Mumford–Shah Model for One-to-one Edge Matching

Jingfeng Han, Benjamin Berkels, Marc Droske, Joachim Hornegger,
Martin Rumpf, Carlo Schaller, Jasmin Scorzin and Horst Urbach

***Abstract*—This paper presents a new algorithm based on the Mumford–Shah model for simultaneously detecting the edge features of two images and jointly estimating a consistent set of transformations to match them. Compared the current asymmetric methods in the literature, this fully symmetric method allows one to determine one-to-one correspondences between the edge features of two images. The entire variational model is realized in a multi-scale framework of the Finite Element approximation. The optimization process is guided by an EM type algorithm and an adaptive generalized gradient flow to guarantee a fast and smooth relaxation. The algorithm is tested on T1 and T2 magnetic resonance image (MRI) data to study the parameter setting. We also present promising results of four applications of the proposed algorithm: inter-object mono-modal registration, retinal image registration, matching digital photographs of neurosurgery with its volume data and motion estimation for frame interpolation.**

***Index Terms*—Image registration, edge detection, Mumford–Shah model**

I. Introduction

IN 1989 the general Mumford–Shah model [1] was first proposed in the literature. In this model, an image is approximated by a cartoon (u, K): u is a piecewise smooth image with sharp edges and K is the discontinuity set in the image domain. This model has been extensively studied for segmentation, image denoising and shape modelling, see i.e. [2], [3], [4], [5] and the references therein.

In 2005 Droske et. al [6], [7] expanded the Mumford–Shah model with the capability of matching the edge features of two images. The edge features are represented by two different cartoon approximations of the images. A smooth dense warping function defines the mapping between the edge features. The modified Mumford–Shah model seeks to simultaneously tackle two highly interdependent tasks: edge segmentation and non-rigid registration. An important benefit of such a joint model is that the intermediate results of one task serve as prior knowledge to the solution of the other task. This advantage has already been pointed out by Yezzi, Zöllei and Kapur [8], who simultaneously segmented edges in different images based on affine matching deformations and an active edge model for the segmentation of implicit curves and surfaces in images, similar to the one proposed by Vese and Chan [9].

A major drawback of the above Mumford Shah based matching is its asymmetry with respect to edge features and the spatial mapping between them. The scheme of the model is shown in Fig. 1. The definition of the similarity measure is not symmetrical: a joint discontinuity set K is used to estimate the edges of the restored template image T and the deformed

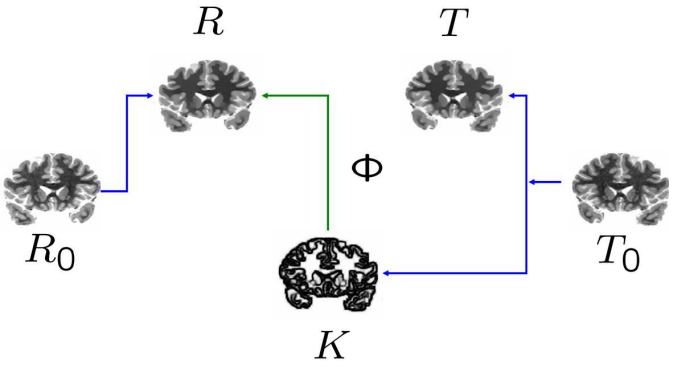

Fig. 1. Non-symmetric Mumford–Shah model for edge matching. R_0 and T_0 are the given reference and template images. R and T are the restored, piecewise smooth functions of image R_0 and image T_0. K is the combined discontinuity set of both images. Function Φ represents the spatial transformation from image T to image R.

edges of the restored reference image R. The model of the spatial mapping between the two images is not symmetrical: the transformation Φ in Fig. 1 is only defined in one direction, from the image T_0 to the image R_0. The asymmetry of the similarity measure and the single directional transformation, as pointed out in [10], cannot ensure that the method is consistent, i.e., if one uses it to compute the transformation Φ from T_0 to R_0 and then switches the roles of T_0 and R_0 to compute the transformation Ψ from R_0 to T_0, it is uncertain whether these transformations are inverse to each other.

In this paper we propose a new symmetric model for edge matching again based on the Mumford–Shah model. Fig. 2 shows the scheme of this symmetric model. We use two relatively separated discontinuity sets (K_R and K_T in Fig. 2) to explicitly represent the edge sets of the associated images. For the ambiguity problem of the correspondence, we apply the idea of consistent registration [11], [12] to simultaneously estimate the forward and reverse transformations and to constrain one transformation to be the inverse of the other one. In this way, the edge sets K_R and K_T of the images R and T, respectively, have equal influence on the edge registration. Thus, the proposed method is one-to-one in the sense, that it allows to determine one-to-one correspondences between the edge features of two images.

Symmetric one-to-one edge matching is not only more sound in the mathematical sense, but also very significant in a broad range of applications, where one is interested in determining the correspondence of the same structure in

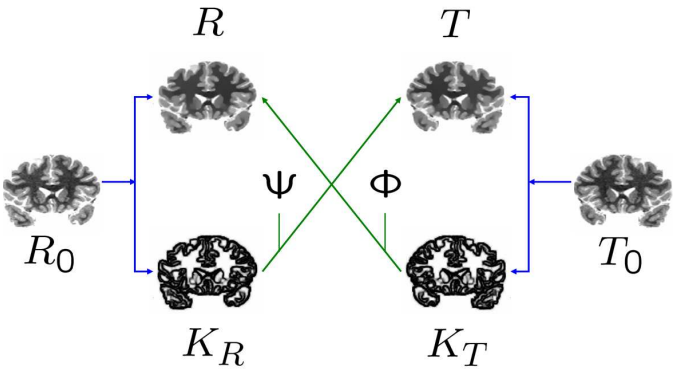

Fig. 2. Symmetric Mumford–Shah model for one-to-one edge matching. R_0 and T_0 are the given images. R and T are the restored, piecewise smooth functions of image R_0 and image T_0. K_R and K_T are the discontinuity sets of the images R and T, respectively. Function Φ represents the transformation from image T to image R and function Ψ represents the transformation from image R to image T.

different images (e.g., non-rigid registration for atlas construction [13], [14], historical biological images [15], [16] or motion estimation).

The paper is organized as follows: In Section II, we introduce some basic knowledge about the classic Mumford–Shah model, the approximation proposed by Ambrosio and Tortorelli and the Finite Element approximation as a preparation for the discussion of the proposed method. Afterwards, in Section III we present the method of one-to-one edge matching, including the functional definitions, variational formulations, numerical implementations and algorithm. In Section IV, we study the parameter setting of the algorithm and show experimental results for several applications. Finally, we draw conclusions in Section V. We note that a preliminary version of part of the work reported in this article has appeared in our conference paper [17].

II. Background

A. Mumford–Shah Model

For a function $u_0 : \Omega \to \mathbb{R}$ on an image domain $\Omega \subset \mathbb{R}^d$ with $d = 2$ or 3 and nonnegative constants α, β and ν, the Mumford–Shah (MS) functional is given by

$$\begin{aligned} E_{\text{MS}}(u, K) =& \frac{\alpha}{2} \int_\Omega (u - u_0)^2 \\ &+ \frac{\beta}{2} \int_{\Omega \setminus K} |\nabla u|^2 \, \mathrm{d}x + \frac{\nu}{2} \mathcal{H}^{d-1}(K). \end{aligned} \tag{1}$$

The first term measures the degree of fidelity of the approximation u with respect to the input data u_0. The second term acts as a kind of "edge-preserving smoother", which penalizes large gradients of u in the homogeneous regions while not smoothing the image in the edge set. The last term $\mathcal{H}^{d-1}$ denotes the $d-1$ dimensional Hausdorff measure, which is used to control the length of the edge set.

B. Ambrosio–Tortorelli Approximation

It is difficult to minimize the original Mumford–Shah functional (1) because of its implicit definition of the discontinuity set K. Various approximations have been proposed during the last two decades. In this article we focus on the Ambrosio–Tortorelli approximation with elliptic functionals [18].

In the Ambrosio–Tortorelli (AT) approximation the discontinuity set K is expressed by a phase-field function v. This scalar function v approximates the characteristic function of the complement of K, $(1 - \chi_K)$, i.e., $v(x) \approx 0$ if $x \in K$ and $v(x) \approx 1$ otherwise. The entire approximation functional is defined as follows:

$$\begin{aligned} E^\epsilon_{\text{AT}}[u, v] =& \frac{\alpha}{2} \int_\Omega (u - u_0)^2 \, \mathrm{d}x + \frac{\beta}{2} \int_\Omega v^2 \|\nabla u\|^2 \, \mathrm{d}x \\ &+ \frac{\nu}{2} \int_\Omega (\epsilon \|\nabla v\|^2 \ + \frac{1}{4\epsilon}(v-1)^2) \, \mathrm{d}x, \end{aligned} \tag{2}$$

The second term, still working as an "edge-preserving smoother", couples zero-regions of v with regions where the gradient of u is large. The following "coupling" between u and v is energetically preferable:

$$v(x) \begin{cases} \approx 0 & \text{where } \|\nabla u\| \gg 0, \\ \approx 1 & \text{where } \|\nabla u\| \approx 0. \end{cases} \tag{3}$$

The last term approximates the edge length, i. e. the $d-1$ dimensional measure $\mathcal{H}^{d-1}(K)$ of the edge set K. The parameter ϵ controls the "width" of the diffusive edge set. Mathematically speaking, the sequence of functionals E^ϵ_{AT} $\Gamma-$converges to the Mumford–Shah functional, i.e.

$$\Gamma - \lim_{\epsilon \to 0} E^\epsilon_{\text{AT}} = E_{\text{MS}}.$$

For a rigorous proof and further explanation we refer to [19].

C. Finite element method

Finite Element (FE) methods are used in this work to discretize the equations. The whole image domain Ω is covered by an uniform rectangular mesh $\mathcal{C}$, on which a standard multi-linear Lagrange finite element space is defined. We consider all images as sets of voxels, where each voxel corresponds to a grid node of $\mathcal{C}$. Let $\mathcal{N} = \{x_1, ..., x_n\}$ denote the nodes of $\mathcal{C}$. The FE basis function of node x_i is defined as the piecewise multi-linear function that fulfills:

$$\varphi_i(x_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

The FE-space $\mathcal{V}$ is the linear hull of the φ_i, i.e.

$$\mathcal{V} := \text{span}(\varphi_1, ..., \varphi_n).$$

The FE-space of vector valued functions is $\mathcal{V}^d$, the canonical basis of this space, is

$$\varphi_1 e_1, ..., \varphi_n e_1, ..., \varphi_1 e_d, ..., \varphi_n e_d,$$

where e_i is the i-th canonical basis vector of $\mathbb{R}^d$. In the FE-space scalar and vector valued functions, e.g. U and Φ, are approximated as follows:

$$U \approx \mathbf{U} := \sum_{i=1}^n \mathbf{U}(x_i)\varphi_i(x) \quad \text{and}$$

$$\Phi = \begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_d \end{bmatrix} \approx \mathbf{\Phi} := \begin{bmatrix} \sum_{i=1}^n \mathbf{\Phi}_1(x_i)\varphi_i(x) \\ \vdots \\ \sum_{i=1}^n \mathbf{\Phi}_d(x_i)\varphi_i(x) \end{bmatrix}.$$

The FE approximation of a function can also be represented by the vector of the function values on the nodes, e.g. $\overline{U} := (\mathbf{U}(x_i), \cdots, \mathbf{U}(x_n))^T$ and $\overline{\Phi} := (\overline{\Phi_1}, \cdots, \overline{\Phi_d})^T$ where $\overline{\Phi_j} = (\mathbf{\Phi}_j(x_1), \cdots, \mathbf{\Phi}_j(x_n))^T$. In this paper we denote continuous functions by uppercase letters (e.g. U or Φ), their FE representation by boldface uppercase letters (e.g. $\mathbf{U}$ or $\mathbf{\Phi}$) and their vector representation by "over-lined" uppercase letters (e.g. $\overline{U}$ or $\overline{\Phi}$).

III. One-to-one edge Matching

A. Problem Statement

The major task of image registration is stated as follows: Find an appropriate transformation Φ such that the transformed template image $T_0 \circ \Phi$ becomes similar to the reference image R_0 [20]. The degree of similarity (or dissimilarity) is evaluated using the gray values R_0 and T_0 or certain features such as edges. We consider a edge based matching method that seeks to register two images based on a joint edge extraction and registration. Thus, the algorithm simultaneously has to fulfill the two following tasks:

- Detection of the edge features from two noisy images.
- Registration of two images using these detected edge features.

The first task is more related to image denoising and edge detection, for which we simply employ the Mumford–Shah model as the feature representation. In practice, the discontinuity sets are approximated by phase-field functions as in the Ambrosio–Tortorelli approximation. Thus, four unknowns $\{R, T, V, W\}$ are estimated, where (R, V) and (T, W) are the feature representations of R_0 and T_0, respectively.

The second task is more related to image registration. The non-rigid transformation from image R_0 to image T_0 is frequently different from the inverse function of the transformation from T_0 to R_0. In order to overcome such correspondence ambiguities, we follow the method of consistent registration [11] to jointly estimate the transformations in both forward and reverse directions. We denote the transformation from T_0 to R_0 as Φ and the transformation from R_0 to T_0 as Ψ. Functions Φ and Ψ are estimated to match the two feature representations (R, V) and (T, W) to each other. Additionally, Φ and Ψ are required to be smooth and approximately inverse to each other. For the desired spatial properties, a regularization functional and a consistency functional are used to constrain the transformations to satisfy these requirements.

B. Functional Definitions

The six unknowns - the restored reference image R, the restored template image T, the edge describing phase-fields V and W of the reference and the template image, respectively, and the deformations Φ and Ψ from the template to the reference domain and vice versa - are estimated by minimizing a joint functional with the following structure:

$$E_{\text{SYM}} = E_{\text{AC}} + \mu E_{\text{CC}} + \lambda E_{\text{REG}} + \kappa E_{\text{CON}}, \tag{4}$$

where μ, λ and κ are nonnegative constants which control the contributions of the associated functionals. The detailed definitions of these functionals are as follows:

1) Auto-coupling Functional:

$$\begin{aligned} E_{\text{AC}} =& C_{\text{AC}}[R, V] + C_{\text{AC}}[T, W] \\ :=& E^{\epsilon}_{\text{AT}}[R, V] + E^{\epsilon}_{\text{AT}}[T, W]. \end{aligned} \tag{5}$$

Here E^{ϵ}_{AT} denotes the functional of the Ambrosio–Tortorelli approximation whose definition has been given in equation (2), where u_0 is replaced by R_0 or T_0 respectively. The single auto-coupling cost function, e.g., $C_{\text{AC}}[R, V]$, essentially makes use of the mechanisms of the Mumford–Shah model and its Ambrosio–Tortorelli approximation to estimate the feature representation (R, V) of the image R_0, such that the piecewise smooth function R optimally couples with the phase-field function V in a manner similar to equation (3). Roughly speaking, this auto-coupling functional is responsible for detecting the edge features of each image and for defining the internal relation between the phase-field function V (or W respectively) and the piecewise smooth function R (or T respectively). In this functional the segmented edge features of the two images, i.e. (R, V) and (T, W), are totally independent from each other.

2) Cross-coupling Functional:

$$\begin{aligned} E_{\text{CC}} =& C_{\text{CC}}[R, W, \Phi] + C_{\text{CC}}[T, V, \Psi] \\ :=& \frac{1}{2} \int_{\Omega} (W \circ \Phi)^2 \left\| \nabla R \right\|^2 \, \mathrm{d}x \\ &+ \frac{1}{2} \int_{\Omega} (V \circ \Psi)^2 \left\| \nabla T \right\|^2 \, \mathrm{d}x. \end{aligned} \tag{6}$$

This functional is responsible for matching the edge features of the two images. It favors spatial transformations Φ and Ψ which optimally couple the feature representations (R, V) and (T, W) in the following way:

$$W \circ \Phi \approx \begin{cases} 0 & \text{where } \|\nabla R\| \gg 0, \\ 1 & \text{where } \|\nabla R\| \approx 0. \end{cases}$$

$$V \circ \Psi \approx \begin{cases} 0 & \text{where } \|\nabla T\| \gg 0, \\ 1 & \text{where } \|\nabla T\| \approx 0. \end{cases}$$

By definition, this functional jointly treats segmentation and registration: For the registration, the functional can act as the similarity measure based on the intermediately segmented edge features. Instead of directly matching the phase-fields functions ($V \leftrightarrow W$) and the smooth functions ($R \leftrightarrow T$), the functional seeks to match the gradient field of the smooth function of one image to the phase-field function of the other image ($V \leftrightarrow \nabla T, W \leftrightarrow \nabla R$). For the segmentation, this functional also imposes the influence of the edge features segmented in the other image. In the following subsection we will see that both spatial transformations are controlled by regularization. The regularized spatial transformations lead to local edge feature correspondence.

3) Regularization Functional:

$$\begin{aligned} E_{\text{REG}} =& C_{\text{REG}}[\Phi] + C_{\text{REG}}[\Psi] \\ :=& \frac{1}{2} \int_{\Omega} \left\| D(\Phi - \mathbb{1}) \right\|^2 \, \mathrm{d}x \\ &+ \frac{1}{2} \int_{\Omega} \left\| D(\Psi - \mathbb{1}) \right\|^2 \, \mathrm{d}x. \end{aligned} \tag{7}$$

Here $\mathbb{1} : x \mapsto x$ denotes the identity mapping and $\Phi - \mathbb{1}$, $\Psi - \mathbb{1}$ the displacement fields corresponding to Φ and Ψ.

Generally speaking, the regularization functional is used to rule out singular transformations which may lead to cracks, foldings, or other undesired properties. In this work the regularization constraint is the sum of the norm of the Jacobian of both displacement fields (see [21] for further explanations of regularization based on the Jacobians of transformations).

Other candidates for regularization constraints are linear elastic [22], [23] and viscous fluid [24], [23] regularizations. These two constraints make use of the continous mechanical model to regularize the transformations [25]. Another alternative, which already ensures a homeomorphism property, is the nonlinear elastic regularization which separately cares about length, area and volume deformation and especially penalizes volume shrinkage [26].

4) Consistency Functional:

$$\begin{aligned} E_{\text{CON}} =& C_{\text{CON}}[\Phi,\Psi] + C_{\text{CON}}[\Psi,\Phi] \\ :=& \frac{1}{2}\int_\Omega \|\Phi\circ\Psi(x) - x\|^2 \,\mathrm{d}x \\ &+ \frac{1}{2}\int_\Omega \|\Psi\circ\Phi(x) - x\|^2 \,\mathrm{d}x. \end{aligned} \tag{8}$$

The forward and reverse transformations Φ and Ψ are purely independent of each other in E_{AC} and E_{REG} and are implicitly correlated in E_{CC} via the matching of the two image / phase-fields pairs, i.e. $(R, W\circ\Phi) \leftrightarrow (T, V\circ\Psi)$. The consistency functional E_{CON} in equation (8) explicitly specifies the relationship between forward and reverse transformations: E_{CC} is minimal if and only if $\Phi\circ\Psi(x) = x = \Psi\circ\Phi(x)$, i.e., $\Phi \approx \Psi^{-1}$ and $\Psi \approx \Phi^{-1}$. The transformation in one direction has to be the inverse function of the transformation in the other direction. For the registration, this consistency constraint favors an invertible and bijective correspondence of the segmented edge features.

C. Variational Formulation

We assume that the minimum of the entire energy E_{SYM} is the zero crossing of its variation with respect to all the unknowns $\{R, T, V, W, \Phi, \Psi\}$. The definition of the entire functional E_{SYM}, as well as each individual functional E_{AC}, E_{CC}, E_{REG} and E_{CON}, is symmetric with respect to the two groups of unknowns: $\{R, V, \Phi\}$ and $\{T, W, \Psi\}$. Thus, we restrict ourself to the description of the computation of variations with respect to $\{R, V, \Phi\}$. The variational formulas of the other group can be deduced in a complementary way.

Given an arbitrary scalar test function $\vartheta \in C_0^\infty(\Omega)$, we obtain the variations with respect to R and V:

$$\begin{aligned} \langle \partial_R E_{\text{SYM}}, \vartheta\rangle &= \langle \partial_R E_{\text{AT}}, \vartheta\rangle + \langle \partial_R E_{\text{CC}}, \vartheta\rangle \\ &= \int_\Omega \alpha(R - R_0)\vartheta + \beta V^2 \nabla R\cdot\nabla\vartheta \,\mathrm{d}x \\ &\quad + \int_\Omega \mu(W\circ\Phi)^2 \nabla R\cdot\nabla\vartheta \;\mathrm{d}x, \end{aligned} \tag{9}$$

$$\begin{aligned} \langle \partial_V E_{\text{SYM}}, \vartheta\rangle &= \langle \partial_V E_{\text{AT}}, \vartheta\rangle + \langle \partial_V E_{\text{CC}}, \vartheta\rangle \\ &= \int_\Omega \beta \|\nabla R\|^2 V\vartheta \;+ \frac{\nu}{4\epsilon}(V-1)\vartheta \;\mathrm{d}x \\ &\quad + \int_\Omega \nu\epsilon \nabla V\cdot\nabla\vartheta \,\mathrm{d}x \\ &\quad + \int_\Omega \mu \|\nabla T\circ\Psi^{-1}\|^2 V\vartheta \,|\det D\Psi|^{-1} \;\mathrm{d}x. \end{aligned} \tag{10}$$

Here we have used the transformation rule

$$\begin{aligned} &\int_\Omega \frac{\mu}{2}\|\nabla T\|^2 V^2\circ\Psi \,\mathrm{d}x \\ &\qquad = \int_{\Psi(\Omega)} \frac{\mu}{2}\|\nabla T\circ\Psi^{-1}\|^2 V^2 \,|\det D\Psi|^{-1} \;\mathrm{d}x \end{aligned}$$

and $\Psi(\Omega) = \Omega$. Given an arbitrary vector-valued test function $\zeta \in C_0^\infty(\Omega, \mathbb{R}^d)$, we obtain the variation with respect to Φ:

$$\begin{aligned} \langle \partial_\Phi E_{\text{SYM}}, \zeta\rangle &= \langle \partial_\Phi E_{\text{CC}}, \zeta\rangle + \langle \partial_\Phi E_{\text{REG}}, \zeta\rangle + \langle \partial_\Phi E_{\text{CON}}, \zeta\rangle \\ &= \int_\Omega \mu \|\nabla R\|^2 (W\circ\Phi)\nabla(W\circ\Phi)\cdot\zeta \,\mathrm{d}x \\ &\quad + \int_\Omega \lambda D\Phi : D\zeta \,\mathrm{d}x \\ &\quad + \int_\Omega \kappa([\Phi\circ\Psi](x) - x)\cdot[\zeta\circ\Psi](x) \,\mathrm{d}x \\ &\quad + \int_\Omega \kappa([\Psi\circ\Phi](x) - x) D\Psi(\Phi(x))\cdot\zeta(x) \,\mathrm{d}x. \end{aligned} \tag{11}$$

Due to the high complexity of the minimization problem (four scalar functions and two vector-valued functions) the unknowns are estimated in an Estimation-Minimization like (EM) procedure: Let $f_{i=1,\dots,m}$ denote the unknown functions and $E := E[f_1, ..., f_m]$ denote the functional.

while $f_{i=1,\dots,m}$ has not yet converged **do**
  for $i = 1$ to m **do**
    $f_i = \arg\min_f E[f_1, ..., f_{i-1}, f, f_{i+1}, ..., f_m]$.
  end for
end while

D. Solution of the linear part

First we introduce generalized mass and stiffness matrices, which play the key roles in the discretization of equations (9) and (10) using FE approximation.

Given a function $f(x) : \Omega \mapsto R$, the generalized mass $M[f]$ and stiffness matrices $L[f]$ are defined as follows:

$$M[f] = \left(\int_\Omega f(x)\varphi_i(x)\varphi_j(x)\,\mathrm{d}x\right)_{i,j} \tag{12}$$

$$L[f] = \left(\int_\Omega f(x)\nabla\varphi_i(x)\cdot\nabla\varphi_j(x)\,\mathrm{d}x\right)_{i,j} \tag{13}$$

Both matrices are $n\times n$-dimensional, where n is the number of nodes in the FE space. Both matrices are sparse, i.e. most entries are zero. An entry is non-zero, if and only if $i = j$ or node i and j are adjacent in the mesh. To compute the integrals in these non-zero entries we use a numerical Gaussian quadrature scheme of order three (cf. [27]). Obviously, the

common mass matrix M and stiffness matrix L are just special cases of the generalized ones, i.e. $M := M[1]$ and $L := L[1]$.

The variations in equation (9) and (10) are linear with respect to the unknowns R and V respectively. In each iteration of the EM procedure, the zero-crossings are simply calculated by solving the corresponding linear systems. Replacing the continuous functions R_0 and R with their FE approximations $\mathbf{R}_0(x) = \sum_{i=1}^n \bar{R}_0^i \varphi_i(x)$ and $\mathbf{R}(x) = \sum_{i=1}^n \bar{R}^i \varphi_i(x)$ and considering base functions φ_j of the FE space as test functions, the equation for zero crossings of (9) is equivalent to:

$$\begin{aligned}
&\alpha \sum_{i=1}^n \sum_{j=1}^n R^i \int_\Omega \varphi_i(x)\varphi_j(x)\,\mathrm{d}x \\
&+ \beta \sum_{i=1}^n \sum_{j=1}^n R^i \int_\Omega \mathbf{V}^2(x)\nabla\varphi_i(x)\cdot\nabla\varphi_j(x)\,\mathrm{d}x \\
&+ \mu \sum_{i=1}^n \sum_{j=1}^n R^i \int_\Omega (\mathbf{W}\circ\mathbf{\Phi})^2(x)\nabla\varphi_i(x)\cdot\nabla\varphi_j(x)\,\mathrm{d}x \\
&= \alpha \sum_{i=1}^n \sum_{j=1}^n R_0^i \int_\Omega \varphi_i(x)\varphi_j(x)\,\mathrm{d}x
\end{aligned} \tag{14}$$

Using the notations of generalized mass (12) and stiffness matrices (13), equation (14) can be rewritten as:

$$\left(\alpha M + \beta L\left[\mathbf{V}^2\right] + \mu L\left[(\mathbf{W}\circ\mathbf{\Phi})^2\right]\right)\overline{R} = \alpha M\overline{R_0}. \tag{15}$$

Similarly (10) leads to:

$$\begin{aligned}
\Big(&\mu M\left[\|\nabla\mathbf{T}\circ\mathbf{\Psi}^{-1}\|^2\left|\det D\mathbf{\Psi}\right|^{-1}\right] \\
&+\beta M\left[\|\nabla\mathbf{R}\|^2\right] + \frac{\nu}{4\epsilon}M + \nu\epsilon L\Big)\overline{V} = \frac{\nu}{4\epsilon}M\underline{1}.
\end{aligned} \tag{16}$$

Here $\underline{1}$ denotes the one-vector, i.e. $(1,\cdots,1)^T$. Analogously, we get the linear systems for $\overline{T}$ and $\overline{W}$:

$$\left(\alpha M + \beta L\left[\mathbf{W}^2\right] + \mu L\left[(\mathbf{V}\circ\mathbf{\Phi})^2\right]\right)\overline{U_T} = \alpha M\overline{T_0}. \tag{17}$$

$$\begin{aligned}
\Big(&\mu M\left[\|\nabla\mathbf{R}\circ\mathbf{\Psi}^{-1}\|^2\left|\det D\mathbf{\Psi}\right|^{-1}\right] \\
&+\beta M\left[\|\nabla\mathbf{T}\|^2\right] + \frac{\nu}{4\epsilon}M + \nu\epsilon L\Big)\overline{V_T} = \frac{\nu}{4\epsilon}M\underline{1}.
\end{aligned} \tag{18}$$

The linear systems (15) - (18) are solved with a preconditioned Conjugate-Gradient (CG) method.

E. Solution of the nonlinear part

Equation (11) shows that the variation of the energy is nonlinear with respect to one of the transformations. Thus the unknown transformation cannot be estimated by solving a linear system. Instead we employ a regularized gradient descent method to iteratively find the zero-crossing:

$$\mathbf{\Phi}^{k+1} = \mathbf{\Phi}^k - \tau^k \cdot \mathrm{grad}_{\mathbf{\Phi}}^{g_\sigma} E[\mathbf{\Phi}^k], \tag{19}$$

where $\mathrm{grad}_{\mathbf{\Phi}}^{g_\sigma} E[\mathbf{\Phi}^k]$ is the regularized gradient with respect to the unknown $\mathbf{\Phi}$ and a metric g_σ, and τ^k is the step size.

1) $\mathrm{grad}_{\mathbf{\Phi}}^{g_\sigma} E[\Phi^k]$: This regularized gradient combined with the time discretization is closely related to iterative Tikhonov regularization, which leads to smooth paths from the initial deformations towards the set of minimizers of the matching energy. As metric we choose:

$$g_\sigma(\mathbf{\Phi}_1,\mathbf{\Phi}_2) = (\mathbf{\Phi}_1,\mathbf{\Phi}_2)_{L^2} + \frac{\sigma^2}{2}(D\mathbf{\Phi}_1, D\mathbf{\Phi}_2)_{L^2}.$$

For theoretical details, we refer to [28], [29], [30]. In our implementation, the regularized gradient $\mathrm{grad}_{\mathbf{\Phi}}^{g_\sigma} E[\mathbf{\Phi}^k]$ is computed in two steps:

- Compute the variation
$$\partial_{\mathbf{\Phi}} E[\mathbf{\Phi}^k] = \left\langle \partial_{\mathbf{\Phi}} E_{\mathrm{SYM}}[\mathbf{\Phi}^k], \zeta \right\rangle$$
according to equation (11), where the integrals are computed with a Gaussian quadrature scheme of order three and the test functions are the canonical basis functions of $\mathcal{V}^d$, see Section II-C.
- The representation of the metric in FE-terms is
$$g_\sigma(\mathbf{\Phi}_1,\mathbf{\Phi}_2) = \left(M_{\text{bl.}} + \tfrac{\sigma^2}{2}L_{\text{bl.}}\right)\overline{\Phi_1}\cdot\overline{\Phi_2}$$
which leads to
$$\mathrm{grad}_{\mathbf{\Phi}}^{g_\sigma} E[\mathbf{\Phi}^k] = \left(M_{\text{bl.}} + \tfrac{\sigma^2}{2}L_{\text{bl.}}\right)^{-1}\left(\partial_{\mathbf{\Phi}_i} E[\mathbf{\Phi}^k]\right)_i.$$
Here $M_{\text{bl.}}$ and $L_{\text{bl.}}$ denote $d\times d$ block matrices with the standard mass and stiffness matrices respectively on the diagonal positions, and zero matrices on the off diagonal positions. We use $\sigma = \sqrt{10h}$, where h is the mesh resolution. The solution of the linear system is computed by a single V-cycle of a multigrid solver.

At this point, we see that the principle difference from "classical" gradient descent methods is that the regularized method does not use the primitive variation but a regularized (smoothed) one as descent direction.

2) τ^k: The step size of the gradient flow is determined by the Armijo-rule[31], choosing the largest τ^k such that energy is minimized in a successive reduction rule. The natural way in the EM procedure is to estimate the step size for each transformation individually, i.e. estimating $\tau_{\mathbf{\Phi}}$ for the transformation $\mathbf{\Phi}$ then estimating $\tau_{\mathbf{\Psi}}$ for $\mathbf{\Psi}$. However, if $\tau_{\mathbf{\Phi}}$ and $\tau_{\mathbf{\Psi}}$ are estimated sequentialy in each iteration, the consistency functional in (6) will prevent $\tau_{\mathbf{\Phi}}$ and $\tau_{\mathbf{\Psi}}$ from being large, because large individual step sizes will increase the consistency functional significantly. Consequently, the regularized gradient descent requires a large number of iterations to approach the minimum. In order to solve this problem, we simultaneously estimate both transformations and compute one step size for both of them:

$$\begin{bmatrix} \mathbf{\Phi}^{k+1} \\ \mathbf{\Psi}^{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}^{k} \\ \mathbf{\Psi}^{k} \end{bmatrix} - \tau^k \begin{bmatrix} \mathrm{grad}_{\mathbf{\Phi}}^{g_\sigma} E[\mathbf{\Phi}^k, \mathbf{\Psi}^k] \\ \mathrm{grad}_{\mathbf{\Psi}}^{g_\sigma} E[\mathbf{\Phi}^k, \mathbf{\Psi}^k] \end{bmatrix}. \tag{20}$$

Since $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are updated at the same time, the consistency energy does not penalize a large step size τ^k any more.

Let $\Theta := [\mathbf{\Phi}, \mathbf{\Psi}]^T$, $\mathrm{grad}E[\Theta] := [\mathrm{grad}_{\mathbf{\Phi}}^{g_\sigma} E, \mathrm{grad}_{\mathbf{\Psi}}^{g_\sigma} E]^T$ and $E'[\Theta] := [E'[\mathbf{\Phi}], E'[\mathbf{\Psi}]]^T$. We define the condition for the Successive Reduction Rule (SRR) as:

$$\frac{E_{\mathrm{REG}}[\Theta^k - \tau^k\cdot\mathrm{grad}E[\Theta^k]] - E_{\mathrm{REG}}[\Theta^k]}{\langle E'[\Theta^k], \mathrm{grad}E[\Theta^k]\rangle} > \frac{1}{4}.$$

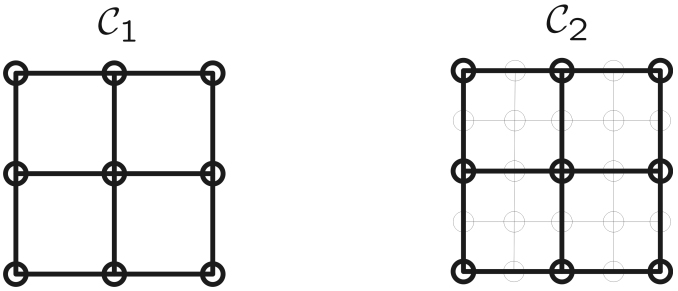

Fig. 3. A simple 2-d example of nested mesh hierarchy. The nodes of the coarse mesh $\mathcal{C}_1$ are a subset of the nodes of the fine mesh $\mathcal{C}_2$. The prolongation of a function from the mesh $\mathcal{C}_1$ to the mesh $\mathcal{C}_2$ only requires the interpolation of the function values on the new nodes.

The step size τ^k is estimated as follows:

% Initialize τ^k from previous iteration.
if $k = 0$ **then** $\tau^k = 1.0$
else $\tau^k = \tau^{k-1}$
% Find the largest τ^k fulfilling SSR.
if SSR succeeds **then**
  do $\tau^k = 2\tau^k$ **until** SSR fails
else
  do $\tau^k = 0.5\tau^k$ **until** SSR succeeds
end if

The regularization of the gradient and the adaptive estimation of the step size allow the regularized gradient descent method to perform more efficiently than the classical ones. In most cases we use five gradient descent steps to estimate the transformations in each iteration of the EM procedure.

F. Multi-scale Algorithm

In order to avoid being trapped in local minima, the algorithm employs a spatial multi-scale scheme, in which global structures are segmented and matched before local ones.

The image domain $\Omega := [0,1]^d$ is discretized by a rectangular mesh $\mathcal{C}_m$, which has $2^m + 1$ equidistant nodes in each axis, thus $n := (2^m + 1)^d$ nodes total. m is called the level of the mesh. A discrete function on the mesh $\mathcal{C}_m$ can also be called a function on level m. Fig. 3 shows a 2-d example of two nested meshes $\mathcal{C}_1$ and $\mathcal{C}_2$, in which the feature representations $\{R, V, T, W\}$ and the transformations $\{\Phi, \Psi\}$ are first computed on the coarse mesh $\mathcal{C}_1$. Then the results are prolongated to the next higher level on the finer mesh $\mathcal{C}_2$.

Although such a nested mesh hierarchy is not natural for finite difference methods, where commonly discrete images with 2^m voxels in each axis are used, it is common for the canonical hierarchy in the Finite Element context. This way the prolongation from one level to the next higher level is very convenient. Let $\mathcal{N}_m$ denote the set of nodes of the m-th mesh, as shown in Fig. 3. The nested mesh hierarchy ensures $\mathcal{N}_{m-1} \subset \mathcal{N}_m$. During prolongation from level $m-1$ to m the function values stay the same on the nodes in $\mathcal{N}_{m-1}$ and the function values on the nodes in $\mathcal{N}_m \setminus \mathcal{N}_{m-1}$ are determined by multi-linear interpolation from the values on the neighboring nodes in $\mathcal{N}_{m-1}$.

The entire multi-scale algorithm is summarized as follows:

given images $\mathbf{R}_0$ and $\mathbf{T}_0$.
given starting level m_0 and ending level m_1.
given number of iterations on each level k_1.
intialize $[\mathbf{R}^{m_0}, \mathbf{V}^{m_0}, \mathbf{T}^{m_0}, \mathbf{W}^{m_0}]$ with 0.
intialize $[\mathbf{\Phi}^{m_0}, \mathbf{\Psi}^{m_0}]$ with $\mathbb{1}$.
for $m = m_0$ **to** m_1 **do**
  for $k = 1$ **to** k_1 **do**
    update $\mathbf{R}^m$ through equation (15)
    update $\mathbf{V}^m$ through equation (16)
    update $\mathbf{T}^m$ through equation (17)
    update $\mathbf{W}^m$ through equation (18)
    update $[\mathbf{\Phi}^m, \mathbf{\Psi}^m]$ with 5 regularized gradient descent steps through equation (20)
  end for
  if $m \neq m_1$ **then**
    intialize $[\mathbf{R}^{m+1}, \mathbf{V}^{m+1}, \mathbf{T}^{m+1}, \mathbf{W}^{m+1}, \mathbf{\Phi}^{m+1}, \mathbf{\Psi}^{m+1}]$ through prolongation from $[\mathbf{R}^m, \mathbf{V}^m, \mathbf{T}^m, \mathbf{W}^m, \mathbf{\Phi}^m, \mathbf{\Psi}^m]$
  end if
end for

IV. Results

Five experiments are performed to demonstrate the one-to-one edge matching algorithm. The first one is designed to study the parameter settings of the algorithm. We have chosen T1- and T2-MRI volumes of the same patient as input data. The second one is designed to show the effect of the algorithm in 3-d inter-object monomodal registrations, whose major task is to build up anatomical correspondence between different individuals. The third experiment shows the application of the algorithm in the registration of retinal images. Then we present results of matching 2-d photographs taken during neurosurgery to the projections of 3-d MRI volume data. Finally we show that the method can be used in the application of frame interpolation. In order to comply with the mesh hierarchy introduced in section III.F, the data sets in experiments A, B, C and D are resampled, while the data set in experiment E is cropped. We have chosen multi-linear interpolation, i.e. bilinear for 2-d data and trilinear for 3-d data, because it conveniently fits into our Finite Element framework and gives acceptable accuracy. But the method does not depend on the way the data is resampled nor on the concrete construction of a multi-scale.

A. Parameter study for 3-d data

Two MRI volumes are acquired from the same individual and with the same machine but with the different scan parameters (T1/T2). The original T1-MRI (our reference image R_0) and T2-MRI (our template image T_0) volumes are already nearly perfectly matched to each other. In order to demonstrate the effect of registration, the T2-MRI volume is artificially deformed by a given elastic transformation. We specified the displacement vectors on eight points and computed the displacement vectors in the remaining part of the data using thin-plate spline interpolation. Both of the given volumes are of size $512 \times 512 \times 101$ and have been resampled to $129 \times 129 \times 129$ pixels to comply with the mesh hierarchy presented before. We performed 18 experiments with different parameter settings. For each experiment 10 EM-iterations were run on the $129 \times 129 \times 129$ mesh. It took approximately two hours for each experiment on a standard PC with Intel Pentium

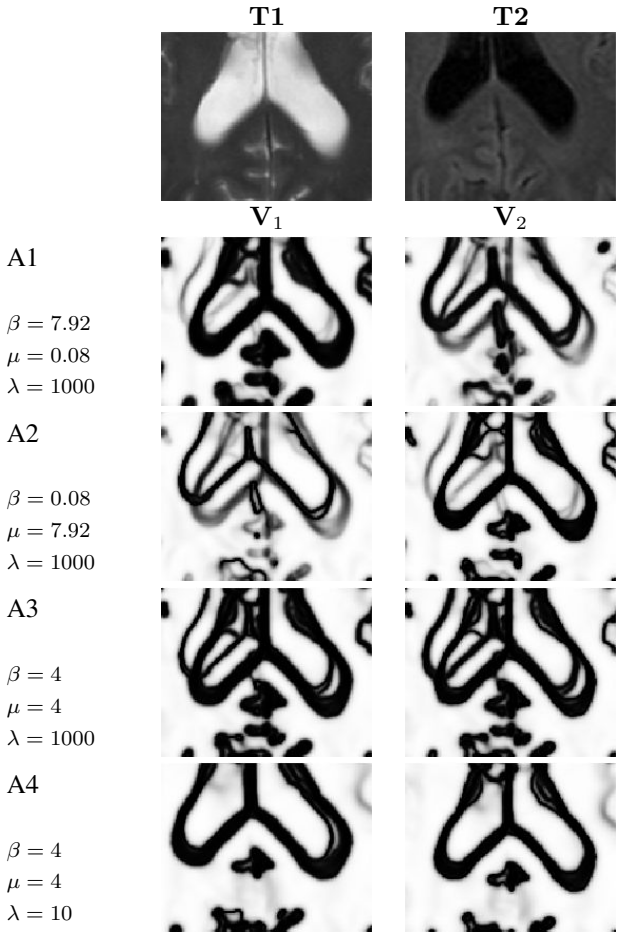

Fig. 4. Experiments A1-A4 show the influence of the parameters β, μ and λ on the phase-field functions. In experiments A1-A3, the overwhelmingly large λ disables the edge matching functionality and allows only edge detections. Furthermore, the ratio between β and μ determines whether the phase-fields represent edge features of its own image or the features of its counterpart. In experiment A4 edge matching as well as edge detection are enabled. Note that edge matching merged the phase-fields of both sides compared to experiment A3.

4 processor 2.26 GHz and 2.0 GB RAM. It is expected that the computational time will decrease significantly by further optimization of the code. Although these parameters are only tested for T1-/T2-MRI edge matching, they can also be used to determine the parameters for edge matching of the other modalities.

Experiments A1-A4 demonstrate how the parameters β, μ and λ balance the edge detection and the edge matching in the algorithm. The other parameters are fixed at $\alpha = 2550$, $\nu = 0.1$, $\kappa = 100$, $\epsilon = 0.01$. In this example, we denote the phase field functions of T1- and T2-MRI volumes as $\mathbf{V}_1$ and $\mathbf{V}_2$ respectively. Fig. 4 shows how the two phase-field functions varied in a local region with different parameters. In the experiments A1-A3, the overwhelmingly large regularization weighting parameter λ $(= 1000)$ prevents the algorithm from matching the edge features of the two images. Without consideration of the edge matching, the detection of edge features is controlled by the ratio between the auto-coupling weighting parameter β and the cross-coupling weighting parameter μ. In experiment A1, since β is much larger than μ, the auto-coupling functional E_{AC} has more influence than the cross-coupling functional E_{CC}. The resulting phase-field functions are more likely to describe its own edge feature. Experiment A2 is exactly the opposite case of A1. With small β and large μ the phase-field function is more likely to represent the edge feature of its counterpart. Namely, $\mathbf{V}_1$ shows the edge features of the image T2 and $\mathbf{V}_2$ shows the edge features of the image T1. The parameters β and μ need to be customized to specific applications. A general principle: β and μ need to be set in such a way that the resulting phase field functions W and V clearly describe the edge features of both images, as shown in experiment A3. For the T1-/T2-MRI data in this experiment, it is reasonable to set β and μ equal. However, when the intensity patterns of images are largely different, like in the neurosurgery photographes and the brain MR projection in section IV.D, it can be necessary to choose the parameters β and μ differently. In experiment A4, we activate the edge matching through a relative small regularization weighting parameter λ $(= 10)$. Each phase-field function describes, not only its own edge features but also, the transformed edge features of the other image. From the figure, one can observe that the phase-field functions are merged with respect to experiment A3.

Experiments B1-B7 and C1-C7 were used to study the setting of the parameters λ and κ. We measured the cross-coupling cost C_{CC}, regularization cost C_{REG} and consistency cost C_{CON} for each experiment. The values of these costs are shown in Table I and II and have been scaled by 10000 for presentation purposes. The minimum and the inverse of the maximum of the determinant of the Jacobians of the forward and reverse transformations are computed to measure the degree of preservation of the topology. If a transformation is regular, these determinants should be close to 1.

Experiments B1-B7 demonstrate the effect of the regularization functional as the weight parameter λ is varied. In experiments B1 and B2, there are minor regularization constraints. A negative Jacobian of the transformations appeared. This means that the estimated transformation failed to preserve the topology of the images. As λ increases, the regularization constraints improve the transformations because the minimum Jacobian and the inverse of the maximum Jacobian are far from being singular. Experiments C1-C7 demonstrate the effect of the consistency functional as the weight parameter κ varied. In the experiment C1, the consistency functional E_{CON} has no influence on the registration. The forward and reverse transformations are relatively independently estimated. The inconsistency of the two transformations are confirmed by the relatively large cost of the consistency functional. As κ increases, the cost of the consistency functional approaches to zero. This means that one transformation is more likely to be the inverse function of the other one. Notice that the cost of cross-coupling functional increases when the consistency constraints and regularization constraints become strong, which indicates a worse matching of edge features between the two images. The optimal parameters should be chosen so as to achieve optimal feature matching, least amount of topological distortion and acceptable inconsistency of the transformations. According to our experience, it is safe to roughly fix five

of the parameters in most 2-d and 3-d applications, i.e. $\lambda = 10, \kappa = 100, \alpha = 2550, \nu = 0.1 \sim 1, \epsilon = 0.01$ usually achieves good results.

B. Volumes of different individuals

In the following two experiments we use the one-to-one edge matching method to solve the inter-object mono-modal registration problem: registering two MR data sets (MR-to-MR) and two CT data sets (CT-to-CT). The two MR data show healthy brains of two individuals. The two CT data show two other patients, one normal and one abnormal. The data sets are collected by the same MR and CT scanners with the same scanning parameters. The MR data sets are preprocessed by segmenting the brain from the head using MRIcro[1].

The original sizes of the two CT data sets were $512 \times 512 \times 58$ and $512 \times 512 \times 61$ while the two MR data sets were $256 \times 256 \times 160$ and $256 \times 256 \times 170$. All of them have been resampled into a $257 \times 257 \times 257$ voxel lattice with the same resolution in all three directions. The experiments were performed with the previously described multi-scale scheme, with 10 iterations for each of the levels: $33 \times 33 \times 33$, $65 \times 65 \times 65$, $129 \times 129 \times 129$ and $257 \times 257 \times 257$. It took approximately 1 minute, 10 minutes, 90 minutes and 5 hours respectively for each level. The parameters of the algorithm were set as follows: $\alpha = 2550$, $\beta = 1$, $\nu = 0.1$, $\mu = 1$, $\lambda = 10$, $\kappa = 100$, $\epsilon = 0.01$.

The matching results of the data sets are visualized by a pattern of "interlace-stripe", showing the two data sets in turns within a single volume. As shown in Fig. 5 and Fig. 6, the titles of each sub-figure keep consistent with the notations in the paper: R and T denote the two original data sets, while $\mathbf{\Phi}$ and $\mathbf{\Psi}$ denote the forward and reverse transformations. The sub-figures $\mathbf{R_0} \| \mathbf{T_0}$ and $\mathbf{T_0} \| \mathbf{R_0}$ show the interlace-stripe volumes of the original data sets R and T, while the sub-figures $\mathbf{R_0} \| \mathbf{T_0} \circ \mathbf{\Phi}$ and $\mathbf{T_0} \| \mathbf{R_0} \circ \mathbf{\Psi}$ show the interlace-stripe volumes of the registered data in forward and reverse directions.

From visual inspection, the algorithm of one-to-one edge matching successfully registers MR-to-MR and CT-to-CT volume data sets of different individuals in both directions. Fig. 5 shows precise alignments of the edges such as the brain's volume shape, hemispheric gap and ventricular system for inter-object MR-to-MR registration. In the inter-object CT-to-CT registration the main interest is to obtain the fitting shape of the bone. In Fig. 6 axial cuts of the 3-d CT data set are shown. Fig. 7 shows that the initial mismatch of the data sets, visible by the discontinued bone edges in the top row, is dissolved with the computed transformation, as is evident from the continuous bone edges in the bottom row.

C. Retinal images

A concurrent representation of the optic nerve header and the neuroretinal rim in various retina image modalities is significant for a definite diagnosis of glaucoma. Several modalities of retinal images have been used in the ophthalmic

[1]http://www.sph.sc.edu/comd/rorden/mricro.html

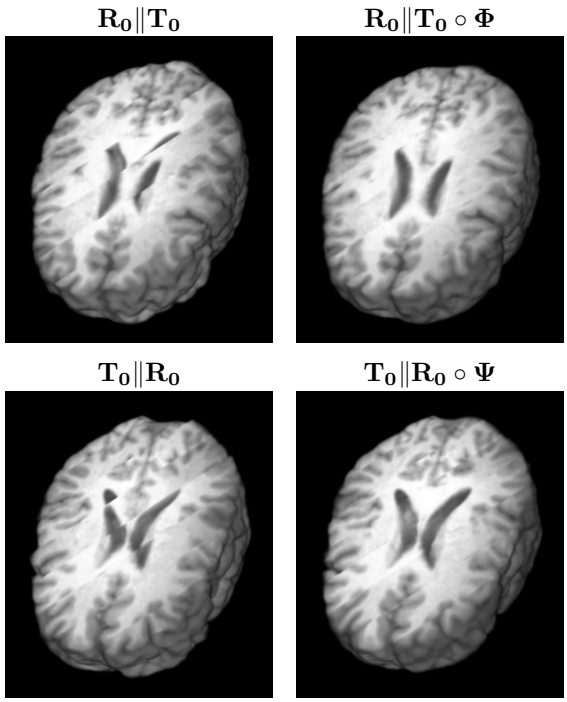

Fig. 5. Inter-object MR-to-MR registration using one-to-one edge matching method. The sub-figure $\mathbf{R_0} \| \mathbf{T_0}$ and the sub-figure $\mathbf{T_0} \| \mathbf{R_0}$ show the interlace-stripe volumes of the original data sets R_0 and T_0, while the sub-figures $\mathbf{R_0} \| \mathbf{T_0} \circ \mathbf{\Phi}$ and $\mathbf{T_0} \| \mathbf{R_0} \circ \mathbf{\Psi}$ show the interlace-striped volumes of registered data sets in forward and reverse directions.

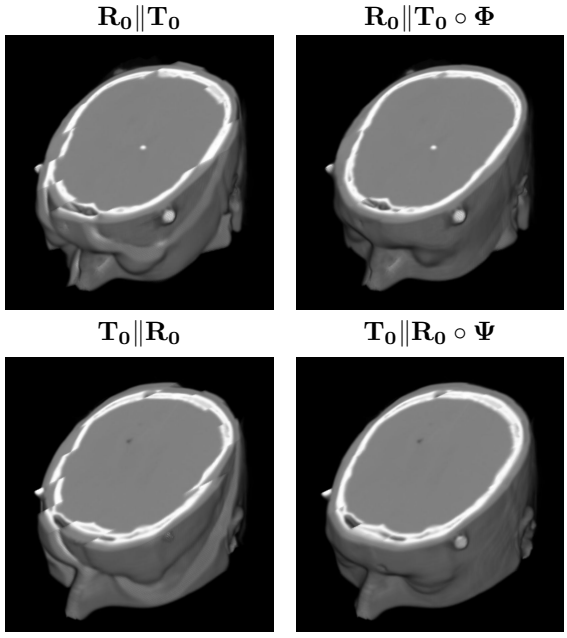

Fig. 6. Inter-object CT-to-CT registration using one-to-one edge matching method. The sub-figure $\mathbf{R_0} \| \mathbf{T_0}$ and the sub-figure $\mathbf{T_0} \| \mathbf{R_0}$ show the interlace-stripe volumes of the original data sets R_0 and T_0, while the sub-figures $\mathbf{R_0} \| \mathbf{T_0} \circ \mathbf{\Phi}$ and $\mathbf{T_0} \| \mathbf{R_0} \circ \mathbf{\Psi}$ show the interlace-striped volumes of registered data sets in forward and reverse directions.

TABLE I
STUDY OF THE WEIGHT OF THE REGULARIZATION FUNCTIONAL E_{REG}

Exp.	λ	C_{CC}		C_{REG}		C_{CON}		$\det D\boldsymbol{\Phi}$		$\det D\boldsymbol{\Psi}$	
		For	Rev	For	Rev	For	Rev	$1/\max$	min	$1/\max$	min
$B1$	0.01	1982	2939	780.9	128.5	3.700	3.667	0.4736	-0.057	0.6933	0.4197
$B2$	0.1	2221	2944	517.3	94.27	2.965	2.940	0.5671	0.088	0.7617	0.5195
$B3$	1	2709	2971	181.2	59.70	2.032	2.029	0.7348	0.4737	0.7358	0.4899
$B4$	5	3120	3050	44.24	27.27	1.149	1.146	0.8738	0.7296	0.8518	0.7464
$B5$	10	3328	3165	20.02	11.00	0.9419	0.9415	0.9253	0.8209	0.9070	0.8706
$B6$	20	3517	3243	6.180	3.031	0.7674	0.7699	0.9403	0.9000	0.9479	0.9301
$B7$	50	3550	3314	1.053	0.5344	0.1792	0.1833	0.9832	0.9599	0.9802	0.9724

C_{CC}: Cross-coupling functional, C_{REG}: Regularization functional, C_{CON}: Consistency functional. The other weight parameters were set as follows: $\alpha = 2550, \beta = 1, \nu = 0.1, \mu = 0.5, \kappa = 100$.

TABLE II
STUDY OF THE WEIGHT OF CONSISTENCY FUNCTIONAL E_{CON}

Exp.	κ	C_{CC}		C_{REG}		C_{CON}		$\det D\boldsymbol{\Phi}$		$\det D\boldsymbol{\Psi}$	
		For	Rev	For	Rev	For	Rev	$1/\max$	min	$1/\max$	min
$C1$	0	3044	3121	28.85	41.87	3.054	3.047	0.8824	0.7507	0.8475	0.7950
$C2$	50	3072	3137	27.19	45.13	0.7922	0.7891	0.8782	0.7251	0.8548	0.8136
$C3$	100	3088	3157	27.24	42.26	0.3255	0.3249	0.8751	0.7495	0.8569	0.8230
$C4$	200	3236	3115	32.69	25.19	0.1720	0.1720	0.8996	0.8032	0.8624	0.8246
$C5$	300	3279	3154	27.72	17.06	0.1430	0.1426	0.9061	0.8046	0.8971	0.8824
$C6$	400	3291	3169	26.82	17.50	0.1118	0.1165	0.9079	0.8086	0.8977	0.8758
$C7$	500	3334	3182	24.74	32.82	0.0803	0.0803	0.9115	0.8170	0.9917	0.8794

E_{CC}: Cross-coupling functional, E_{REG}: Regularization functional, E_{CON}: Consistency functional. The other weight parameters were set as follows: $\alpha = 2550, \beta = 1, \nu = 0.1, \mu = 0.5, \lambda = 10$.

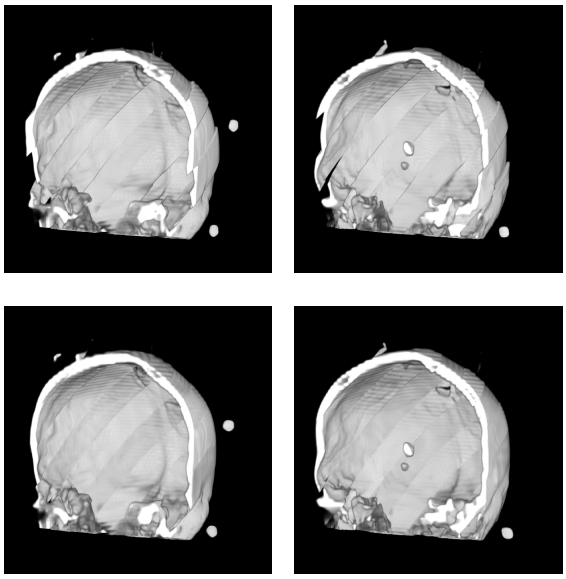

Fig. 7. The matching of skulls in CT-to-CT registration. Above: Interlace-stripe volumes of skulls of original data sets. Bottom: Interlace-stripe volumes of matched skulls.

clinic: the reflection-free photographs with an electronic flash illumination and the depth/reflectance retina images acquired by scanning-laser-tomograph. By acquisition, the depth and reflectance images normally have been perfectly matched to each other. Thus the task of this application is the registration of multi-modal retina images, i.e. to match the reflectance and depth images with the photograph. For the registration of mono-modal retina images we refer to [32], [33]. Fig. 8 shows an example of multi-modal retina images of a same patient. In a recent paper [34], an affine transformation model and an extended mutual information similarity are applied for the registration of bi-modal retina images. However, as shown in Fig. 9 (first column), this method (using the software described in [34]) still cannot recover the minor deviations in the domain of vessels and neuroretinal rims. In this experiment we employ our one-to-one edge matching algorithm as a post-registration to compensate such small deviations of fine vessels.

The images are pre-processed in the following way: extracting the green channel of the photograph as the input for the registration, affinely pre-registering the photograph to reflectance and depth images using the automatic software described in [34], sampling the pre-registered images in a mesh of 257×257. The algorithm is run for 10 iterations in three levels, which takes less than three minutes altogether. The parameters of the algorithm are set as follows: $\alpha = 2550$, $\beta = 1$, $\nu = 0.1$, $\mu = 0.5$, $\lambda = 10$, $\kappa = 100$, $\epsilon = 0.01$. From Fig. 9, one can observe that most minor deviations in the domain of vessels are compensated by the computed non-rigid transformations. Note that in this example with fine elongated structures, different from more volumetric image structures in the other applications, an affine pre-registration is used to compensate the large initial mismatch and to avoid getting stuck in a local minimum.

D. Photographs of neurosurgery

In neocortical epilepsy surgery, the tumor (lesion) may be located adjacent to, or partly within, so-called eloquent

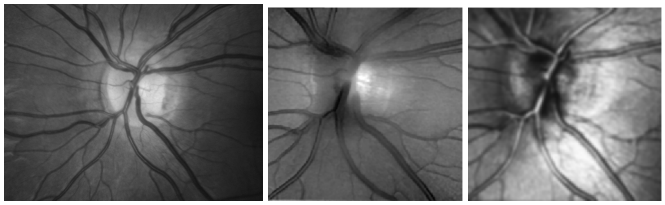

Fig. 8. Multi-modal retina images of a same patient: photograph (left), depth image (middle) and reflectance image (right).

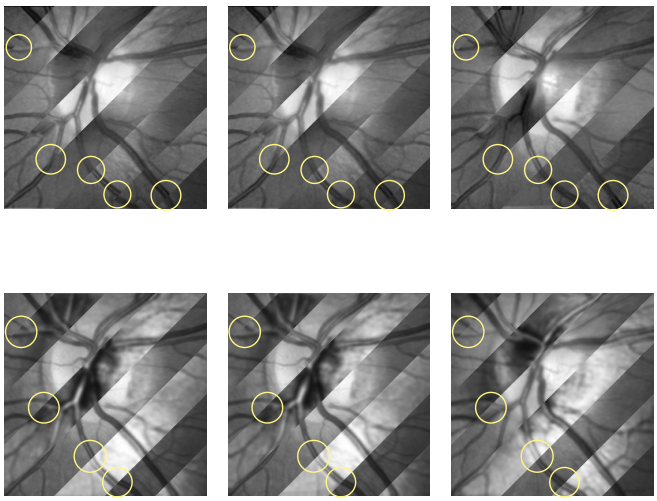

Fig. 9. The example of post-registration of bi-modal retina images using one-to-one edge matching. The photograph is registered with the depth image (top) and the reflectance image (bottom). A published registration method for bi-modal retina images cannot fully recover the minor deviations of fine structures (first column). The forward and reverse transformations estimated by the one-to-one edge matching successfully remove such minor mismatching.

(functionally very relevant) cortical brain regions. For a safe neurosurgical planning, the physician needs to map the appearance of the exposed brain to the underlying functionality. Usually, an electrode is placed on the surface of the brain in the first operation for electrophysiological examination of the underlying brain functionalities, then the photograph within the tested anatomical boundaries is colored according to the function of electrode contacts. On the other hand, the pre-operative 3-d MR data set contains the information of the underlying tumor and healthy tissue as well. In the second operation, the registered photograph and MRI volume are used together to perform the cutting without touching eloquent areas. At the moment a neocortical expert needs to manually rotate the 3-d MR to find the best 2-d projection matching to the photographs. However, due to the different acquisitions and the brain shift during surgery, the photograph and MR projection cannot be accurately aligned. In this experiment, we make use of our one-to-one edge matching algorithm to refine the matching between a 2-d digital photograph of epilepsy surgery to the projection of 3-d MR data of the same patient.

The digital photographs of the exposed cortex are taken with a handheld Agfa e1280 digital camera (Agfa, Cologne, Germany) from the common perspective of the neurosurgeon's view. The high-resolution 3-d data set is acquired according to the T1-weighted MR imaging protocol (TR 20, TE 3.6, flip angle 30°, 150 slices, slice thickness 1mm) using 1.5 Tesla Gyroscan ACS-NT scanner (Philips Medical Systems). The brain is automatically extracted from the MRI volume using the SISCOM module of the Analyze software (Mayo Foundation, Rochester, MN). For both the photograph and the MR projection, the regions of interest are manually selected by a physician.

Fig. 10 shows the input images, preprocessed images, interlace-stripe registered and unregistered images. In sub-figure $\mathbf{R}_0$ the digital photograph shows the exposed left hemisphere from an intraoperative viewpoint, the frontal lobe on the upper left, the parietal lobe on the upper right and parts of the temporal lobe on the bottom. The surface with the gyri and sulci and the overlying vessels are clearly visible. Alongside, sub-figure $\mathbf{T}_0$ displays the left-sided view of the rendered MR volume in the corresponding parts. Comparing sub-figures $\mathbf{R}_0$ and $\mathbf{T}_0$, one can notice that the undesired surface vessels and reflectance flash are strongly presented in the digital photograph, while the MR projection images clearly display the desired edge features. The photographic image and the projection image were preprocessed by appropriate GIMP filter chains for edge enhancement. The preprocessed images are displayed in sub-figures "$\mathbf{R}_0$ preprocessed" and "$\mathbf{T}_0$ preprocessed", respectively. Both images were resampled to 2049×2049 pixels. The algorithm was run from level 3 to level 11. We note that the values of the parameters β and μ are quite different from the other examples. The reason is that the image modalities of the photograph and the MR projection differ largely from each other. The two parameters are set as $\beta = 100$ and $\mu = 0.1$, so that both phase field functions W and V clearly represent the edge features on the brain and have comparable influence on the registration. In sub-figures $\mathbf{T_0}\|\mathbf{R_0}$ and $\mathbf{R_0}\|\mathbf{T_0}$, the interlace-stripe images illustrate the mismatch of photograph and MR projection. Sub-figures $\mathbf{T_0}\|\mathbf{R_0} \circ \mathbf{\Psi}$ and $\mathbf{R_0}\|\mathbf{T_0} \circ \mathbf{\Phi}$ show that the method greatly refines the matching of the desired edge features. Especially the brain sulci and gyri, which are significant for neurosurgery, are nearly perfectly aligned to each other. We have implemented a mutual information algorithm in the same Finite Element framework (including the step sized controlled, regularized, multi-scale descent) for a comparison. Overall, our method gives comparable results in most cases, especially when dealing with coarse structures. However, in this example that contains a large number of fine structures, the edge-based matching gives better alignment. The zoom views of local regions in Fig. 11 show that the edge-matching method can achieve a better alignment of fine structures than the mutual information based registration.

E. Motion estimation for frame interpolation

Temporal interpolation of video frames in order to increase the frame rate requires the estimation of a motion field (transformation). Then pixels in the intermediate frame are interpolated along the path of the motion vector. In this section,

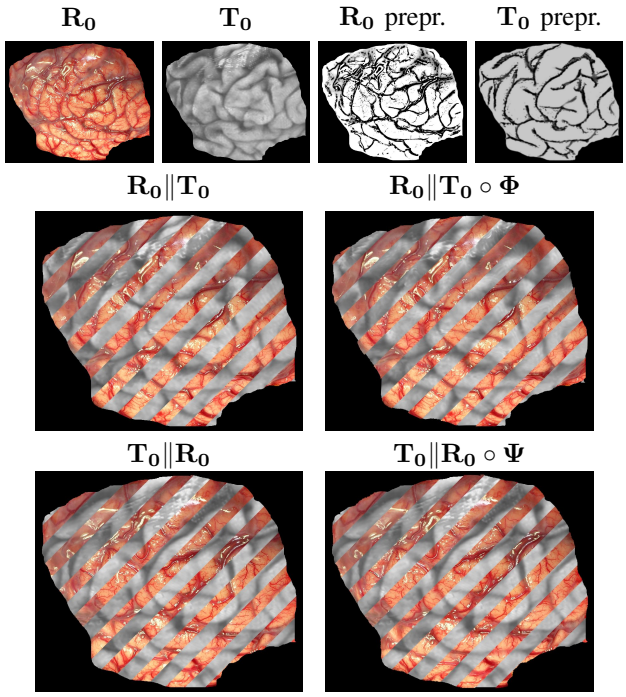

Fig. 10. Experimental results of matching a neurosurgery photograph of a section of the brain with its MR projection. All the sub-figures only display the region of interest: the exposed cortex. $\mathbf{R_0}$: The photograph of the exposed left hemisphere from an intraoperative view point. $\mathbf{T_0}$: The projection of the MR volume, whose orientation is specified by physicians. Preprocessed $\mathbf{R_0}$ and Preprocessed $\mathbf{T_0}$: The preprocessed photograph and MR projection. $\mathbf{R_0}\|\mathbf{T_0}$ and $\mathbf{T_0}\|\mathbf{R_0}$: The interlace-strip images of unregistered photograph and MR projection. $\mathbf{R_0}\|\mathbf{T_0}\circ\mathbf{\Phi}$ and $\mathbf{T_0}\|\mathbf{R_0}\circ\mathbf{\Psi}$: The interlace-strip images of registered photograph and MR projection.

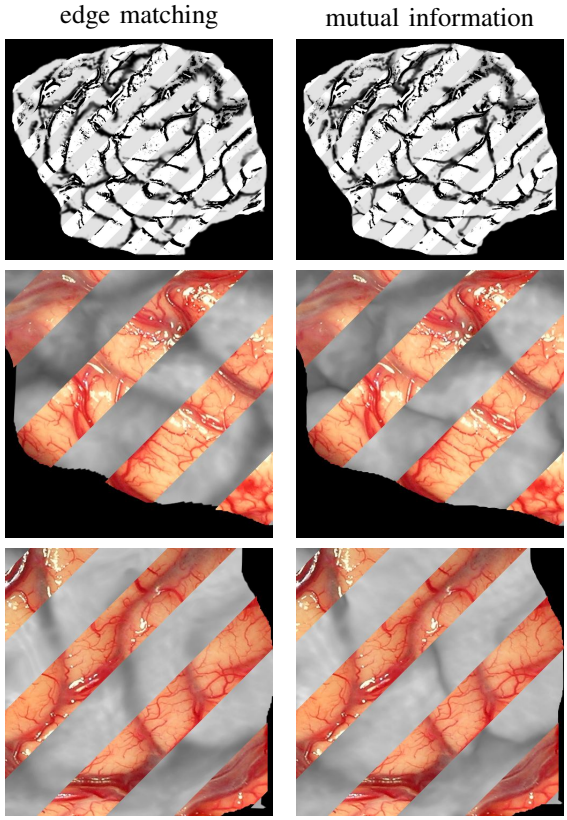

Fig. 11. Comparison of one-to-one edge matching (left) and the mutual information based matching (right). The two algorithms are implemented in a same Finite Element framework including the step size controlled, regularized multi-scale descent. The first row shows how the pre-processed images are registered by the two methods. The last two rows show zoomed views of local regions in the registered images. The comparison shows that one-to-one edge matching comes along with a significantly better registration of fine structures.

we give a proof of concept that the one-to-one edge matching method can be used for this application. For a review of techniques of frame interpolation, we refer to [35], [36].

We perform our test on the Susie sequence and interpolate frame 58 in Fig. 12. We use a 257×257 cropped version for the experiment. Frames 57, 58 and 59 are denoted as F_{57}, F_{58} and F_{59} respectively. The forward transformation $\Phi : F_{57} \rightarrow F_{59}$ and reverse transformation $\Psi : F_{59} \rightarrow F_{57}$ are estimated by the one-to-one edge matching with the parameter setting: $\alpha = 2550$, $\beta = 1$, $\nu = 0.1$, $\mu = 1$, $\lambda = 10$, $\kappa = 100$, $\epsilon = 0.01$. Frame 58 is interpolated as: $F_{58} = 0.5 \times (F_{57} \circ 0.5\Phi + F_{59} \circ 0.5\Psi)$. It is compared with a standard block matching algorithm using an adaptive rood pattern search [37], 16×16 blocks and a search range of $[-16, 16]$ in the horizontal and vertical directions. The experimental results show that the block matching algorithm produces blocky and noisy motion fields, while the one-to-one edge matching based motion estimation gives an excellent visual quality of frame interpolation.

Fig. 12. Motion estimation for frame interpolation. Top: Original frame 57, 58 and 59 of Susie sequence. Bottom: the interpolated frame 58 using simply averaging (left), one-to-one edge matching motion estimation (middle) and standard block matching motion estimation (right).

V. Conclusion and Summary

This paper presents a new algorithm for the edge matching problem. It simultaneously performs the following three tasks: detecting the edge features from two images, computing two dense warping functions in both forward and reverse directions to match the detected features, and constraining each dense warping function to be the inverse of the other.

An adaptive regularized gradient descent, in the framework of multi-resolution Finite Element approximation, enables the algorithm to efficiently find the pair of dense transformations. The algorithm has been tested on T1-/T2-MR volume data. It is found that the proposed algorithm successfully preserved the topology of the images and the bijectivity of the mappings. The paper also shows that the algorithm has been successfully used in four applications: registration of inter-object volume data, registration of retinal images, matching photographs of neurosurgery with its volume data and motion estimation for frame interpolation.

ACKNOWLEDGMENT

The work is supported by Deutsche Forschungsgemeinschaft (DFG) under the grant SFB 603, TP C10 and by the DFG priority program 1114 on time sequence analysis and image processing. The authors acknowledge Rüdiger Bock (Chair of Pattern Recognition, Erlangen-Nürnberg University) for providing the retinal images and the retina registration software for comparison. The authors thank Michael Fried (Chair of Applied Mathematics, Erlangen-Nürnberg University) for his valuable comments and suggestions. The authors thank HipGraphics Inc. for providing the software (InSpace) for volume rendering.

REFERENCES

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

[2] J. M. Morel and S. Solimini, *Variational methods in image segmentation*. Cambridge, MA, USA: Birkhauser Boston Inc., 1995.

[3] T. F. Chan, B. Y. Sandberg, and L. A. Vese, "Active contours without edges for vector–valued images," *Journal of Visual Communication and Image Representation*, vol. 11, pp. 130–141, 2000.

[4] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[5] M. Fried, "Multichannel image segmentation using adaptive finite elements," *Computing and Visualization in Science*, to appear.

[6] M. Droske and W. Ring, "A Mumford-Shah level-set approach for geometric image registration," *SIAM Appl. Math.*, 2007, to appear.

[7] M. Droske, W. Ring, and M. Rumpf, "Mumford-shah based registration," *Computing and Visualization in Science manuscript*, 2007, to appear.

[8] T. Kapur, L. Yezzi, and L. Zöllei, "A variational framework for joint segmentation and registration," *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (IEEE CVPR - MMBIA)*, pp. 44–51, 2001.

[9] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266 – 277, 2001.

[10] P. Rogelj and S. Kovačič, "Symmetric image registration," *Medical Image Analysis*, vol. 10, pp. 484–493, 2006.

[11] G. E. Christensen and H. J. Johnson, "Consistent image registration," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, pp. 568–582, 2001.

[12] H. J. Johnson and G. E. Christensen, "Consistent landmark and intensity-based image registration," *IEEE Transactions on Medical Imaging*, vol. 21, no. 5, pp. 450–461, 2002.

[13] D. Rueckert, A. Frangi, and J. Schnabel, "Automatic construction of 3-d statistical deformation models of the brain using nonrigid registration," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 1014–1025, 2003.

[14] S. Marsland, C. Twining, and C. Taylor, "Groupwise non-rigid registration using polyharmonic clamped-plate splines," in *Sixth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCIA'03)*, ser. LNCS, R. E. Ellis and T. M. Peters, Eds. Montreal: Springer Verlag, 2003, pp. 246–250.

[15] C. Sorzano, P. Thevenaz, and M. Unser, "Elastic registration of biological images using vector-spline regularization," *IEEE Transactions on Biomedical Engineering*, April 2005.

[16] I. Arganda-Carreras, C. Sorzano, R. Marabini, J. Carazo, C. Ortiz-De-Solorzano, and J. Kybic, "Consistent and elastic registration of histological sections using vector-spline regularization," in *workshop of the 9th European Conference on Computer Vision (CVAMIA-06)*, R. E. Ellis and T. M. Peters, Eds. In Press, 2006.

[17] J. Han, B. Berkels, M. Rumpf, J. Hornegger, M. Droske, M. Fried, J. Scorzin, and C. Schaller, " A Variational Framework for Joint Image Registration, Denoising and Edge Detection," in *Proceedings Bildverarbeitung für die Medizin* , H. Handels, J. Ehrhardt, A. Horsch, H. Meinzer, and T. Tolxdorff, Eds. Berlin: Springer, 2006, pp. 246–250.

[18] L. Ambrosio and V. M. Tortorelli, "Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence," *Communications on Pure and Applied Mathematics*, vol. 43, pp. 999–1036, 1990.

[19] R. March, "Visual reconstructions with discontinuities using variational methods," *Image and Vision Computing*, vol. 10, no. 1, pp. 30–38, 1992.

[20] J. Modersitzki, *Numerical Methods for Image Registration*. Oxford: Oxford University Press, 2004.

[21] J. Ashburner, J. Andersson, and K. Friston, "High-dimensional nonlinear image registration using symmetric priors," *NeuroImage*, vol. 9, pp. 619–628, 1999.

[22] C. Broit, "Optimal registration of deformed images," Ph.D. dissertation, University of Pensylvania, 1981.

[23] G. E. Christensen, S. C. Joshi, and M. I. Miller, "Volumetric transformation of brain anatomy," *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, pp. 864–877, 1997.

[24] M. Bro-Nielsen and C. Gramkow, "Fast fluid registration of medical images." *Lecture Notes in Computer Science*, vol. 1131, pp. 267–276, 1996.

[25] M. E. Gurtin, *An Introduction to Continuum Mechanics*. Orlando, Florida: Academic Press, 1981.

[26] M. Droske and M. Rumpf, "A variational approach to non-rigid morphological registration," *SIAM Appl. Math.*, vol. 64, no. 2, pp. 668–687, 2004.

[27] R. Schaback and H. Werner, *Numerische Mathematik*, 4th ed. Berlin: Springer-Verlag, 1992.

[28] O. Scherzer and J. Weickert, "Relations between regularization and diffusion filtering," *Journal of Mathematical Imaging and Vision*, vol. 12, no. 1, pp. 43–63, 2000.

[29] U. Clarenz, S. Henn, and K. Rumpf, M. Witsch, "Relations between optimization and gradient flow methods with applications to image registration," in *Proceedings of the 18th GAMM Seminar Leipzig on Multigrid and Related Methods for Optimisation Problems*, 2002, pp. 11–30.

[30] 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*, vol. 313. AMS, 2002, pp. 67–84.

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

[32] A. Can, C. V. Stewart, B. Roysam, and H. L. Tanenbaum, "A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 3, 2002.

[33] ——, "A feature-based technique for joint, linear estimation of high-order image-to-mosaic transformations: mosaicing the curved human retina," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 3, 2002.

[34] L. Kubecka and J. Jan, "Registration of bimodal retinal images - improving modifications," in *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2004)*, vol. 3, 2004, pp. 1695– 1698.

[35] R. Krishnamurthy, J. Woods, and P. Moulin, "Frame interpolation and bidirectional prediction of video usingcompactly encoded optical-flow fields and label fields," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, pp. 713–726, 1999.

[36] H. A. Karim, M. Bister, and M. U. Siddiqi, "Multiresolution motion estimation for low-rate video frame interpolation," *Journal on Applied Signal Processing*, vol. 11, pp. 1708–1720, 2004.

[37] Y. Nie and K.-K. Ma, "Adaptive rood pattern search for fast block-matching motion estimation," *IEEE Transactions on Image Processing*, vol. 11, pp. 1442 – 1449, 2002.