dots,n}, v^k)_{k=1,\dots}$ be a minimizing sequence. In what follows we will replace minimizing sequences by subsequences without any explicit subsequence indexing.

Ball [5] has shown that—due to the growth condition on $\bar{W}$ — ϕ_i^k is a homeomorphism with $\det \mathcal{D}\phi_i^k > 0$ a. e. and that the transformation rule

$$\int_{\Omega} f \circ \phi_i^k \det \mathcal{D}\phi_i^k \, dx = \int_{\Omega} f \, dx$$

holds if any of both integrals exists. Also, for elastic deformations with finite elastic energy and any bounded subset $\mathcal{B} \subset W^{1,t}(\Omega)$ the closure of $\mathcal{B}_{\phi_i^k} := \{f \circ \phi_i^k \mid f \in \mathcal{B}\}$ is compact in $L^{\tilde{t}}(\Omega)$ for all $\tilde{t} < \frac{s}{s+1} \left(\frac{1}{t} - \frac{1}{3}\right)^{-1}$, since $f \circ \phi_i^k$ is integrable for ϕ_i^k being a homeomorphism and by the transformation rule and Hölder's inequality,

$$\begin{aligned} \|f \circ \phi_i^k\|_{L^{\tilde{t}}}^{\tilde{t}} &= \int_{\Omega} |f \circ \phi_i^k|^{\tilde{t}} \, dx = \int_{\Omega} |f|^{\tilde{t}} \frac{1}{\det \mathcal{D}\phi_i^k \circ (\phi_i^k)^{-1}} \, dx \\ &\leq \left(\int_{\Omega} |f|^{\frac{\tilde{t}(s+1)}{s}} \, dx \right)^{\frac{s}{s+1}} \left(\int_{\Omega} \frac{1}{(\det \mathcal{D}\phi_i^k \circ (\phi_i^k)^{-1})^{s+1}} \, dx \right)^{\frac{1}{s+1}} = \|f\|_{L^{\frac{s+1}{\tilde{t}}}}^{\tilde{t}} \left(\int_{\Omega} (\det \mathcal{D}\phi_i^k)^{-s} \, dx \right)^{\frac{1}{s+1}}, \end{aligned}$$

where $W^{1,t}(\Omega)$ is compactly embedded in $L^{\frac{s+1}{\tilde{t}}}$, and the last integral is bounded by the elastic energy $\mathcal{W}^e[\mathcal{O}_i, \phi_i^k]$.

Now we construct a different minimizing sequence, still denoted $((\phi_i^k)_{i=1,\dots,n}, v^k)_{k=1,\dots}$, by letting $v^k = v[(\phi_i^k)_{i=1,\dots,n}]$ be the minimizer of $\mathcal{E}^{\gamma,\epsilon}[\cdot, (\phi_i^k)_{i=1,\dots,n}]$. However, we first have to verify the existence of such a minimizer v^k . One way to see this - which we find particularly instructive - is to investigate more closely compactness of the concatenations of phase fields and deformations with bounded elastic energy. Hence, for given $(\phi_i^k)_{i=1,\dots,n}$ let $v^{j,k}$ be a minimizing sequence of $\mathcal{E}^{\gamma,\epsilon}[\cdot, (\phi_i^k)_{i=1,\dots,n}]$. Due to the boundedness of $\mathcal{L}^e[v^{j,k}]$, $v^{j,k}$ is bounded in $W^{1,2}(\Omega)$ so that for a subsequence $v^{j,k} \rightharpoonup v^k$ in $W^{1,2}(\Omega)$ for some $v^k \in W^{1,2}(\Omega)$. Since the lower semi-continuity of $\mathcal{L}^e$ is obvious, for v^k to be the minimizer it only remains to prove lower semi-continuity of $\mathcal{F}^e$ as $v^{j,k} \rightharpoonup v^k$. We have

$$\begin{aligned} \mathcal{F}^e[v_i, \phi_i^k, v^{j,k}] &= \mathcal{F}^e[v_i, \phi_i^k, (v^{j,k} - v^k) + v^k] = \mathcal{F}^e[v_i, \phi_i^k, v^k] \\ &\quad + \int_{\Omega} (v_i^2 + (1 - v_i)^2) ((v^k - v^{j,k}) \circ \phi_i^k)^2 \, dx + 2 \int_{\Omega} (v_i^2 (v^k - 1) \circ \phi_i^k + (1 - v_i)^2 v^k \circ \phi_i^k) (v^{j,k} - v^k) \circ \phi_i^k \, dx. \end{aligned}$$

The second term is larger than or equal to zero, while for a subsequence the final term converges to zero, as it follows from the following estimate

$$\begin{aligned} &\left| \int_{\Omega} (v_i^2 (v^k - 1) \circ \phi_i^k + (1 - v_i)^2 v^k \circ \phi_i^k) (v^{j,k} - v^k) \circ \phi_i^k \, dx \right| \\ &\leq (\|v_i\|_{L^6}^2 \|v^k \circ \phi_i^k - 1\|_{L^3} + \|1 - v_i\|_{L^6}^2 \|v^k \circ \phi_i^k\|_{L^3}) \|(v^{j,k} - v^k) \circ \phi_i^k\|_{L^3}. \end{aligned}$$

In fact, still not using the strict bounds for v_i , we already observe by Sobolev embedding that v_i is uniformly bounded in $L^6(\Omega)$. Due to $s > 2$, we can apply our above compactness argument for $t = 2$ and $\tilde{t} = 3$ to v^k and $v^{j,k} - v^k$ so that the right hand side converges to zero for a subsequence.

For the truncated phase field $\tilde{v}^k = \max(0, \min(1, v^k))$, we observe $\mathcal{E}^{\gamma,\epsilon}[v^k, (\phi_i^k)_{i=1,\dots,n}] \geq \mathcal{E}^{\gamma,\epsilon}[\tilde{v}^k, (\phi_i^k)_{i=1,\dots,n}]$. Since v^k is already a minimizer, this implies that the sequence $(v^k)_{k=1,\dots}$ is uniformly bounded, i. e. $0 \leq$

$v^k \leq 1$. Furthermore, for each $k = 1, \dots$ the phase field v^k satisfies the Euler Lagrange equation

$$-\epsilon \mu \Delta v^k = -\frac{\mu}{4\epsilon}(v^k - 1) - \frac{\gamma}{\epsilon} \sum_{i=1}^n (v^k(1 - v_i \circ (\phi_i^k)^{-1})^2 - (v_i \circ (\phi_i^k)^{-1})^2(1 - v^k)) |\det \mathcal{D}(\phi_i^k)^{-1}| \quad (6.1)$$

in a weak sense. The right hand side of (6.1) is uniformly bounded in $L^{s+1}(\Omega)$ (due the above transformation rule with $f \equiv |\det \mathcal{D}(\phi_i^k)^{-1}|^{s+1} = |\det \mathcal{D}\phi_i^k \circ (\phi_i^k)^{-1}|^{s+1}$), and applying classical elliptic regularity theory [33] we observe that $(v^k)_{k=1, \dots}$ is uniformly bounded in $W^{2, s+1}(\Omega)$.

Next, we consider the functional

$$(\phi_i)_{i=1, \dots, n} \mapsto \mathcal{E}^{\gamma, \epsilon}[v[(\phi_i)_{i=1, \dots, n}], (\phi_i)_{i=1, \dots, n}]$$

on $\mathcal{A}^n$. Due to the growth condition on $\bar{W}$ we get that $((\mathcal{D}\phi_i^k, \text{cof}\mathcal{D}\phi_i^k, \det \mathcal{D}\phi_i^k))_{k=1, \dots}$ is uniformly bounded in $L^p \times L^q \times L^r$. By Poincaré's inequality applied to $(\phi_i^k - 1)$ we obtain that $\{\phi_i^k\}_{k \in \mathbb{N}}$ is uniformly bounded in $W^{1, p}(\Omega)$. Thus, we can extract a weakly convergent subsequence. Applying the nowadays classical compensated compactness result by Ball [4] we observe that

$$(\mathcal{D}\phi_i^k, \text{cof}\mathcal{D}\phi_i^k, \det \mathcal{D}\phi_i^k) \rightharpoonup (\mathcal{D}\phi_i, \text{cof}\mathcal{D}\phi_i, \det \mathcal{D}\phi_i)$$

in $L^p \times L^q \times L^r$ for some $\phi_i \in \mathcal{A}$. This holds for any $i = 1, \dots, n$. In addition, $\mathcal{W}^\epsilon[\mathcal{O}_i, \cdot]$ is sequentially weakly lower semicontinuous.

Furthermore, from the uniform boundedness of $v^k = v[(\phi_i^k)_{i=1, \dots, n}]$ in $W^{2, s+1}(\Omega)$ as $k \rightarrow \infty$ it follows by Sobolev embedding that $v^k \rightarrow v$ in $C^{1, \alpha}(\bar{\Omega})$ for some $v \in C^{1, \alpha}(\bar{\Omega})$, which holds for any $0 \leq \alpha < 1 - \frac{3}{s+1}$. Hence, $\mathcal{L}^\epsilon[v^k]$ converges to $\mathcal{L}^\epsilon[v]$. Next, we justify the continuity of $\mathcal{F}^\epsilon$ on a subsequence, i.e.

$$\mathcal{F}^\epsilon[v_i, \phi_i^k, v[(\phi_i^k)_{i=1, \dots, n}]] \rightarrow \mathcal{F}^\epsilon[v_i, \phi_i, v]$$

for $k \rightarrow \infty$. Due to Sobolev's embedding theorem, a subsequence of the ϕ_i^k converges strongly in $C^{0, \beta}(\bar{\Omega})$ for $\beta < 1 - \frac{3}{p}$. For x and $x + z$ in $\bar{\Omega}$ we estimate

$$|(v^k \circ \phi_i^k)(x + z) - (v^k \circ \phi_i^k)(x)| \leq C \|v^k\|_{C^{0, 1}} \|\phi_i^k\|_{C^{0, \beta}} |z|^\beta.$$

Thus, we obtain that the concatenation of v^k and ϕ_i^k is uniformly Hölder continuous for a positive Hölder exponent β . Now, we can apply the Arzela Ascoli theorem to establish the pointwise convergence of $v^k \circ \phi_i^k$ against $v \circ \phi_i$ for another subsequence and thence the requested continuity of the penalty functional $\mathcal{F}^\epsilon$. Finally, from the continuity of $\phi_i \mapsto \mathcal{F}^\epsilon[v_i, \phi_i, v[(\phi_i)_{i=1, \dots, n}]]$ and the lower semicontinuity of $\frac{1}{n} \sum_{i=1}^n \mathcal{W}^\epsilon[\mathcal{O}_i, \phi_i] + \mu \mathcal{L}^\epsilon[v[(\phi_i)_{i=1, \dots, n}]]$ we deduce the sequentially weak lower semicontinuity of the functional

$$(\phi_i)_{i=1, \dots, n} \mapsto \mathcal{E}^{\gamma, \epsilon}[v[(\phi_i)_{i=1, \dots, n}], (\phi_i)_{i=1, \dots, n}].$$

From this, the existence of a minimizing phase field v and a minimizing set of elastic deformations $(\phi_i)_{i=1, \dots, n}$ follows immediately. Furthermore, the stated regularity is a consequence of the above arguments, and as already mentioned the homeomorphism property is a direct consequence of the boundedness of the elastic energy [5]. For further details on that we refer to [62]. $\square$

Furthermore, there is a similar existence result for the joint model (4.3):

THEOREM 6.2 (Existence of minimizers for the joint model). *Let $\alpha, \beta, \nu > 0$ and assume that the assumptions of Theorem 6.1 hold. Furthermore, let $u_i^0 \in L^2(\Omega)$ for $i = 1, \dots, n$. Then the energy*

$$\mathcal{E}_{\text{joint}}^{\gamma, \epsilon}[v, (u_i, v_i, \phi_i)_{i=1, \dots, n}] = \frac{1}{n} \sum_{i=1}^n (\mathcal{E}_{AT}^\epsilon[u_i, v_i, u_i^0] + \mathcal{W}^\epsilon[\Omega, \phi_i] + \gamma \mathcal{F}^\epsilon[v_i, \phi_i, v]) + \mu \mathcal{L}^\epsilon[v]$$

attains its minimum over n -tuples of images $u_i \in W^{1,2}(\Omega)$, phase fields $v_i \in W^{1,2}(\Omega)$, and deformations $\phi_i \in \mathcal{A}$ with $i = 1, \dots, n$, and over phase fields $v \in W^{1,2}(\Omega)$. The minimizing v, v_i and ϕ_i for $i = 1, \dots, n$ are embedded in classical function spaces, namely $v_i \in C^{1,\alpha'}(\bar{\Omega})$, $v \in C^{1,\alpha''}(\bar{\Omega})$, $\phi_i \in C^{0,\beta'}(\bar{\Omega})$, $v \circ \phi_i \in C^{0,\beta'}$ for all $0 < \alpha' < 1$, $0 < \alpha'' < 1 - \frac{3}{s+1}$, $0 < \beta' < 1 - \frac{3}{p}$. Furthermore, the minimizing deformations are homeomorphisms.

Proof. The required arguments are closely related to those in the proof of Theorem 6.1. Hence, we give here only a brief sketch. Let $((u_i^k, v_i^k, \phi_i^k)_{i=1,\dots,n}, v^k)_{k=1,\dots}$ be a minimizing sequence, where we assume that for fixed $(u_i^k, \phi_i^k)_{i=1,\dots,n}, v^k$ the n -tuple of phase fields $(v_i^k)_{i=1,\dots,n}$ is a minimizer over all n -tuples of phase fields in $W^{1,2}(\Omega)$. The existence of these phase fields is straightforward, and once more by truncation we observe that $0 \leq v_i^k \leq 1$ (and that just as above v_i^k converges strongly in $C^{1,\alpha'}$ for $0 < \alpha' < 1$). Hence, v_i^k is now an admissible phase field for the description of the input shapes in Theorem 6.1. Thus, we can again modify the minimizing sequence, and suppose that for fixed $(v_i^k)_{i=1,\dots,n}$ the other components $(u_i^k, \phi_i^k)_{i=1,\dots,n}, v^k$ minimize the global energy. To prove this we follow exactly the above proof and remark that the lower semicontinuity of $\mathcal{E}_{\text{AT}}^\epsilon[\cdot, v_i^k]$ is obvious. Now, we let $k \rightarrow \infty$ for the modified sequence, and repeating arguments we observe the requested lower semicontinuity of the total energy. Indeed, to establish the continuity of $\mathcal{F}^\epsilon$ on a subsequence we make use of the a priori ensured L^∞ bound for the v_i^k , which is required to prove the strong $C^{1,\alpha''}$ regularity of the v^k . $\square$

7. Finite element implementation. We now describe the actual spatial discretization by finite elements and the implementation of a discrete multi-scale method. To simplify the exposition, we will confine ourselves to the central model derived in this paper and consider the energy (4.2). We consider the images u_i , phase fields v, v_i , and deformations ϕ_i as being represented by continuous, piecewise multilinear (trilinear in 3D and bilinear in 2D) finite element functions on the image domain $\Omega = [0, 1]^d$. Each pixel or voxel value corresponds to a node of the regular mesh.

For ease of implementation we suppose dyadic resolutions of the images with $2^L + 1$ pixels or voxels in each direction corresponding to a grid size $h = 2^{-L}$. Concerning notation, in contrast to the non-discretized functions, which were represented by lower case letters in the above exposition, we use upper case letters to denote their counterparts in the finite-dimensional finite element function spaces. For their corresponding nodal vector representation with respect to the canonical basis of hat functions we use bold letters. Specifically, the finite element approximation of a phase field v is represented by $\mathbf{V} = (\mathbf{V}_j)_{j \in I_h}$ with $V = \sum_{j \in I_h} \mathbf{V}_j \varphi_j$, where $\{\varphi_j\}_{j \in I_h}$ is the nodal basis and I_h the grid node index set corresponding to the grid with grid size h . For a discretized deformation Φ and the corresponding nodal vector Φ we get $\Phi = \sum_{j \in I_h} \sum_{k=1,\dots,d} \Phi_{jk} \varphi_j e_k$, where $e_1, \dots, e_d$ is the canonical basis in $\mathbb{R}^d$.

Fully discrete nonlinear system of equations. From the spatially continuous Euler Lagrange conditions (5.1) for v and $\phi_1, \dots, \phi_n$ we derive a fully discrete nonlinear system of equations. A finite element discretization of the first equation in (5.1) leads to the system of equations

$$\begin{aligned} 0 = & \left(\frac{\gamma}{n\epsilon} \sum_{i=1}^n M \left[((1 - V_i \circ \Phi_i^{-1})^2 + (V_i \circ \Phi_i^{-1})^2) |\det(\mathcal{D}\Phi_i^{-1})| \right] + \mu\epsilon L[1] + \frac{\mu}{4\epsilon} M[1] \right) \mathbf{V} \\ & - \left(\frac{\gamma}{n\epsilon} \sum_{i=1}^n M \left[(V_i \circ \Phi_i^{-1})^2 |\det(\mathcal{D}\Phi_i^{-1})| \right] + \frac{\mu}{4\epsilon} M[1] \right) \mathbf{1} =: A_{V_i, \Phi_i} \mathbf{V} - b_{V_i, \Phi_i}, \end{aligned} \quad (7.1)$$

where for some weighting function $\omega : \Omega \rightarrow \mathbb{R}$ the generalized mass matrix $M[\omega]$ and the generalized stiffness matrix $L[\omega]$ are defined as

$$M[\omega] = \left(\int_{\Omega} \omega \varphi_i \varphi_j \, dx \right)_{ij}, \quad L[\omega] = \left(\int_{\Omega} \omega \nabla \varphi_i \cdot \nabla \varphi_j \, dx \right)_{ij},$$

and where $\mathbf{1}$ is the nodal vector with entries all equal to 1. This equation is obviously linear in $\mathbf{V}$ and nonlinear in Φ_i . The second equation in (5.1) yields the finite element discretization

$$\begin{aligned} 0 = & \frac{2\gamma}{n\epsilon} \int_{\Omega} ((1 - V_i)^2 (V \circ \Phi_i) - V_i^2 (1 - V \circ \Phi_i)) \Psi \cdot (\nabla V \circ \Phi_i) dx \\ & + \frac{1}{n} \int_{\Omega} \left((1 - \delta) \chi_{\mathcal{O}_i}^\epsilon + \delta \right) W_{,A}(\mathcal{D}\Phi_i) : \mathcal{D}\Psi dx, \end{aligned} \quad (7.2)$$

where Ψ is running over all vector valued basis functions $\varphi_j e_k$. The unknowns are $\mathbf{V} \in \mathbb{R}^N$ and the set of discrete deformations $\Phi_1, \dots, \Phi_n \in \mathbb{R}^{dN}$, where $N = (2^L + 1)^d$ is the number of grid nodes.

Numerical quadrature. Integral evaluation will be performed by Gaussian quadrature of third order on each grid cell. At various places in (7.1) and (7.2) we have to evaluate pushforwards $(V \circ \Phi_i)$ or pullbacks $(V_i \circ \Phi_i^{-1})$ of the discretized phase fields V and V_i under a discretized deformation Φ_i . In both cases we replace the exact evaluation of the integrals by simple and effective quadrature schemes. Let us emphasize that it is not sufficient just to compute nodal interpolants $\mathcal{I}_h(V \circ \Phi_i)$, $\mathcal{I}_h(V_i \circ \Phi_i^{-1})$ of $V \circ \Phi_i$ and $V_i \circ \Phi_i^{-1}$, respectively, where $\mathcal{I}_h$ denotes the classical Lagrangian interpolation. In fact, artificial displacements near the shape edges are then observed, accompanied with strong tensions and generated while alternating between optimizing the average phase field V and the deformations Φ_i . In our algorithm, we compute $V \circ \Phi_i$ exactly at the quadrature points and approximate $V_i \circ \Phi_i^{-1}$ by $V_i \circ \mathcal{I}_h^\Delta(\Phi_i^{-1})$. Here, $\mathcal{I}_h^\Delta$ is a slightly modified interpolation operator motivated by an effective implementation of a discretized inverse deformation, which proceeds as follows. We map each grid cell under the deformation Φ_i onto the image domain. Next we identify all grid nodes which are located within this deformed cell. Now, we would like to apply interpolation within this deformed cell to retrieve the requested nodal values of the discretized inverse deformation Φ_i . However, inversion of multilinear deformation leads to nonlinear equations. To avoid this shortcoming, we cut the cell into virtual simplices. We then replace the regular cell in the retrieval algorithm by the set of simplices and obtain piecewise affine inverse mappings via linear interpolation. This scheme is designed to be consistent on locally affine transformations, which is what is required not to destroy the expected accuracy of a method based on multi-linear finite elements.

Alternating minimization approach. The complexity of the minimization problem suggests to perform a fixed point iteration, alternatingly computing the different unknowns. In each iteration, we first solve for the phase field V and afterwards for the deformations Φ_i . The updated values are then used in the next iteration until the algorithm converges. Equation (7.1) for fixed deformations Φ_i is linear in the vector $\mathbf{V}$ and is therefore solved by conjugate gradient iteration. Concerning the deformations Φ_i and equation (7.2) for fixed phase field V , a direct solution is ruled out due to the involved nonlinear elasticity and the nonlinearity in the concatenation of deformations with phase fields. We resort to a regularized gradient descent for the discrete counterpart $\mathcal{E}_h^{\gamma,\epsilon}$ of the energy $\mathcal{E}^{\gamma,\epsilon}$, updating the discrete deformation vector Φ_i according to

$$\Phi_i = \Phi_i^{\text{old}} - \tau \text{grad}_{\Phi_i} \mathcal{E}_h^{\gamma,\epsilon}, \quad (7.3)$$

where τ is the current step size and grad_{Φ_i} represents the finite element approximation to the regularized gradient with respect to the weighted H^1 inner product

$$(\Psi_1, \Psi_2)_\sigma := (\Psi_1, \Psi_2)_{L^2} + \frac{\sigma^2}{2} (\mathcal{D}\Psi_1, \mathcal{D}\Psi_2)_{L^2}. \quad (7.4)$$

More precisely, $\text{grad}_{\Phi_i} \mathcal{E}_h^{\gamma,\epsilon}$ is the vector valued finite element function implicitly defined as the solution of $(\text{grad}_{\Phi_i} \mathcal{E}_h^{\gamma,\epsilon}, \Psi)_\sigma = \langle \delta_{\Phi_i} \mathcal{E}_h^{\gamma,\epsilon}, \Psi \rangle$ for all discrete displacement fields Ψ . The right hand side of this system of equations coincides with the right hand side of (7.2). Here, the effect of the smoothing metric (7.4) is related to the convolution with a Gauss kernel or equivalently the application of one time step

for the heat equation semigroup. Consequently, information flow across the image is enhanced, and the deformations equilibrate faster. Also, the descent algorithm becomes more resistant to being trapped in local minima. Different from Miller et al. [47], injectivity is already ensured by the variational approach itself. The H^1 -gradient is used to avoid numerical oscillations in conjunction with the concatenation of discrete deformations and discrete phase fields. In the algorithm it turned out to be sufficient to approximate the gradient performing a single multigrid V-cycle for the system

$$\left(M[1] + \frac{\sigma^2}{2} L[1] \right) (\text{grad}_{\Phi_i} \mathcal{E}_h^{\gamma, \epsilon})_j = (\langle \delta_{\Phi_i} \mathcal{E}_h^{\gamma, \epsilon}, \Psi \rangle)_j ,$$

where $(\cdot)_j$ for $j = 1, \dots, d$ indicates one particular component sub vector of the nodal vector and with a slight misuse of notation $\text{grad}_{\Phi_i} \mathcal{E}_h^{\gamma, \epsilon}$ and $\langle \delta_{\Phi_i} \mathcal{E}_h^{\gamma, \epsilon}, \Psi \rangle$ are interpreted as nodal vectors. For details we refer to [21, 26]. The step size τ in (7.3) is obtained according to Armijo's rule, which ensures sufficient agreement between the objective functional and its linearization. If the actually observed energy decay in one time step is smaller than $\frac{1}{4}$ of the decay estimated from the gradient (the Armijo condition is then violated), then the time step τ is halved for the next trial, else it is doubled as often as possible without violating the Armijo condition.

Multi-scale optimization. The variational problem considered here is highly nonlinear, and the proposed scheme is expected to have very slow convergence; also it might end up in some nearby local minimum. Here, a multi-level approach (initial optimization on a coarse scale and successive refinement) turns out to be indispensable in order to accelerate convergence and not to be trapped in local minima far from the global minimum. Due to our assumption of a dyadic resolution $2^L + 1$ in each grid direction we are able to build a hierarchy of grids with $2^l + 1$ nodes in each direction for $l = L, \dots, 0$. Via a trivial restriction operation we restrict every finite element function to any of these coarse grid spaces. Starting the optimization on a coarse grid, the results from coarse scales are successively prolonged onto the next grid level for a refinement of the solution. Hence, the construction of a multigrid hierarchy allows to solve coarse scale problems in our multi-scale approach on coarse grids. Since the width ϵ of the diffusive phase field edges should naturally scale with the grid width h , we choose $\epsilon = Ch$ on each level for some constant factor C . In our implementation $C = 1$.

Algorithm. The entire algorithm in pseudo code notation reads as follows:

```

EnergyRelaxation  $((U_i^0)_{i=1, \dots, n})$  {
  initialize  $\Phi_i = \mathbb{1}$  on grid level  $l_0$  for all  $i = 1, \dots, n$ ;
  for grid level  $l = l_0$  to  $L$  {
    do {
      segment the images  $(U_i^0)_{i=1, \dots, n}$  to obtain phase fields  $(V_i)_{i=1, \dots, n}$ ;
       $\mathbf{V}^{\text{old}} = \mathbf{V}$ ;
      solve the linear system
         $A_{V_i, \Phi_i} \mathbf{V} = b_{V_i, \Phi_i}$ 
      for the phase field vector  $\mathbf{V}$ ;
      for image  $i = 1$  to  $n$ 
        for count  $k = 1$  to  $K$  {
           $\Phi_i^{\text{old}} = \Phi_i$ ;
          perform a gradient descent step
             $\Phi_i = \Phi_i^{\text{old}} - \tau \text{grad}_{\Phi_i^{\text{old}}} \mathcal{E}_h^{\gamma, \epsilon}[\mathbf{V}, (\Phi_j)_{j=1, \dots, n}]$ 
          with Armijo step size control for  $\tau$ ;
        }
      } while  $(\sum_{i=1}^n |\Phi_i^{\text{old}} - \Phi_i| + |\mathbf{V}^{\text{old}} - \mathbf{V}| \geq \text{Threshold})$ ;
      if  $(l < L)$  prolongate  $\mathbf{V}$ ,  $\Phi_i$  for all  $i = 1, \dots, n$  onto the next grid level;
    }
  }
}
```

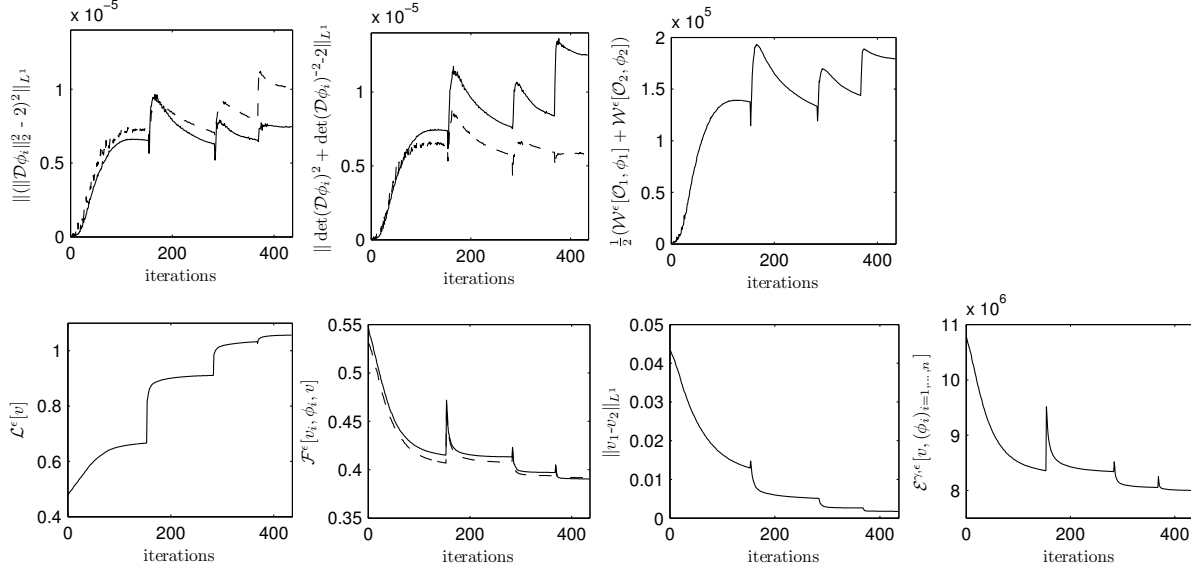


FIG. 7.1. For the example presented in Figure 2.4 we show the progression of the various energy contributions during the solution iteration. The top row shows the hyperelastic energy contributions due to length and volume variation (left and middle, respectively; the solid line corresponds to $i = 1$, the dashed one to $i = 2$) as well as the total hyperelastic energy (right), i. e. the sum of length and volume contributions weighted with $a_1 = a_3 = 10^{10}$. The bottom row shows the length regularization (left) and the mismatch penalty for both images (second graph), which are weighted with $\mu = 10^{-2}$ and $\gamma = 10^7$ respectively in the overall energy (right). In all graphs, the spikes correspond to the prolongation to the next grid level. The third graph in the bottom row shows the L^1 -difference between $v_1 \circ \phi_1^{-1}$ and $v_2 \circ \phi_2^{-1}$.

}
 It has turned out to be appropriate to choose $l_0 \approx 4$ as the coarsest grid level in our multi-scale approach. On each scale we apply the alternating minimization approach. In each inner iteration we solve once for the discrete phase field $\mathbf{V}$ and perform $K = 5$ gradient descent steps. As a threshold value for the convergence test we choose $\text{Threshold} = 2 \cdot 10^{-4}$. For the example considered in Figure 2.4, one iteration of our scheme on grid level $L = 9$ takes 10 seconds on a Pentium IV PC at 1.8 GHz running under Linux. The complete method typically converges after roughly 100 such iterations on each grid level. In Figure 7.1 we depict the progression of the various components of the energy $\mathcal{E}^{\gamma, \epsilon}$ for the averaging problem shown in Figure 2.4. The strong decay of the global energy at the beginning of the algorithm is clearly visible. Apparently, the mismatch penalty strongly dominates the total energy. Also, we show the L^1 -difference between $v_1 \circ \phi_1^{-1}$ and $v_2 \circ \phi_2^{-1}$, which strongly decreases, indicating a good match between both deformed shapes.

8. Applications. We have applied our shape averaging approach to various collections of 2D and 3D shapes and to image morphologies.

Shape averaging in 2D. As first illustrative examples, we computed the average of different 2D objects as shown in Figures 8.1 to 8.2. Furthermore, Figures 2.2 and 2.4 have already shown that due to the invariance of the hyperelastic energy with respect to local rotations the computed averages try to locally preserve isometries. Effectively, the different characteristics of the input shapes, both on the global and a local scale, are averaged in a physically intuitive way, and the scheme proves to be fairly robust due to the diffusive approximation based on the phase field model and the multi-scale relaxation. Nevertheless, a lack of topological equivalence of the given input shapes might lead to corresponding local artifacts in the shape average result, compare for instance the twentieth input shape in Figure 8.2 and

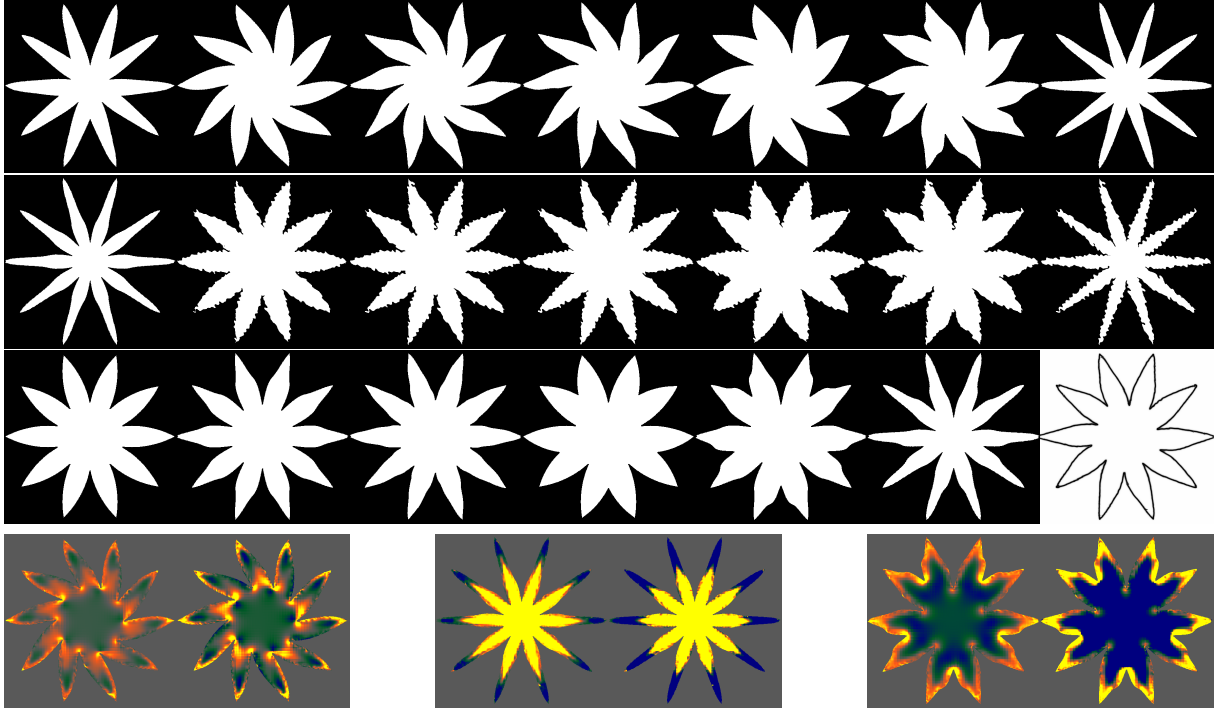


FIG. 8.1. 20 shapes “device7” from the MPEG7 shape database and their average phase field. The bottom line shows $|\mathcal{D}\phi_i|_2$ and $\det(\mathcal{D}\phi_i)$ for shape 2, 8, and 19, with ranges of $[0.8\sqrt{2}, 1.2\sqrt{2}]$ and $[0.8, 1.2]$ color-coded as . (Resolution 513×513 , $\gamma = 10^7$, $\mu = 10^{-2}$, $(a_1, a_2, a_3) = (10^6, 0, 10^6)$.)

its locally spurious impact on the average phase field between the two legs.

Figure 8.3 shows two more examples, using input shapes from [22] and the shape database at the Centre for Vision, Speech, and Signal Processing, University of Surrey. Averaging the hand shapes yields very similar results to the averages obtained in [22] and [30] as the Euclidean and the Fréchet mean of vectors of landmark positions, respectively. The average fish shape has also been computed in [14]. Note that our result preserves more fine structures as opposed to the quite rounded mean shape in [14].

Shape averaging in 3D. In what follows we will consider the averaging of 3D shapes originally given as triangulated surfaces and first converted to an implicit representation as binary images. A set of 48 kidneys and a set of 24 feet will serve as input data. The first five original kidneys and their computed average are shown in Figure 8.4. Local structures seem to be quite well represented and preserved during the averaging process compared to e.g. the average of kidney shapes in medial representation in [29].

Via the deformation of a given object onto the average during the averaging algorithm, the method also yields local and global information about the distance of the object to the average. For the first two of the original kidneys, the corresponding local elastic energy density or rather the three deformation invariants are depicted in Figure 8.5. The inside of the first kidney apparently gets slightly dilated, whereas the second one is compressed. Also, it can be observed that the dilation or compression is reduced at the shape boundaries, which is caused by the finite width phase field description of the edges: For the deformed phase fields to match, they all have to have the same thickness and hence may only be deformed significantly in tangential, but not in normal direction.

Naturally, any averaging will involve some smoothing, eliminating fine details which differ from shape to shape. It is hence of interest whether features, common to all shapes but differing slightly, pertain if the number of samples is increased: Figure 8.6 shows the result of averaging different numbers of kidneys.

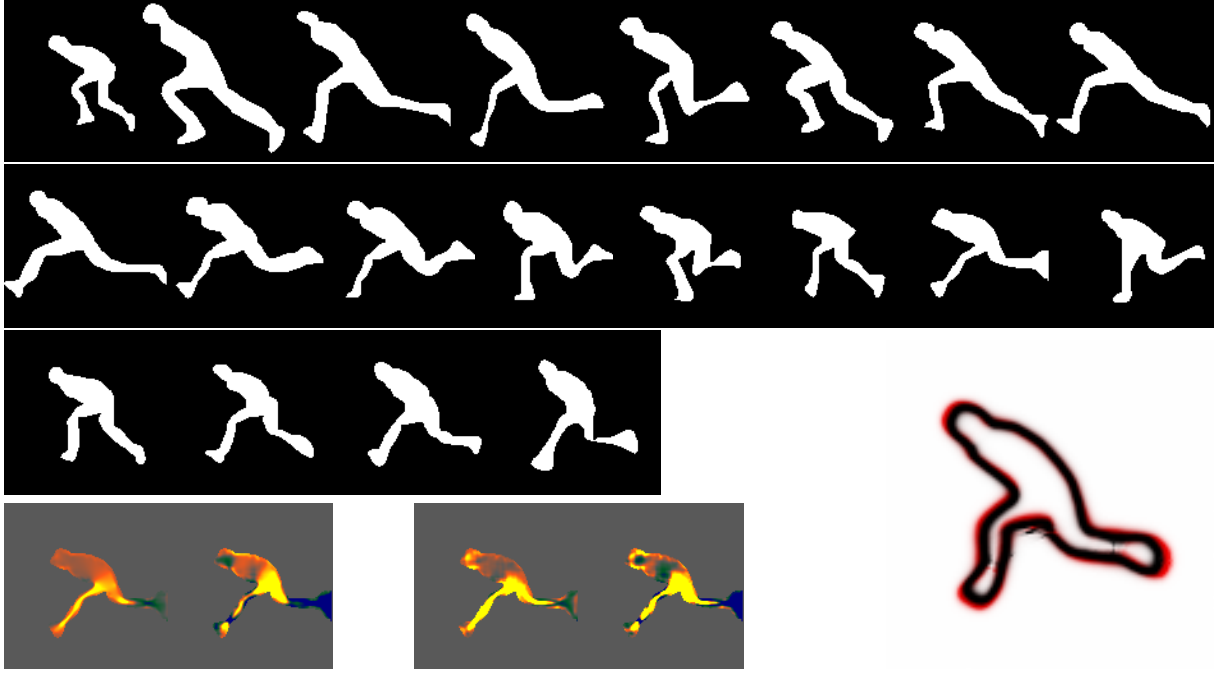


FIG. 8.2. 20 shapes “stef” from the MPEG7 shape database and their average phase field (bottom right) for elastic parameters $(a_1, a_2, a_3) = (10^7, 0, 10^6)$ (black, on top) and $(a_1, a_2, a_3) = (10^5, 0, 10^6)$ (red, underneath). The bottom line shows $|\mathcal{D}\phi_i|_2$ and $\det(\mathcal{D}\phi_i)$ for shape 15 in the case of $a_1 = 10^7$ and 10^5 , respectively, with ranges of $[0.8\sqrt{2}, 1.2\sqrt{2}]$ and $[0.8, 1.2]$ color-coded as . Obviously, the larger the ratio between the weights of volume and length variation penalty, the more elongated the shapes become. (Resolution 129×129 , $\gamma = 10^7$, $\mu = 10^{-2}$.)

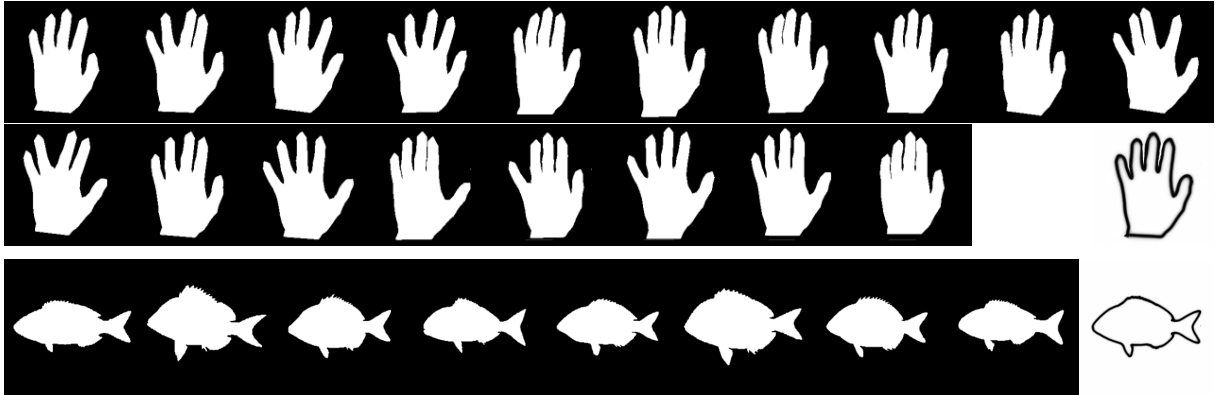


FIG. 8.3. Average of 18 hand and 8 fish silhouettes, taken from [22] and the shape database at the Centre for Vision, Speech, and Signal Processing, University of Surrey, respectively.

Also, we would like to know how the method performs for relatively large numbers of input shapes, since the lack of a triangle inequality for a hyperelastic “distance” measure prevents a law of large numbers. It is indeed observed that the middle dent of the average kidney is a little less pronounced than in each single kidney, even in the case of averaging just two kidneys. Also, the influence of each additional original shape seems significantly strong, however, a kidney-like shape is doubtlessly preserved up to the average of all 48 kidneys.

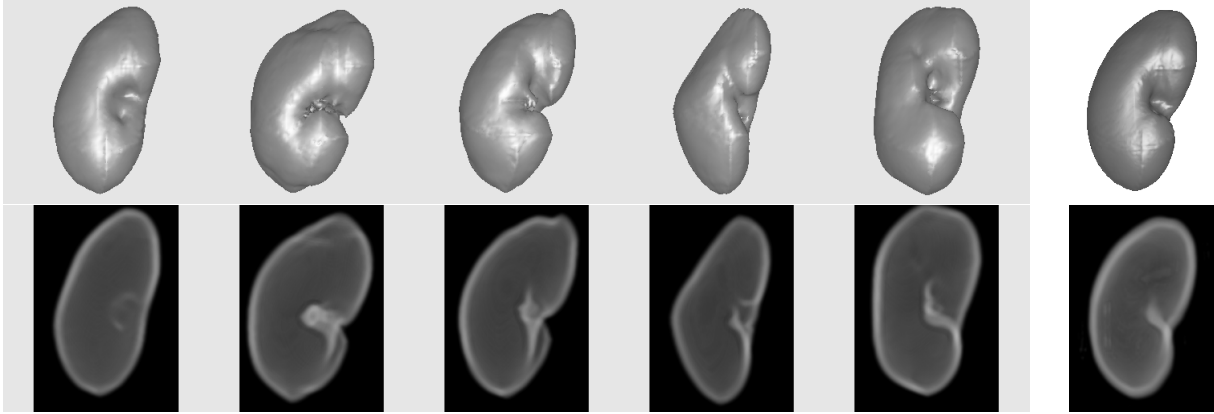


FIG. 8.4. Five segmented kidneys and their average (right). The lower row shows the volume renderings of the corresponding phase fields. (Result obtained for resolution $257 \times 257 \times 257$ and parameter values $\gamma = 10^7$, $\mu = 1$, $(a_1, a_2, a_3) = (10^8, 0, 10^7)$.)

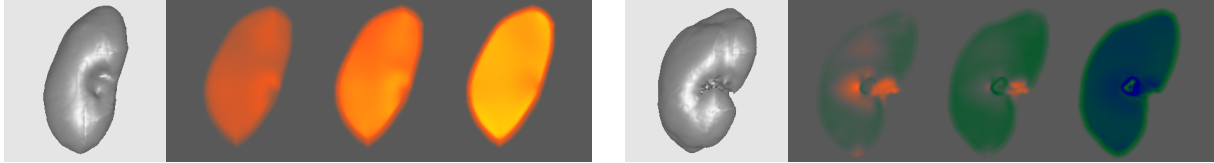


FIG. 8.5. Sagittal cross-section through the distribution of $|\mathcal{D}\phi_i|_2$, $|\text{cof}(\mathcal{D}\phi_i)|_2$, and $\det(\mathcal{D}\phi_i)$ for the first two kidneys from Figure 8.4 (the ranges of $[0.85\sqrt{3}, 1.15\sqrt{3}]$, $[0.85\sqrt{3}, 1.15\sqrt{3}]$, and $[0.85, 1.15]$ are color-coded as ). While the first kidney is dilated towards the average, the second is compressed. In the thin diffusive interface region, the dilation or compression is reduced.

The next example consists of a set of feet, where the average may help to design an optimal shoe. The 24 original feet are displayed in Figure 8.7. Their surface is colored according to the local distance to the surface of the computed average shape, which helps to identify regions of strong variation. For that purpose the foot shapes have been optimally aligned with the average for the final visualization. Let us emphasize that the algorithm itself robustly deals with even quite large rigid body motions. Apparently, the instep differs comparatively little between the given feet, whereas the toes show a rather strong variation. Note that—since we only display normal distance to the surface of the average foot—any potential tangential displacement is not visible, but could of course also be visualized when examining shape variation.

For real applications, one has to be careful when dealing with shapes of different volume. Depending on the chosen hyperelastic parameters, there may be a bias towards larger or smaller shapes, and appropriate parameters will have to be chosen carefully. Also, for too soft hyperelastic material models, buckling instabilities will occur during compression of large volume shapes. Some influence of different hyperelastic parameters is illustrated in Figure 8.2.

Weighted averaging. Returning back to the kidneys, it is also possible to compute a weighted average, where the deformation energy of the different input shapes is weighted differently. The averaging functional then is modified to

$$\mathcal{E}^{\gamma, \epsilon}[v, (\phi_i)_{i=1, \dots, n}] = \sum_{i=1}^n \left(\lambda_i \mathcal{W}^\epsilon[\mathcal{O}_i, \phi_i] + \frac{\gamma}{n} \mathcal{F}^\epsilon[v_i, \phi_i, v] \right) + \mu \mathcal{L}^\epsilon[v],$$

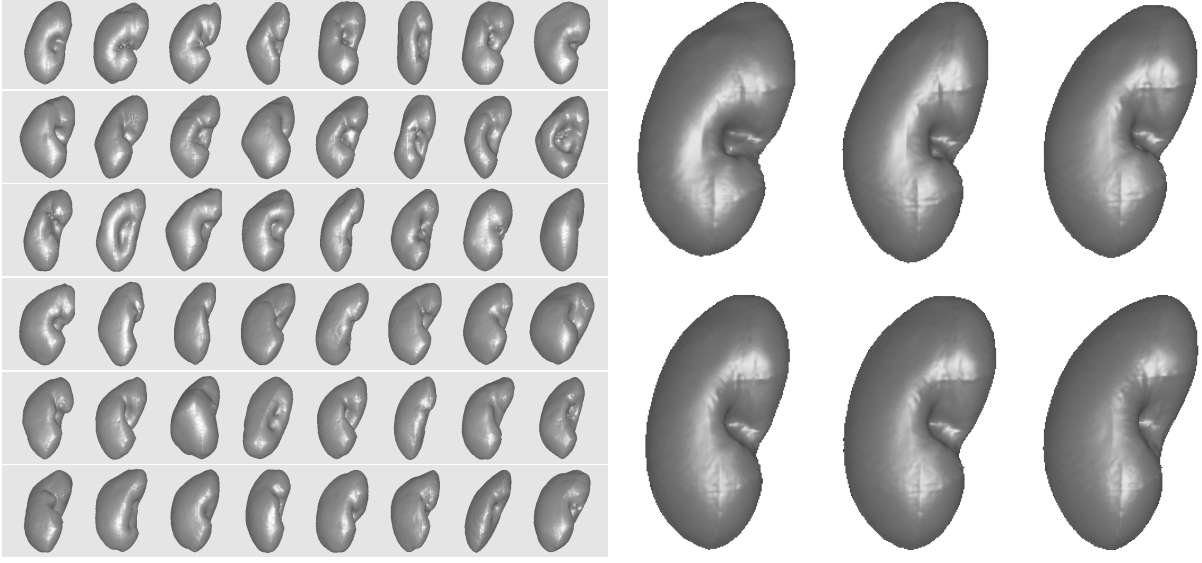


FIG. 8.6. On the left, 48 kidney shapes are shown. On the right, from top left to bottom right the averaged shape of the first two, four, five, six, eight and of all 48 kidneys are depicted. The parameter values are as for Figure 8.4.

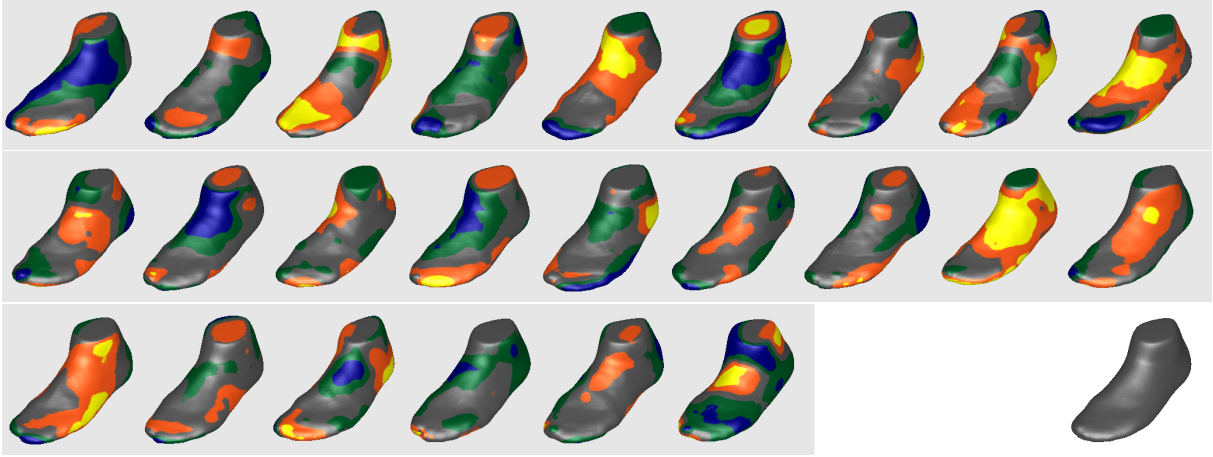


FIG. 8.7. 24 given foot shapes, textured with the distance to the surface of the average foot (bottom right). Values range from 6 mm inside the average foot to 6 mm outside, color-coded as . The front of the instep can be identified as a region of comparatively low variation. (Result obtained for resolution $257 \times 257 \times 257$ and parameter values $\gamma = 10^7$, $\mu = 1$, $(a_1, a_2, a_3) = (10^8, 0, 10^8)$.)

where the weights λ_i are nonnegative and add up to 1. Such a “nonlinear convex-combination” of three kidneys is presented in Figure 8.8.

Averaging image morphologies. To illustrate that the approach can also be applied to average image morphologies, the input of our final example consists of two-dimensional, transversal CT scans of the human thorax from four different patients (Figure 8.9, left). Unlike the previous examples, these images do not encode volumetric shapes homeomorphic to the unit ball, but contain far more complicated structures. Also, the quality of contrast differs between the images, and—even more problematic—the images do not show a one-to-one correspondence, i.e. several structures (the scapula, ribs, parts of the

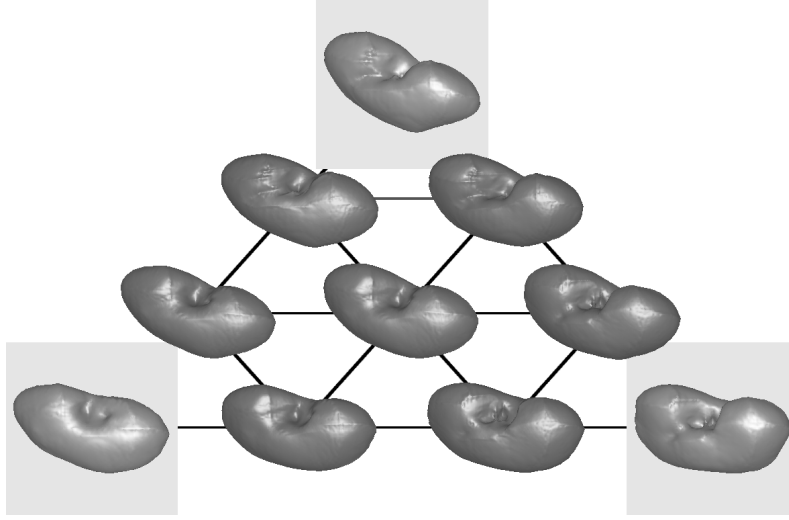


FIG. 8.8. Given the three kidney geometries placed in the corners of the triangle, seven differently weighted averages are placed at corresponding positions in the triangle.

liver) are only visible in some images, but not in others, implying that the underlying shapes are not even homeomorphic. Nevertheless, the algorithm manages to segment and align the main features (the heart, the spine, the aorta, the sternum, the ribs, the back muscles, the skin), yielding sensible average contours (Figure 8.9, right). In order to achieve this, we this time jointly segmented and averaged the original CT scans, i. e. we augmented our objective function (4.2) by the Ambrosio–Tortorelli energy (4.1) for each image and then alternatingly optimized for the v_i , u_i , v , and ϕ_i . The second to fourth column of Figure 8.9 depict the corresponding deformations ϕ_i and the deformation invariants. Obviously, the deformation behaves quite regularly: Not only is it homeomorphic, but also too large and distorting deformations are prevented by the hyperelastic regularization. This enables the method to be applied to images containing also distinct structures, whereas for viscous flow regularization as in [8, 18] such individual structures are at risk of being matched with anything nearby. The deformation energy is quite evenly distributed over the images and only peaks at pronounced features, where a local exact fit can be achieved (e. g. at the back muscles). Outside the thorax, the energy rapidly decreases to zero, justifying that in this example we did not weight the elastic energy differently inside and outside the body.

9. Conclusions. We have proposed a model for shape averaging, based on elastic deformation of the given shapes to the same configuration. That configuration which yields minimum accumulated hyperelastic deformation energy is defined as average. The method combines a geometric perspective on the shapes $\mathcal{S}_1, \dots, \mathcal{S}_n$ to be averaged with a physical perspective on the objects whose boundary and possibly interior edges constitute the shapes. Thus, not only the intrinsic geometry of the shapes or the Euclidian distances between different components of the shapes matter but also the elastic stress induced by the required deformation onto the averaged shape. It is crucial for the approach that the underlying elastic model is geometrically nonlinear, coping with large deformations, and the material law is nonlinear as well, distinguishing the major deformation effects such as the change of length, area, and volume. The nonlinear hyperelastic functional is invariant with respect to rigid body motions, and isometries are minimizers of the functional. This implies that the proposed method effectively measures the local distance from an isometry in a physically motivated way. It can be combined with the segmentation of image morphology in a joint approach, which is capable of enhancing the low contrast image edges in parallel with the built-in groupwise registration.

The proposed numerical scheme behaves very robustly due to the implicit description of the shapes in

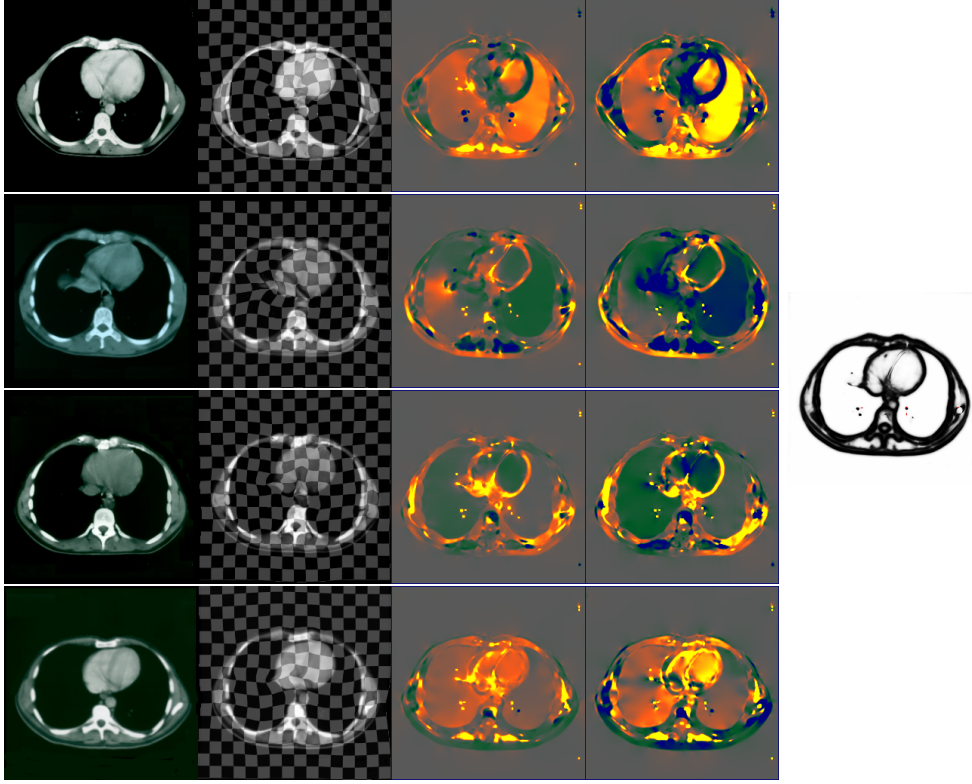


FIG. 8.9. Averaging CT scan slices of the thorax from four different patients. From left to right: Original images, deformations ϕ_i (applied to a checkerboard on which the original image was printed), $|\mathcal{D}\phi_i|_2$ and $\det(\mathcal{D}\phi_i)$ (color-coded as $|\mathcal{D}\phi_i|_2$ with ranges $[0.8\sqrt{2}, 1.2\sqrt{2}]$ and $[0.8, 1.2]$), and average phase field. (Result obtained for resolution 257×257 and parameter values $\gamma = 10^7$, $\mu = 0.1$, $(a_1, a_2, a_3) = (10^6, 0, 10^6)$.)

terms of phase fields and the use of a multi-scale relaxation scheme. As opposed to level set descriptions the phase field approach allows for non-closed shape geometries. Different applications for 2D and 3D shapes and for image morphology demonstrate the performance of the proposed shape averaging. Finally, the considered phase field model is mathematically well-posed and corresponding existence results are discussed.

As possible future research perspectives we see the following. Currently, we confine with a single well phase field model and assume that inside and outside of the objects are a priori given. This approach is in particular computationally beneficial, because the variational problem is quadratic in the phase field v . A Modica–Mortola type phase field functional would allow a more flexible modeling of objects and at the same time a straightforward evaluation of boundary stresses. The latter is particularly interesting if we would like to study cost functionals involving boundary stresses in the definition of the shape average. Furthermore, we plan to investigate a principle component analysis based on the induced set of elastic stress responses on the averaged shape. Finally, the relation of the diffused and the sharp interface models is to the best of our knowledge unclear. Due to the coupling of phase field v and deformations ϕ_i an analytical treatment in the spirit of Γ -convergence seems to be currently out of reach.

Acknowledgments. We are grateful to Werner Bautz, radiology department at the university hospital Erlangen, Germany, for providing CT data of kidneys, as well as to Heiko Schlarb from Adidas, Herzogenaurach, Germany, for providing 3D scans of feet, and to Bruno Wirth, urology department at the Hospital zum Hl. Geist, Kempen, Germany, for providing thorax CT scans. Furthermore, we thank

Stan Osher for pointing us to the issue of elastic shape averaging and Marc Droske and Sergio Conti for discussions about the phase field approach.

REFERENCES

- [1] L. Ambrosio and V. M. Tortorelli. Approximation of functionals depending on jumps by elliptic functionals via Γ -convergence. *Comm. Pure Appl. Math.*, 43:999–1036, 1990.
- [2] L. Ambrosio and V. M. Tortorelli. On the approximation of free discontinuity problems. *Bollettino dell’Unione Matematica Italiana, Sezione B*, 6(7):105–123, 1992.
- [3] T.J. Baker. Three dimensional mesh generation by triangulation of arbitrary point sets. In *Computational Fluid Dynamics Conference, 8th, Honolulu, HI, June 9-11, 1987*, volume 1124-CP, pages 255–271, 1987.
- [4] J. M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Archive of Rational Mechanics and Analysis*, 63:337–403, 1977.
- [5] J.M. Ball. Global invertibility of sobolev functions and the interpenetration of matter. *Proceedings of the royal Society of Edinburgh*, 88A:315–328, 1981.
- [6] K. K. Bhatia, J. V. Hajnal, A. Hammers, and D. Rueckert. Similarity metrics for groupwise non-rigid registration. In N. Ayache, S. Ourselin, and A. Maeder, editors, *Medical Image Computing and Computer-Assisted Intervention, MICCAI 2007*, volume 4792 of *LNCS*, pages 544–552, 2007.
- [7] 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.
- [8] M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. In K. H. Höhne and R. Kikinis, editors, *Visualization in Biomedical Computing: 4th International Conference, VBC*, volume 1131 of *LNCS*, pages 267–276, 1996.
- [9] A. Bronstein, M. Bronstein, and R. Kimmel. *Numerical Geometry of Non-Rigid Shapes*. Monographs in Computer Science. Springer, 2008.
- [10] Martin Burger and Roman Stainko. Phase-field relaxation of topology optimization with local stress constraints. *SIAM Journal on Control and Optimization*, 45(4):1447–1466, 2006.
- [11] Tony F. Chan and Luminita A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.
- [12] Tony F. Chan and Luminita A. Vese. A level set algorithm for minimizing the Mumford-Shah functional in image processing. In *IEEE/Computer Society Proceedings of the 1st IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pages 161–168, 2001.
- [13] 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.
- [14] Guillaume Charpiat, Olivier Faugeras, and Renaud Keriven. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Foundations of Computational Mathematics*, 5(1):1–58, 2005.
- [15] M. Chipot and L. C. Evans. Linearization at infinity and lipschitz estimates in the calculus of variations. *Proceedings of the Royal Society of Edinburgh A*, 102(3–4):291–303, 1986.
- [16] G. Christensen, Richard D Rabbitt, and Michael I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1435–1447, October 1996.
- [17] G. E. Christensen, S. C. Joshi, and M. I. Miller. Volumetric transformations of brain anatomy. *IEEE Trans. Medical Imaging*, 16, no. 6:864–877, 1997.
- [18] G. E. Christensen, R. D. Rabbitt, and M. I. Miller. A deformable neuroanatomy textbook based on viscous fluid mechanics. In J. Prince and T. Runolfsson, editors, *Proc. 27th Annual Conf. Information Sci. and Systems*, pages 211–216, 1993.
- [19] G. E. Christensen, R. D. Rabbitt, and M. I. Miller. 3D brain mapping using a deformable neuroanatomy. *Phys. Med. Biol.*, 39(3):609–618, 1994.
- [20] P. G. Ciarlet. *Three-dimensional elasticity*. Elsevier Science Publishers B. V., 1988.
- [21] U. Clarenz, M. Droske, and M. Rumpf. Towards fast non-rigid registration. In *Inverse Problems, Image Analysis and Medical Imaging, AMS Special Session Interaction of Inverse Problems and Image Analysis*, volume 313, pages 67–84. AMS, 2002.
- [22] 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.
- [23] B. Dacorogna. *Direct methods in the calculus of variations*. Springer-Verlag, New York, 1989.
- [24] G. Dal Maso, J.M. Morel, and S. Solimini. A variational method in image segmentation: existence and approximation results. *Acta Math.*, 168(1-2):89–151, 1992.
- [25] E. De Giorgi, M. Carriero, and A. Leaci. Existence theorem for a minimum problem with free discontinuity set. *Arch. Rat. Mech. and Anal.*, 108:195–218, 1989.
- [26] M. Droske and M. Rumpf. A variational approach to non-rigid morphological registration. *SIAM Appl. Math.*, 64(2):668–687, 2004.

- [27] M. Droske and M. Rumpf. Multi scale joint segmentation and registration of image morphology. *IEEE Transaction on Pattern Recognition and Machine Intelligence*, 29(12):2181–2194, 2007.
- [28] O. Faugeras and G. Hermosillo. Well-posedness of two nonrigid multimodal image registration methods. *SIAM J. Appl. Math.*, 64:1550–1587, 2004.
- [29] P. Thomas Fletcher, Conglin Lu, and Sarang Joshi. Statistics of shape via principal geodesic analysis on lie groups. In *Computer Vision and Pattern Recognition, 2003 (CVPR 2003)*, 2003.
- [30] Tom Fletcher, Suresh Venkatasubramanian, and Sarang Joshi. Robust statistics on riemannian manifolds via the geometric median. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.
- [31] M. Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. H. Poincaré*, 10:215–310, 1948.
- [32] M. Fuchs, B. Jüttler, O. Scherzer, and H. Yang. Shape metrics based on elastic deformations. Technical report, FWF, Joint Research Program of Industrial Geometry, 2008.
- [33] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order*. Grundlehren der Mathematischen Wissenschaften. 224. Berlin-Heidelberg-New York: Springer- Verlag, 1992.
- [34] S. Henn and K. Witsch. A multigrid approach for minimizing a nonlinear functional for digital image matching. *Computing*, 64(4):339–348, 2000.
- [35] Byung-Woo Hong, Stefano Soatto, and Luminita Vese. Enforcing local context into shape statistics. *Advances in Computational Mathematics*, online first, 2008.
- [36] S. Joshi, B. Davis, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, 23:S151–S160, 2004. Supplement 1.
- [37] T. Kapur, L. Yezzi, and L. Zöllei. A variational framework for joint segmentation and registration. *IEEE CVPR - MMBIA*, pages 44–51, 2001.
- [38] H. Karcher. Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, 30(5):509–541, 1977.
- [39] S. L. Keeling and W. Ring. Medical image registration and interpolation by optical flow with maximal rigidity. *Journal of Mathematical Imaging and Vision*, 23 (1):47–65, 2005. to appear.
- [40] D. G. Kendall. Shape manifolds, procrustean metrics, and complex projective spaces. *Bull. London Math. Soc.*, 16:81–121, 1984.
- [41] J. E. Marsden and T. J. R. Hughes. *Mathematical foundations of Elasticity*. Prentice–Hall, Englewood Cliffs, 1983.
- [42] S. Marsland, C. J. Twining, and C. J. Taylor. Groupwise non–rigid registration using polyharmonic clamped–plate splines. In R. E. Ellis and T. M. Peters, editors, *Medical Image Computing and Computer–Assisted Intervention, MICCAI*, volume 2879 of *LNCS*, pages 771–779, 2003.
- [43] F. Mémoli and G. Sapiro. A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics*, Volume 5:313–347, 2005.
- [44] Facundo Mémoli. Gromov-hausdorff distances in euclidean spaces. In *Workshop on Non-Rigid Shape Analysis and Deformable Image Alignment (CVPR workshop, NORDIA’08)*, 2008.
- [45] M. I. Miller, G. E. Christensen, Y. Amit, and U. Grenander. Mathematical textbook of deformable neuroanatomies. *Proc. Natl. Acad. Sci. USA*, 90(24):11944–11948, 1993.
- [46] 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.
- [47] M.I. Miller, A. Trounev, and L. Younes. On the metrics and euler-lagrange equations of computational anatomy. *Annual Review of Biomedical Engineering*, 4:375–405, 2002.
- [48] Luciano Modica and Stefano Mortola. Un esempio di Γ^- -convergenza. *Boll. Un. Mat. Ital. B (5)*, 14(1):285–299, 1977.
- [49] J.-M. Morel and S. Solimini. Segmentation of images by variational methods: a constructive approach. *Revista Matematica de la Universidad Complutense de Madrid*, 1(1):169–182, 1988.
- [50] J.M. Morel and S. Solimini. *Variational models in image segmentation*. Birkäuser, 1994.
- [51] D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Communications on Pure Applied Mathematics*, 42:577–685, 1989.
- [52] R. W. Ogden. *Non-Linear Elastic Deformations*. John Wiley, 1984.
- [53] X. Pennec. Left-invariant riemannian elasticity: a distance on shape diffeomorphisms? In *Mathematical Foundations of Computational Anatomy - MFCA 2006*, pages 1–14, 2006.
- [54] X. Pennec, R. Stefanescu, V. Arsigny, P. Fillard, and N. Ayache. Riemannian elasticity: A statistical regularization framework for non-linear registration. In *Medical Image Computing and Computer-Assisted Intervention ? MICCAI 2005*, LNCS, pages 943–950, 2005.
- [55] R. D. Rabbitt, J. A. Weiss, G. E. Christensen, and M. I. Miller. Mapping of hyperelastic deformable templates using the finite element method. In *Proc. of SPIE*, volume 2573, pages 252–265, 1995.
- [56] P. Rogelj and S. Kovačič. Symmetric image registration. *Medical Image Analysis*, 10(3):484–493, 2006.
- [57] D. Rueckert, A. F. Frangi, and J. A. Schnabel. Automatic construction of 3D statistical deformation models using nonrigid registration. In W. Niessen and M. Viergever, editors, *Medical Image Computing and Computer–Assisted Intervention, MICCAI*, volume 2208 of *LNCS*, pages 77–84, 2001.
- [58] J. Sokolowski and J-P. Zolésio. *Introduction to shape optimization*. Springer-Verlag, Berlin, 1992. Shape sensitivity

- analysis.
- [59] C. O. S. Sorzano, P. Thévenaz, and M. Unser. Elastic registration of biological images using vector-spline regularization. *IEEE Transactions on Biomedical Engineering*, 52(4):652–663, 2005.
 - [60] C. Studholme. Simultaneous population based image alignment for template free spatial normalisation of brain anatomy. In J. C. Gee, J. B. A. Maintz, and M. W. Vannier, editors, *Second International Workshop, WBIR, Biomedical Image Registration*, volume 2717 of *LNCS*, pages 81–90, 2003.
 - [61] G. Sundaramoorthi, A. Yezzi, and A. Mennucci. Sobolev active contours. *International Journal of Computer Vision.*, Volume 73, Number 3:345–366, 2007.
 - [62] V. Šverák. Regularity properties of deformations with finite energy. *Arch. Rat. Mech. Anal.*, 100:105–127, 1988.
 - [63] P. M. Thompson and A. W. Toga. *Handbook of Medical Image Processing*, chapter “Elastic Image Registration and Pathology Detection”. Academic Press, 2000.
 - [64] A. W. Toga and P. M. Thompson. *Handbook of Medical Image Processing*, chapter “Brain Atlases and Registration”. Academic Press, 2000.
 - [65] Gozde Unal, Greg Slabaugh, Anthony Yezzi, and Jason Tyan. Joint segmentation and non-rigid registration without shape priors. Scr-04-tr-7495, Georgia Institute of Technology, Atlanta, 2004.
 - [66] B.C. Vemuri, J. Ye, Y. Chen, and C.M. Leonard. Image registration via level-set motion: Applications to atlas-based segmentation. *Medical Image Analysis*, 7:1–20, 2003.
 - [67] W. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1):35–51, 1996.
 - [68] P. P. Wyatt and J. A. Noble. MAP MRF joint segmentation and registration. In T. Dohi and R. Kikinis, editors, *MICCAI*, volume 2488 of *LNCS*, pages 580–587, 2002.
 - [69] A. Yezzi, L. Zöllei, and T. Kapur. A variational framework for integrating segmentation and registration through active contours. *Medical Image Analysis*, 7(2):171–185, 2003.
 - [70] Laurent Younes. Computable elastic distances between shapes. *SIAM J. Appl. Math.*, 58:565–586, 1998.
 - [71] Yuan-Nan Young and Doron Levy. Registration-based morphing of active contours for segmentation of ct scans. *Mathematical biosciences and Engineering*, Volume 2, Number 1,:79–96, 2005.