SCHOOL OF ENGINEERING - STI
ELECTRICAL ENGINEERING INSTITUTE
SIGNAL PROCESSING LABORATORIES

Dominique ZOSSO

EPFL - STI - IEL - LTS5
Station 11
CH-1015 LAUSANNE

Phone: +41 21 693 46 22
Fax: +41 21 693 76 00
e-mail: dominique.zosso@epfl.ch

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

GEODESIC ACTIVE FIELDS – A GEOMETRIC FRAMEWORK FOR IMAGE REGISTRATION

Dominique Zosso, Xavier Bresson, Jean-Philippe Thiran

Ecole Polytechnique Fédérale de Lausanne (EPFL)
Signal Processing Laboratories LTS5

Technical Report LTS-2010-001

January 22, 2010

Part of this work will be submitted to “IEEE Transactions on Image Processing”.
This work has been supported by the National Competence Centre in Biomedical Imaging (NCCBI).

Geodesic Active Fields – A Geometric Framework for Image Registration

Dominique Zosso*, Xavier Bresson, Jean-Philippe Thiran

January 22, 2010

Abstract

In this paper we present a novel framework called geodesic active fields for general image registration on Riemannian manifolds. In image registration, one looks for the underlying deformation field that best maps one image onto another. This is a classic ill-posed inverse problem, which is usually solved by adding a regularization term. In this paper, we define a new geometric approach to image registration. More specifically, we propose to embed the deformation field in a higher dimensional manifold. Then, the deformation field is driven by a minimization flow towards a harmonic map corresponding to the solution of the registration problem, much like geodesic active contours in image segmentation. The energy of the deformation field is measured with the Polyakov energy weighted by a suitable image distance, borrowed from standard registration models. We investigate three different weighting functions, the squared error and the approximated absolute error for monomodal images, and the local joint entropy for multimodal images. Further, the concept of exponential maps of vector fields is integrated into the framework, to allow for smoothly invertible diffeomorphic deformation fields. Finally, we illustrate the validity of our framework on a set of different image registration tasks, including non-flat and multiscale images.

Biomedical image processing, Computational geometry, Differential geometry, Diffusion equations, Image registration, Image representations, Image shape analysis, Partial differential equations, Scale-spaces, Surfaces.

1 Introduction

Image registration is the concept of mapping homologous points of different images, representing a same object. Homology, in turn, is defined as the relation between "organs deriving from the same embryonic blanks"[1]. This fundamental concept is illustrated in Fig. 1.

In practice, however, it is highly difficult to establish homology in images strictly based on this definition. For automatic image registration, it is therefore commonplace to substitute homology by a measurable criterion of image dissimilarity, which is to be minimized. Depending on the nature of the images to be registered, different metrics are used to assess image distances. But also the deformation model and constraints that are applied on that deformation field can vary, as well as the optimization technique that is used to solve the minimization [2–5].

*This work is supported by the Swiss National Competence Center in Biomedical Imaging (NCCBI), and by ONR N00014-03-1-0071 and ONR MURI subcontract from Stanford University. D. Zosso and J.-P. Thiran are with the Signal Processing Laboratory (LTS5), Ecole Polytechnique Fédérale de Lausanne (EPFL), Station 11, CH-1015 Lausanne, Switzerland, {dominique.zosso,jp.thiran}@epfl.ch. X. Bresson is with the Department of Mathematics, University of California (UCLA), Los Angeles, CA 90095-1555, USA, xbresson@math.ucla.edu.

[1] *Homologue:"sont homologues les organes dérivant des mêmes ébauches embryonnaires"*. Le trésor de la langue française informatisé.

Let a deformation field $\mathbf{u}(\mathbf{x}) : \mathbf{x} \in \Omega \subset \mathbb{R}^n \rightarrow (u_1, \ldots, u_p) \in \mathbb{R}^p$ describe the spatial displacement along $p \leq n$ dimensions of an n-dimensional image of support Ω. The determination of this underlying deformation field between two images is an ill-posed inverse problem, requiring additional prior knowledge. On one hand, parametric deformation models, including rigid and affine transformations, which are defined globally for the whole image space, restrict the degrees of freedom to a small number of parameters. On the other hand, freeform deformations allow for an individual local displacement of each point in the image domain. To restrict such deformation fields to what is believed to be "physically meaningful" deformations, constraints on the field regularity are introduced. Typical regularization constraints reduce the variations of the deformation field $\mathbf{u}(\mathbf{x})$ by defining an additional penalty, e.g., the vectorial total variation [6]:

$$\int_\Omega |\nabla \mathbf{u}(\mathbf{x})| \mathrm{d}\mathbf{x} = \int_\Omega \sqrt{\sum_{k=1}^{n} |\nabla u_k|^2} \mathrm{d}\mathbf{x}. \tag{1}$$

Thus, image distance metric and regularization penalty are commonly incorporated into a single energy minimization model, a.k.a. variational model, e.g., [7]. The energy functionals are of the general form

$$E = E_{\text{data}} + \alpha \cdot E_{\text{regularization}}, \quad \alpha > 0. \tag{2}$$

As a typical instance consider an energy functional consisting of the L^2-norm between the fix and the moving image, $F, M : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, regularized by the forementioned vectorial total variation:

$$E = \int_\Omega \left(M(\mathbf{x}+\mathbf{u}) - F(\mathbf{x})\right)^2 \mathrm{d}\mathbf{x} + \alpha \cdot \int_\Omega |\nabla \mathbf{u}(\mathbf{x})| \mathrm{d}\mathbf{x}. \tag{3}$$

The balancing parameter α can have a severe impact on the registration result. Its choice is arbitrary and the optimal depends on several conditions.

In their seminal work [8], Sochen, Kimmel and Malladi introduced the powerful Beltrami framework for image denoising and enhancement. This model is a generalization of the Polyakov model [9] introduced in string theory for physics. The Polyakov model represents strings as harmonic maps in high-dimensional and curved spaces defined by Riemannian manifolds. Adopting this pure geometric point of view amounts to seeing objects such as images, shapes, or vector fields as harmonic maps. Recently, a new regularisation criterion derived from the Beltrami framework was introduced in optical flow modeling [10]. There, the authors embed the optical flow field as harmonic map and propose to use the Polyakov energy as regularisation term.

Independently, a weighted version of the Polyakov energy was used to perform segmentation of images defined on Riemannian manifolds using active contours [11]. The weighting allows to combine the edge detection function and the regularisation term by multiplication, and provides an elegant framework for e.g. multiscale image segmentation.

The goal of this work is to define a novel image registration scheme using a geometric approach. We embed the deformation field in a higher dimensional space and define a variational model using the weighted Polyakov energy. While the Polyakov energy itself only provides a regularity constraint – harmonic map –, the weighting allows to drive the deformation field towards low image dissimilarity. In analogy to geodesic active contours in segmentation [12], we call this model *geodesic active fields* for image registration. The proposed approach directly generalizes to non-Euclidean images, and thus automatically allows to work e.g. with non-flat or multiscale images.

The next section will present the Polyakov energy and its weighted version, as well as the corresponding minimizing flows. In section 3 we show how the Polyakov framework can be used to define a geometric image registration model. We extend the framework to multiscale image registration in section 4. Section 5 studies different weighting functions. Then, section 6 presents the integration of exponential mapping into the registration scheme to provide diffeomorphic deformation fields. We detail the implementation of a converging mean curvature estimator in section 7. Finally, we show some results obtained with our geodesic active fields framework in section 8 and we discuss our model in section 9.

2 Weighted Polyakov Energy

Sochen, Kimmel and Malladi introduced in [8, 13] a general geometrical framework for low-level vision. In this framework, images are seen as embedding maps between two Riemannian manifolds. An energy functional defined by Polyakov in [9] provides a measure on the embedding maps, and a flow that minimizes this measure.

2.1 Beltrami framework

An n-dimensional manifold Σ with coordinates $\sigma^{1\cdots n}$ is embedded in an m-dimensional manifold M with coordinates $X^{1\cdots m}$, with $m > n$. The embedding map $X : \Sigma \mapsto M$ is given by m functions of n variables. For example, a 2D grey-level image can be seen as a surface embedded in 3D: $X : (x,y) \to (x,y,I)$. A Riemannian structure can be introduced: the metric $g_{\mu\nu}$ locally measures the distances on Σ, whereas on M distances are measured using h_{ij}.

To measure the weight of the mapping $X : \Sigma \mapsto M$, Sochen *et al.* use the Polyakov energy, known from high energy physics [9], as a natural generalization of the L^2-norm on the embedded image to manifolds:

$$S[X^i, g_{\mu\nu}, h_{ij}] = \int \sqrt{g} g^{\mu\nu} \partial_\mu X^i \partial_\nu X^j h_{ij} \mathrm{d}^m\sigma, \tag{4}$$

where the Einstein summation convention is used, g is the determinant of the image metric, and $g^{\mu\nu}$ is its inverse, such that $g^{\mu\nu} g_{\nu\gamma} = \delta^\mu_\gamma$ (δ^μ_γ is the Kronecker delta). Naturally, the metric g is chosen as the induced metric, obtained by the *pullback*-relation: $g_{\mu\nu} = h_{ij} \partial_\mu X^i \partial_\nu X^j$. Under such a metric, the Polyakov energy shortens to:

$$S = \int \sqrt{g} \mathrm{d}^m\sigma, \tag{5}$$

and represents the area of the embedded image surface. Using the Euler-Lagrange equation technique from calculus of variations, the following minimizing flow is obtained:

$$\partial_t X^i = \frac{1}{\sqrt{g}} \partial_\mu \left(\sqrt{g} g^{\mu\nu} \partial_\nu X^i \right) + \Gamma^i_{jk} \partial_\mu X^j \partial_\nu X^k g^{\mu\nu}, \tag{6}$$

where the Levi-Civita connection Γ^i_{jk}, also called the Christoffel symbol, is defined as

$$\Gamma^i_{jk} = \frac{1}{2} h^{il} \left(\partial_j h_{kl} + \partial_k h_{jl} - \partial_l h_{jk} \right). \tag{7}$$

Assuming the embedding is in a Euclidean space with Cartesian coordinates, the Christoffel symbols are all equal to zero, and the corresponding gradient descent equation is

$$\partial_t X^i = -\frac{1}{\sqrt{g}} \frac{\delta S}{\delta X^i} = \frac{1}{\sqrt{g}} \partial_\mu (\sqrt{g} g^{\mu\nu} \partial_\nu X^i) \equiv H^i, \tag{8}$$

known as the Beltrami flow, where H^i denotes the i-th component of the mean curvature vector of the manifold.

An interesting property of the Beltrami framework is the freedom to choose the metric of the embedding space. For example, let us embed a 2D graylevel image in 3D, using the following metric tensor:

$$h_{ij} = \mathrm{diag}(1, 1, \beta^2), \tag{9}$$

where $\beta > 0$ is a constant. This allows to set the scale of the feature dimension independently of the spatial dimensions. The pullback relation yields the metric tensor $g_{\mu\nu}$:

$$g_{\mu\nu} = \begin{bmatrix} 1+\beta^2 I_x^2 & \beta^2 I_x I_y \\ \beta^2 I_x I_y & 1+\beta^2 I_y^2 \end{bmatrix}. \tag{10}$$

Its determinant is given by $g = 1+\beta^2|\nabla I|^2$. Thus, the Polyakov energy of the embedding reads:

$$S = \int \sqrt{1+\beta^2|\nabla I|^2}\,\mathrm{d}^m\sigma. \tag{11}$$

If $\beta \to \infty$, the 1 in this energy becomes negligible, and the energy approaches the TV-norm, well-known in image denoising [14, 15]. If, however, $\beta \to 0$, then the minimizing flow approaches the isotropic heat diffusion [8]. The impact of β on the apparent feature amplitudes of an embedded scalar field is illustrated in Fig. 2a).

Moreover, the features being considered within the Beltrami framework are not restricted to scalar values only, but generalize directly to any co-dimension. For explicit applications of the framework to denoising of color images and textures, we refer the reader to [16]. For a review of the framework over a variety of manifolds and data structures, see [17].

2.2 Weighted Polyakov energy

Here, we present a weighted version of the Polyakov energy. To demonstrate the validity of the concept, we show how it generalizes the geodesic active contour model to images defined on non-Euclidean manifolds.

In [11] the Polyakov energy was tuned by a weighting function f:

$$S_f = \int f\sqrt{g}\,\mathrm{d}^m\sigma, \tag{12}$$

where $f = f(X^i, g_{\mu\nu}, h_{ij})$. Still assuming the embedding is in a Euclidean space with Cartesian coordinates, i.e., h_{ij} is diagonal and constant, the corresponding gradient descent equation is

$$\partial_t X^i = fH^i + \partial_k f g^{\mu\nu}\partial_\mu X^k \partial_\nu X^i - \frac{m\cdot n}{2}\partial_k f h^{ki}, \tag{13}$$

where fH^i corresponds to a weighted mean curvature flow on manifolds.

Now, one can show that the geodesic active contour model (GAC) [12,18] is a special case of the weighted Polyakov model. More interestingly, we generalize the GAC model to evolve on any curved surfaces defined by a Riemannian manifold.

Let us start with the GAC model. Let us consider the mapping

$$X = C : p \to (C^1(p), C^2(p)), \tag{14}$$

p being any curve parameterization. (C^1, C^2) are the coordinates of the curve embedded in the standard flat Euclidean plane described by the metric tensor $h_{ij} = \delta_{ij}$.

Energy (12) of this mapping is the GAC energy:

$$S_f = \int f|C_p|\mathrm{d}p = \int f\mathrm{d}s, \tag{15}$$

if the arclength parameterization s is used. Besides, the minimization flow (13) is also the GAC flow:

$$\partial_t C = f\kappa\mathcal{N} - \langle \nabla f, \mathcal{N}\rangle \mathcal{N}, \tag{16}$$

where $\mathcal{N}, \kappa$ are resp. the normal and the curvature to the curve C.

Let us now consider the extension of the GAC model to any manifold, i.e., a curve C embedded on a 2-D manifold M. The GAC is still represented by a mapping

$$X : s \to C(s) = (C^1(s), C^2(s)) \in M, \tag{17}$$

s being the planar arclength parameter of C. The manifold M is defined by a parametric surface such that:

$$(u^1, u^2) \to X_M(u^1, u^2) = (X^1(u^1, u^2), X^2(u^1, u^2), X^3(u^1, u^2)) \in \mathbb{R}^3, \tag{18}$$

where $(u^1, u^2) \in \mathbb{R}^2$ are the parameterization coordinates on M. The metric tensor of M is given by $h_{ij} = \langle \frac{\partial X_M}{\partial u^i}, \frac{\partial X_M}{\partial u^j} \rangle$. The energy of the GAC on the manifold M is given by (12) as follows:

$$\begin{cases} S_f(C) = \int f |C_s|_M ds \\ |C_s|_M = \sqrt{C_s^i C_s^j h_{ij}} \end{cases}, \tag{19}$$

where $|C_s|_M$ is the norm of the vector C_s on the manifold M. The minimizing flow is provided by (13) as:

$$\partial_t C = \left(f \cdot \frac{\kappa + \langle \Gamma_{ij} C_s^i C_s^j, \mathcal{N} \rangle}{h_{ij} C_s^i C_s^j} - \langle \nabla^M f, \mathcal{N} \rangle \right) \mathcal{N}, \tag{20}$$

where ∇^M is the gradient operator defined on M such that $\nabla_l^M = h^{kl}\partial_k$, and Γ are the Levi-Civita coefficients.

The general variational model (19) along with the minimizing flow (20) have been used in [19] to design active contours on omni-directional images. Omnidirectional images are defined on curved surfaces such as the sphere, the paraboloid and the hyperboloid.

It is also interesting to notice that the evolution flow (20) is the same flow given by [20]. Spira and Kimmel have considered the Riemannian counterparts of the original GAC model (16), whereas our approach has derived the general flow from the Polyakov variational model.

3 Geodesic Active Fields

In this section we define the general evolution equation for the geodesic active fields for image registration. The main concept is to embed a deformation field, that maps two n-dimensional images, in a m-dimensional Euclidean space. A very simple such embedding is illustrated in Fig. 2b). The embedded manifold then evolves towards a minimal surface, while being attracted by configurations that bring the two images into registration.

We first define the model for the general case of Euclidean images, before refining it to specific applications, such as stereo vision or registration on non-flat manifolds.

3.1 The general Euclidean case

Let us first consider the case of n-dimensional images defined on well-known Euclidean domains $\Omega \subset \mathbb{R}^n$ with Cartesian coordinates $\mathbf{x} = (x_1, \ldots, x_n)$. We look for a deformation field acting along $p \leq n$ dimensions, i.e., $\mathbf{u} : \Omega \subset \mathbb{R}^n \mapsto \mathbb{R}^p, \mathbf{u} = (u_1, \ldots, u_p)\{\mathbf{x}\}$.

The embedding X and the metric tensors h_{ij} and $g_{\mu\nu}$ are chosen as follows:

$$\begin{cases} X : (x_1, \ldots, x_n) \to (x_1, \ldots, x_n, u_1, \ldots, u_p) \\ h_{ij} = \mathrm{diag}(1, \ldots, 1, \beta^2, \ldots, \beta^2) \\ g_{\mu\nu} = \partial_\mu X^i \partial_\nu X^j h_{ij} = \delta_{\mu\nu} + \beta^2 \sum_{i=1}^{p} \partial_\mu u_i \partial_\nu u_i, \end{cases} \tag{21}$$

where $x_1, \ldots, x_n$ are the n spatial components of the image and $u_1, \ldots, u_p$ are the components of the dense deformation field. β is the scaling factor applied to the deformation field components to get the desired aspect ratio. These equations are introduced in the weighted Polyakov functional (12) and its minimization flow (13), leading to the following general Euclidean registration energy functional and the minimizing evolution flow:

$$\begin{cases} E_{GAF} = \int f \sqrt{g} \prod_{i=1}^{n} \mathrm{d}x_i \\ \partial_t u_i = f H^{n+i} + \partial_k f g^{\mu\nu} \partial_\mu X^k \partial_\nu u_i - \frac{m \cdot n}{2\beta^2} f_{u_i}, \end{cases} \tag{22}$$

where the weighting function $f = f(\mathbf{x}, \mathbf{u})$ is arbitrary, and will be defined in section 5.

3.2 Stereo vision

In stereo vision, the depth information corresponding to a location is encoded as the lateral shift between its representation in two adjacent image acquisitions. The recovered depth information is used in e.g. satellite imaging or robot vision to reconstruct the observed scene. The lateral shift can be determined by registration of the two images, where only lateral deformation is allowed. That is, the deformation field has only one component, along the x-dimension. We choose the following embedding and metric tensors, corresponding to the illustration in Fig. 2b):

$$\begin{cases} X : (x, y) \to (x, y, u) \\ h_{ij} = \mathrm{diag}(1, 1, \beta^2) \\ g_{\mu\nu} = \begin{bmatrix} 1 + \beta^2 u_x^2 & \beta^2 u_x u_y \\ \beta^2 u_x u_y & 1 + \beta^2 u_y^2 \end{bmatrix}, g = 1 + \beta^2 |\nabla u|^2. \end{cases} \tag{23}$$

Introducing those equations into (22), we get the following energy functional and evolution equation:

$$\begin{cases} E_{GAF} = \int f \sqrt{1 + \beta^2 |\nabla u|^2} \mathrm{d}x\mathrm{d}y \\ \partial_t u = f H^u + \partial_k f g^{\mu\nu} \partial_\mu X^k \partial_\nu u - \frac{3}{\beta^2} f_u, \end{cases} \tag{24}$$

where H^u is simply the 3rd component of the mean curvature vector:

$$H^u = \frac{g_{11} u_{yy} - 2 \cdot g_{12} u_{xy} + g_{22} u_{xx}}{g^2}. \tag{25}$$

3.3 2-D image registration

In the case of registration, involving deformations along all image dimensions, one has $p = n$ and $m = 2n$. Here, as an example without loss of generality, we consider the registration of 2D images. The deformation field is described by u and v, resp. along x and y:

$$(u, v) : (x, y) \in \Omega \subset \mathbb{R}^2 \mapsto \mathbb{R}^2, (u, v) = (u(x, y), v(x, y)). \tag{26}$$

We choose the following embedding and metric tensors:

$$\begin{cases} X : (x,y) \to (x,y,u,v) \\ h_{ij} = \mathrm{diag}(1,1,\beta^2,\beta^2) \\ g_{\mu\nu} = \begin{bmatrix} 1+\beta^2(u_x^2+v_x^2) & \beta^2(u_x u_y + v_x v_y) \\ \beta^2(u_x u_y + v_x v_y) & 1+\beta^2(u_y^2+v_y^2) \end{bmatrix} \\ g = 1+\beta^2(|\nabla u|^2+|\nabla v|^2)+\beta^4(\nabla u,\nabla v)^2, \end{cases} \tag{27}$$

where $(\nabla u, \nabla v) \equiv u_x v_y - u_y v_x$ is the magnitude of the cross product of the gradient vectors ∇u and ∇v. The expression of the determinant g has become quite cumbersome. The term $\beta^4(\nabla u,\nabla v)^2$ measures the misalignment of the gradients between different deformation field components [16]. All these settings put into the general equations produce the following energy functional and minimizing flow:

$$\begin{cases} E_{GAF} = \int f\sqrt{1+\beta^2(|\nabla u|^2+|\nabla v|^2)+\beta^4(\nabla u,\nabla v)^2}\mathrm{d}x\mathrm{d}y \\ \partial_t u = fH^u + \partial_k f g^{\mu\nu}\partial_\mu X^k \partial_\nu u - \frac{4}{\beta^2} f_u \\ \partial_t v = fH^v + \partial_k f g^{\mu\nu}\partial_\mu X^k \partial_\nu v - \frac{4}{\beta^2} f_v. \end{cases} \tag{28}$$

3.4 Registration on non-flat manifolds

One of the main advantages of the proposed framework is that the image domain does not necessarily have to be Euclidean. Indeed, images can be defined on any Riemannian manifold. In the Euclidean case, the spatial coordinates were directly given by the image domain parameters. In the non-Euclidean case, the spatial coordinates of the image are more complicated functions of the domain parameters instead.

To give a simple example, consider a spherical patch S described by two angles, θ and ϕ, on which the images are defined:

$$\begin{cases} S : (\theta,\phi) \in \Omega \subset \mathbb{R}^2 \to (x,y,z) \in \mathbb{R}^3 \\ \Omega = [\theta_{\min},\theta_{\max}] \times [\phi_{\min},\phi_{\max}] \\ 0 < \theta_{\min} < \theta_{\max} < \pi, \quad 0 \le \phi_{\min} < \phi_{\max} < 2\pi \\ x(\theta,\phi) = \sin\theta\cos\phi \\ y(\theta,\phi) = \sin\theta\sin\phi \\ z(\theta,\phi) = \cos\theta \end{cases} \tag{29}$$

The induced metric on S is naturally given by $g_S = \mathrm{diag}(1,\sin^2\theta)$. Further, let the deformation field $(\vartheta(\theta,\phi),\varphi(\theta,\phi))$ act on the two angles describing the patch. This suggests the following embedding:

$$\begin{cases} X : (\theta,\phi) \in \Omega \subset \mathbb{R}^2 \to (\theta,\phi,\vartheta,\varphi) \in \mathbb{R}^4 \\ h_{ij} = \mathrm{diag}(g_S, \beta^2 g_S) = \mathrm{diag}(1,\sin^2\theta,\beta^2,\beta^2\sin^2\theta) \end{cases} \tag{30}$$

where the metric tensor h_{ij} has been set by taking the induced metric g_S of the patch parameterization into account. The pullback relation yields the following metric tensor $g_{\mu\nu}$ in parameter space:

$$\begin{aligned} g_{\mu\nu} = &\begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix} \\ &+\beta^2 \begin{bmatrix} \vartheta_\theta^2 + \sin^2\theta\varphi_\theta^2 & \vartheta_\theta\vartheta_\phi + \sin^2\theta\varphi_\theta\varphi_\phi \\ \vartheta_\theta\vartheta_\phi + \sin^2\theta\varphi_\theta\varphi_\phi & \vartheta_\phi^2 + \sin^2\theta\varphi_\phi^2 \end{bmatrix}. \end{aligned} \tag{31}$$

Given this metric tensor h_{ij}, the embedding space is not Euclidean anymore, and the computation of the mean curvature vector involves the Levi-Civita connection as in (6), to account for the Riemannian part.

For the spherical patch, only two relevant Christoffel symbols computed by (7) differ from zero:

$$\Gamma^{\varphi}_{jk} = \begin{cases} \cot\theta & j \mathrel{\widehat{=}} \varphi, k \mathrel{\widehat{=}} \theta, \\ \cot\theta & j \mathrel{\widehat{=}} \theta, k \mathrel{\widehat{=}} \varphi, \\ 0 & \text{otherwise}, \end{cases} \tag{32}$$

and $\Gamma^{\vartheta}_{jk} = 0 \forall j,k$, where, with some abuse of notation, $\Gamma^{\varphi}_{jk} = \Gamma^{i}_{jk}|_{i \mathrel{\widehat{=}} \varphi}$. This gives the following evolution equations for the deformation field:

$$\begin{cases} \partial_t \vartheta = f H^{\vartheta} + \partial_k f g^{\mu\nu} \partial_\mu X^k \partial_\nu \vartheta - \frac{4}{\beta^2} f_{\vartheta}, \\ \partial_t \varphi = f H^{\varphi} + \partial_k f g^{\mu\nu} \partial_\mu X^k \partial_\nu \varphi - \frac{4}{\beta^2 \sin^2\theta} f_{\varphi}, \\ H^{\vartheta} = \frac{1}{\sqrt{g}} \partial_\mu \left(\sqrt{g} g^{\mu\nu} \partial_\nu \vartheta \right), \\ H^{\varphi} = \frac{1}{\sqrt{g}} \partial_\mu \left(\sqrt{g} g^{\mu\nu} \partial_\nu \varphi \right) + 2\cot\theta (g^{\theta\theta} \varphi_\theta + g^{\theta\phi} \varphi_\phi), \end{cases} \tag{33}$$

where $g^{\theta\theta} = g^{\mu\nu}|_{\mu \mathrel{\widehat{=}} \theta, \nu \mathrel{\widehat{=}} \theta}$, and $g^{\theta\phi} = g^{\mu\nu}|_{\mu \mathrel{\widehat{=}} \theta, \nu \mathrel{\widehat{=}} \phi}$.

4 Multiscale Image Registration

4.1 Motivation

The way we perceive the world depends on the scale of the aperture we use to measure it [21]. This principle is trivially known as "missing the forest for the trees", and has also particular validity in image processing: images are naturally composed of objects which are meaningful only at certain scales of observation [22, 23].

As an example in image registration let us consider the human brain. It exhibits a highly convoluted and irregular structure, with high complexity and variability. For example, sulci and gyri vary a lot between subjects. On the other hand, high level structures of the brain – the "big picture" – are highly conserved, such as the two hemispheres, the lobes and main folds. Hence, a hierarchical representation of these structures is important in the context of intersubject registration: considering the complexity of the cortical surface, directly involving local small-scale features would mislead the registration to be trapped in bad local minima. A robust method needs to rely on large-scale features, describing the main landmarks of the cortex, such as the main gyri or sulci, while small-scale features drive the registration more locally to reach the desired precision [24].

As a consequence of the above, the scale at which one measures a certain property becomes an additional dimension of the imaging space, giving rise to Witkin's patented notion of a *scale-space* [25]. Witkin introduced the concept of artificially generating larger (coarser) scales of an image through low-pass filtering.

The geometry of a large class of scale-spaces can be defined by a general metric tensor [26, 27]:

$$h_{ij} = \operatorname{diag}\left(\underbrace{\frac{1}{c^2}, \ldots, \frac{1}{c^2}}_{n \text{ times}}, \frac{1}{c^2 \rho^2} \right), \tag{34}$$

where the first n elements of the diagonal correspond to the spatial dimensions $x_1, \ldots, x_n$, and the last element refers to the scale σ. c and ρ are two functions that represent the conductance and the density in the general model of heat transfer. The spatial derivative within such a scale-space is now obtained as $c\nabla$, whereas the scale derivative is given

by $\rho c\partial_\sigma$. The natural heat equation, that defines the scale-space, is:

$$(\rho c\partial_\sigma)I = (c\nabla)\cdot(c\nabla)I, \tag{35}$$

$$\partial_\sigma I = \frac{1}{\rho}\nabla\cdot(c\nabla I). \tag{36}$$

Different choices for c and ρ yield different well-known scale-spaces. The linear scale-space, e.g., corresponds to $c=\sigma$ and $\rho=1$: $\partial_\sigma I=\sigma\Delta I$. The Perona-Malik scale-space is reproduced with $\rho=1$ and $c=\exp(-\alpha|\nabla I|^2),\alpha>0$ [28]. The Beltrami flow of Sochen-Kimmel-Malladi requires $c=\rho=\frac{1}{\sqrt{1+\beta^2|\nabla I|^2}}$ [8,13]. The linear and the Beltrami scale-space are illustrated at the example of the fractal image of a *Von Koch snowflake*, and a single slice of a T1-weighted brain MR image in Fig. 3.

4.2 Multiscale active deformation fields

Multiscale images have an additional image dimension: the scale σ. Along this scale-dimension, no deformation takes place. The multiscale deformation field is embedded as follows:

$$\begin{cases} X:(x_1,\ldots,x_n,\sigma)\to(x_1,\ldots,x_n,\sigma,u_1,\ldots,u_n)\\ h_{ij}=\mathrm{diag}(\frac{1}{c^2},\ldots,\frac{1}{c^2},\frac{1}{c^2\rho^2},\frac{\beta^2}{c^2},\ldots,\frac{\beta^2}{c^2})\\ g_{\mu\nu}=\partial_\mu X^i\partial_\nu X^j h_{ij}, \end{cases} \tag{37}$$

where the structure of the metric tensor h_{ij} is arbitrary, and inspired by (34).

Considering a linear scale-space, i.e., $c=\sigma$ and $\rho=1$, the embedding thus looks like:

$$\begin{cases} X:(x_1,\ldots,x_n,\sigma)\to(x_1,\ldots,x_n,\sigma,u_1,\ldots,u_n)\\ h_{ij}=\frac{1}{\sigma^2}\mathrm{diag}(1,\ldots,1,1,\beta^2,\ldots,\beta^2)\\ g_{\mu\nu}=\partial_\mu X^i\partial_\nu X^j h_{ij}=\frac{1}{\sigma^2}\left(\delta_{\mu\nu}+\beta^2\sum_{i=1}^{n}\partial_\mu u_i\partial_\nu u_i\right). \end{cases} \tag{38}$$

Again, as for non-flat image domains, the multiscale embedding is not Euclidean, and the Levi-Civita connection (7) is required to compute the complete mean curvature vector according to (6).

Note, that the deformation field $\mathbf{u}=\mathbf{u}(\mathbf{x},\sigma)$ evolves at all scales *simultaneously*. At each scale, the deformation field is attracted by the corresponding data term, while coherence between scales is obtained thanks to the regularizing power of harmonic maps.

4.3 Multiscale 2-D image registration

In the case of 2D images to be registered, the only relevant non-zero Christoffel symbols computed as (7) are $\Gamma^u_{u\sigma}=\Gamma^u_{\sigma u}=\Gamma^v_{v\sigma}=\Gamma^v_{\sigma v}=-\frac{1}{\sigma}$. The evolution equations for both components (u,v) of the deformation field along (x,y) are

$$\begin{cases} \partial_t u = fH^u+\partial_k f g^{\mu\nu}\partial_\mu X^k\partial_\nu u-\frac{15\sigma^2}{2\beta^2}f_u\\ \partial_t v = fH^v+\partial_k f g^{\mu\nu}\partial_\mu X^k\partial_\nu v-\frac{15\sigma^2}{2\beta^2}f_v\\ H^u=\frac{1}{\sqrt{g}}\partial_\mu\left(\sqrt{g}g^{\mu\nu}\partial_\nu u\right)-\frac{2}{\sigma}\partial_\mu u g^{\mu\sigma}\\ H^v=\frac{1}{\sqrt{g}}\partial_\mu\left(\sqrt{g}g^{\mu\nu}\partial_\nu v\right)-\frac{2}{\sigma}\partial_\mu v g^{\mu\sigma} \end{cases}, \tag{39}$$

where, with some abuse of notation, $g^{\mu\sigma}$ denotes the column ν of the inverse of the metric tensor $g^{\mu\nu}$ corresponding to the scale σ.

5 What Choice of Weighting Function?

The purpose of the weighting function f is to drive the deformation field towards minimal surfaces that bring the two images into registration. As such, the flow must stop when the deformed image perfectly matches the target image. Hence, the weighting function is naturally chosen to be an image distance metric, which approaches zero when the two images match.

In practice, we found useful to extend the weighting function by a positive constant, to give more weight to regularization. The general form of the weighting function is thus

$$f = 1 + \alpha f^{(i)}, \tag{40}$$

where $f^{(i)}$ is a specific image distance metric, as defined below, and α is the balancing parameter. This form represents a Polyakov energy functional, where the image distance metric corresponds to an additional penalty weighting.

5.1 Deformation model

The weighting function is the place, where the deformation field actually gets to act on the images. Therefore, it is crucial to define the particular deformation model we want to use. First, we work with Euler coordinates. That is, for any pixel in the fixed imaged, the corresponding pixel is looked up in the moving image using a coordinate mapping. The corresponding location in the moving image will almost never fall on an exact pixel location and interpolation will be required.

Here, we use a very simple scheme, where the lookup is based on a shift by addition. The transform operator T is thus defined as:

$$T_0 : T_0\mathbf{x} = \mathbf{x} + \mathbf{u}(\mathbf{x}), \tag{41}$$

where addition is implicitely understood only along the $p \leq n$ dimensions of the image that are deformed. Also, for simplicity we shall ignore any boundary issues and finite support. In section 6 we will introduce a more complicated deformation model, based on exponential maps of vector fields.

Once corresponding fixed an moving image locations can be mapped, the matching quality can be quantified using one of several distance metrics.

5.2 Squared error

If the images have been acquired using similar sensors, one can generally assume that the same entities are pictured at the same feature intensity in both images. An intuitive and simple choice for monomodal image registration subject to additive Gaussian noise is the squared error metric [29]:

$$f^{(1)T_0}(\mathbf{x},\mathbf{u}) = (M(T_0\mathbf{x}) - F(\mathbf{x}))^2 = (M(\mathbf{x}+\mathbf{u}(\mathbf{x})) - F(\mathbf{x}))^2, \tag{42}$$

where F and M refer to the fix and moving images respectively. The evolution equation (22) includes the partial derivatives of the weighting function with respect to all components of the embedding. For the function given in (42), these are obtained as follows:

$$\begin{cases} f_{\mathbf{x}}^{(1)T_0} = 2 \cdot (M(T_0\mathbf{x}) - F(\mathbf{x})) \cdot \left(J^T \nabla M(T_0\mathbf{x}) - \nabla F(\mathbf{x})\right) \\ f_{\mathbf{u}}^{(1)T_0} = 2 \cdot (M(T_0\mathbf{x}) - F(\mathbf{x})) \cdot \nabla M(T_0\mathbf{x}) \end{cases} \tag{43}$$

where ∇F and ∇M refer to the gradients of the fix and moving images, respectively, and where J^T denotes the transpose of the Jacobian of the deformed field:

$$J_{ij} = \delta_{ij} + \frac{\partial u_i}{\partial x_j}. \tag{44}$$

5.3 Local joint entropy

If images of different modality are to be registered, the above squared error metric is not a suitable distance metric anymore. Instead, mutual information is a commonly accepted similarity criterion in this case [30–32].

Mutual information is a global measure on the joint (p^{fm}) and marginal (p^f and p^m) histograms of the fixed and moving images:

$$MI = \sum_{i_1,i_2} p^{fm} \ln(p^{fm}) - \sum_{i_1} p^f \ln(p^f) - \sum_{i_2} p^m \ln(p^m), \tag{45}$$

where $p^{fm} = p^{fm}(i_1, i_2)$ etc. Let us assume, that the marginal entropies remain constant throughout the whole registration process, as they only depend on the fix and moving image separately. Maximizing mutual information is thus equal to minimizing the joint entropy.

The same joint entropy, i.e., the expectation of the negative logarithm of the joint probability, can also be computed in the image domain, instead of using the above histograms [33]:

$$H^{fm} = -\sum_{i_1,i_2} p^{fm} \ln(p^{fm}) = \frac{1}{|\Omega|} \sum_{\mathbf{x}} -\ln(p^{fm}), \tag{46}$$

where $p^{fm} = p^{fm}(F(\mathbf{x}), M(\mathbf{x}+\mathbf{u}))$. The negative logarithm denotes the local joint entropy. This local joint entropy has a minimum value of 0 (if the joint probability matches 1), and is unbound positive. This provides us with a local measure that corresponds well to the weighting function criteria stated above.

Consequently, we define the following information theory based weighting function for multi-modal image registration:

$$f^{(2)T_0}(\mathbf{x}, \mathbf{u}) = -\ln(p^{fm}(F(\mathbf{x}), M(T_0\mathbf{x}))). \tag{47}$$

Using this weighting function, the goodness of a local alignment is measured by the frequency of similar intensity pairs in the rest of the image.

The partial derivatives along spatial components $f^{(2)}_{\mathbf{x}}$ are again best estimated numerically. The partial derivatives along deformation field components are obtained using the chain rule:

$$f^{(2)T_0}_{\mathbf{u}} = -\frac{p^{fm}_m(F(\mathbf{x}), M(T_0\mathbf{x}))}{p^{fm}(F(\mathbf{x}), M(T_0\mathbf{x}))} \cdot \nabla M(T_0\mathbf{x}), \tag{48}$$

where $p^{fm}_m(i_1, i_2)$ is the partial derivative of the histogram along the dimension corresponding to the moving image.

Estimation of the joint histogram $p^{fm}(i_1, i_2)$ and its partial derivative $p^{fm}_m(i_1, i_2)$ can be done in several ways. Some authors propose Parzen windowing with Gaussian kernels, e.g., [31], while others use B-spline basis functions for histogram smoothing [34, 35].

The simplest way consists in binning the histograms sufficiently coarse, and estimating the derivative $p^{fm}_m(i_1, i_2)$ of the joint probability distribution numerically, using finite differences.

5.4 Absolute error

For non-smooth deformation fields, e.g., observed in optical flow-based image registration, the L^1-norm may perform better as data term [36]. The L^1-norm measures the absolute error between the two images,

$$f^{(3)T_0}(\mathbf{x},\mathbf{u}) = |M(T_0\mathbf{x}) - F(\mathbf{x})|\,, \tag{49}$$

and it can be approximated by a differentiable function:

$$f^{(3)T_0}(\mathbf{x},\mathbf{u}) = \sqrt{(M(T_0\mathbf{x}) - F(\mathbf{x}))^2 + \varepsilon^2}, \tag{50}$$

where $1 \gg \varepsilon > 0$, typically $\varepsilon \sim 10^{-3}$. The partial derivatives of the approximated function are obtained easily:

$$\begin{cases} f_{\mathbf{x}}^{(3)T_0} = \frac{(M(T_0\mathbf{x}) - F(\mathbf{x}))}{f^{(3)T_0}(\mathbf{x},\mathbf{u})} \cdot \left(J^T \nabla M(T_0\mathbf{x}) - \nabla F(\mathbf{x})\right), \\ f_{\mathbf{u}}^{(3)T_0} = \frac{(M(T_0\mathbf{x}) - F(\mathbf{x}))}{f^{(3)T_0}(\mathbf{x},\mathbf{u})} \cdot \nabla M(T_0\mathbf{x}). \end{cases} \tag{51}$$

6 Diffeomorphic Deformations

6.1 Motivation

The dense deformation field given by $\mathbf{u}$ defines a surjective homomorphism, i.e., it maps an image into another image. Nevertheless, it is not necessarily injective, that is, several points of the initial image can be mapped to the same position after deformation, corresponding to folds and overlaps of the image sheet.

Meanwhile, bijectivity, i.e., simultaneous surjectivity and injectivity, is highly desirable, as it ensures the existence of an inverse mapping, which means that points can be mapped back and forth between two registered images. A bijective homomorphism is called an isomorphism: both the map and its inverse are homomorphisms.

Often, both the map and its inverse should not only be continuous, but also smooth. This leads to a diffeomorphism, i.e., an invertible deformation that maps one image to another, such that both the deformation and its inverse are smooth (differentiable). A diffeomorphic deformation preserves the topology of the image and prevents it from folding.

Different approaches exist, that try to obtain a diffeomorphic deformation field, e.g. [37–42]. Models, which only penalize non-positive determinants of the transform Jacobians, may well guarantee deformations to be free of microfolds locally, but they have no control, however, over macroscopic folding and are not necessarily diffeomorphic. Fluid-like deformation models are shown to prevent folding completely, but require complex computations. The use of exponential maps results in truly diffeomorphic deformations, even for large deformations.

6.2 Exponential map

Diffeomorphisms do not form a vector space with respect to addition: addition of diffeomorphisms does not necessarily produce a diffeomorphism, and the results may not be invertible. Instead, diffeomorphisms can be smoothly composed and inverted [41]. Formally, diffeomorphic deformation fields on the image manifold form a Lie group $\mathcal{G}$. The group element is a smooth manifold, and the associated operation is composition. The neutral element is given by identity I, and an inverse of the diffeomorphism exists by definition.

To this Lie group, a corresponding Lie algebra $\mathfrak{g}$ can be associated. The underlying vector space of this Lie algebra is the tangent space of $\mathcal{G}$ around the neutral element Id. The Lie algebra and the Lie group are linked through the exponential map [41]:

$$\exp : \mathfrak{g} \to \mathcal{G}, \quad \exp(X) = \gamma(1), \tag{52}$$

where $\gamma : \mathbb{R} \rightarrow \mathcal{G}$ is the unique one-parameter subgroup of $\mathcal{G}$ whose tangent vector at identity is given by X. The exponential map can be constructed as the integral curve

$$\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t) = X \circ \gamma(t), \tag{53}$$

where X is a stationary vector field, and with initial condition $\gamma(0) = Id$. Physically interpreted, the diffeomorphic deformation is the movement of particles governed by a stationary velocity field X, after unit time. The exponential map is a smooth mapping from a neighborhood of 0 in $\mathfrak{g}$ to a neighborhood of Id in $\mathcal{G}$. Note that being a one-parameter subgroup implies $\gamma(2t) = \gamma(t) \circ \gamma(t)$, and more generally $\exp(tX) = \gamma(t)$ [43]. In particular, this yields

$$\exp(X) = \exp(2^{-N}X)^{2^N}, \tag{54}$$

which leads to the scaling and squaring algorithm for fast vector field exponential computation [43, 44]:

1. **Scaling:** Divide X by a factor 2^N, so that $2^{-N}X$ is sufficiently close to zero (for example $N = 8$)
2. **Exponentiation:** Perform an explicit first order integration using the Taylor expansion of the exponential, $\exp(2^{-N}X) \approx Id + 2^{-N}X$.
3. **Squaring:** N recursive squarings, $\exp(2^{-n+1}X) = \exp(2^{-n}X) \circ \exp(2^{-n}X)$.

Composition of sampled deformations are obtained using recursive interpolation.

Fig. 4 shows the convergence towards diffeomorphic maps as a function of the scaling and squaring depth N. The difference between warping a mesh with a vector field and its exponential map is illustrated in Fig. 5.

6.3 Diffeomorphic weighting functions

The most obvious way of incorporating exponential maps in diffeomorphic image registration is inspired by the works of Beg and Ashburner [38, 40]. Let $\mathbf{u}$ describe the static velocity field, and the transform be obtained as its exponential map, e.g., $T\mathbf{x} = \exp(\mathbf{u})$. The deformation field is directly guaranteed to be diffeomorphic and invertible – the inverse is obtained as the exponential map of the negative velocity field. But in return, a local change of the velocity field does not only affect the displacement of a single pixel, but all pixels of which the trajectories pass through the support of that change. Consequently, the expression of the metric flow, i.e., the partial derivatives of the weighting function with respect to the velocity field, gets cumbersome [40]. Further, it is suspected that the class of diffeomorphisms that can actually be generated by a static velocity field is quite limited.

Very recently, Vercauteren *et al.* introduced exponential map diffeomorphisms in the Demons framework [41]. There, at each iteration an infinitesimally small update $\mathbf{s}$ is applied through composition of its exponential map with the existing diffeomorphic deformation:

$$T_1 : T_1\mathbf{x} = (\mathbf{x} + \mathbf{u}) \circ \exp(\mathbf{s}), \tag{55}$$

where composition of two diffeomorphisms produces another diffeomorphism, within the limits of discrete numerical computations. At each iteration, the deformation field $\mathbf{u}$ is updated as:

$$\begin{cases} \mathbf{s} \leftarrow \mathbf{0} + dt \cdot \mathrm{d}\mathbf{s} = dt \cdot \mathrm{d}\mathbf{s} \\ \mathbf{x} + \mathbf{u} \leftarrow (\mathbf{x} + \mathbf{u}(\mathbf{x})) \circ \exp(\mathbf{s}) = \mathbf{u}(\mathbf{x}) \circ \exp(\mathbf{s}) + \exp(\mathbf{s}), \end{cases} \tag{56}$$

where dt denotes the time step. This scheme has the advantage of keeping the impact of the update very local. Also, successive composition of diffeomorphisms is commonly believed to generate a wider space of possible deformations, than exponential maps of static velocity fields.

Here, we propose to integrate this deformation scheme into the geodesic active fields model. First, we want to identify the update $\mathbf{s}$ with the minimizing flow $\mathrm{d}\mathbf{u}$ computed in (22). We have:

$$\begin{cases} \exp(\mathbf{s})|_{\mathbf{s}=0} = Id, \\ \frac{\partial}{\partial \mathbf{s}}\Big|_{\mathbf{s}=0} \exp(\mathbf{s}) = Id. \end{cases} \tag{57}$$

Therefore, eq. (56) provides us with the following equality:

$$\mathrm{d}\mathbf{u} = J^T \cdot \mathrm{d}\mathbf{s}, \tag{58}$$

where, once more, J^T refers to the transpose of the transform Jacobian (44).

Given the compositive deformation model T_1, the squared error weighting function becomes:

$$f^{(1)T_1}(\mathbf{x},\mathbf{u}) = \left(M(\mathbf{x}+\mathbf{u}) \circ \exp(\mathbf{s}) - F(\mathbf{x})\right)^2. \tag{59}$$

The spatial gradient of the weighting function, $\nabla f^{(1)T_1}$, is again easily obtained numerically through finite differences. The partial derivatives with respect to the update are:

$$\begin{aligned} f_{\mathbf{s}}^{(1)T_1}(\mathbf{x},\mathbf{u})|_{\mathbf{s}=0} = 2\,&\left(M(\mathbf{x}+\mathbf{u}) \circ \exp(\mathbf{s}) - F(\mathbf{x})\right) \\ &\cdot J^T \nabla M(\mathbf{x}+\mathbf{u}) \cdot \partial_{\mathbf{s}} \exp(\mathbf{s}), \end{aligned} \tag{60}$$

and, more interestingly, we get the metric flow with respect to the deformation field $\mathbf{u}$:

$$f_{\mathbf{u}}^{(1)T_1} = 2\left(M(T_0\mathbf{x}) - F(\mathbf{x})\right) \cdot \nabla M(T_0\mathbf{x}), \tag{61}$$

which is exactly equal to the non-diffeomorphic, additive flow (43).

The diffeomorphic counterparts of the absolute error and local joint entropy weighting function are calculated analogously.

From an implementational point of view, there is only one difference between the additive, non-diffeomorphic T_0, and the compositive diffeomorphic T_1 deformation model: in the latter, the update is obtained as composition with the exponential map of the affinely transformed flow (22):

$$\mathbf{u} \leftarrow (\mathbf{x}+\mathbf{u}(\mathbf{x})) \circ \exp(dt \cdot J^{-T}\mathrm{d}\mathbf{u}) - \mathbf{x}. \tag{62}$$

6.4 Inversibility

The main motivation for diffeomorphic deformations is the existence of a smooth inverse. In this section, we shall compute the inverse deformation along with the regular update of $\mathbf{u}$.

At the end of every iteration $k \geq 0$, the forward transform is composed with the exponential map of the update. It is therefore also natural to compose the inverse of the update with the current inverse transform, written $(\mathbf{x}+\mathbf{p}^k(\mathbf{x}))$. The inverse of the update, in turn, is easily obtained by computing the exponential map of the negative velocity field, and we get:

$$\begin{cases} \mathbf{x}+\mathbf{u}^{k+1} \leftarrow (\mathbf{x}+\mathbf{u}^k(\mathbf{x})) \circ \exp(\mathbf{s}), \\ \mathbf{x}+\mathbf{p}^{k+1} \leftarrow \exp(-\mathbf{s}) \circ (\mathbf{x}+\mathbf{p}^k(\mathbf{x})). \end{cases} \tag{63}$$

The proof using mathematical induction is straightforward. At any iteration $k+1$, inversibility requires:

$$\begin{cases} (\mathbf{x}+\mathbf{u}^k(\mathbf{x})) \circ \exp(\mathbf{s}) \circ \exp(-\mathbf{s}) \circ (\mathbf{x}+\mathbf{p}^k(\mathbf{x})) = Id, \\ \exp(-\mathbf{s}) \circ (\mathbf{x}+\mathbf{p}^k(\mathbf{x})) \circ (\mathbf{x}+\mathbf{u}^k(\mathbf{x})) \circ \exp(\mathbf{s}), \end{cases} \tag{64}$$

which is obviously true, if $(\mathbf{x}+\mathbf{u}^k)^{-1} = (\mathbf{x}+\mathbf{p}^k)$. Indeed, at the first iteration, $\mathbf{u}^0 = \mathbf{0} = \mathbf{p}^0$, and thus $(\mathbf{x}+\mathbf{u}^0)^{-1} = Id = (\mathbf{x}+\mathbf{p}^0)$.

7 Mean Curvature Estimation

In (8), mean curvature is expressed as the anisotropic divergence of the coordinate gradient. Except for the simple stereo case, where an analytical expression of mean curvature was given, explicit expressions are cumbersome. Instead, we propose to estimate the mean curvature vector numerically, by using central differences twice.

The 2D case First we investigate the case of flat and non-flat 2D images. We use the same scheme as was already used by [10]. We include its derivation for the sake of completeness. To simplify the notation, we shorten the expression of the anisotropy tensor:

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} = \sqrt{g} g^{\mu\nu}. \tag{65}$$

The mean curvature component corresponding to u now explicitely writes:

$$H^u = \frac{1}{\sqrt{g}} \left[\partial_x (a \cdot u_x + b \cdot u_y) + \partial_y (b \cdot u_x + c \cdot u_y) \right]. \tag{66}$$

Let the manifold be sampled at integer positions x and y, e.g. u is sampled as $u_{x,y}$. We now apply central differences once to estimate the innermost partial derivatives:

$$\begin{aligned} H^u = \quad & \tfrac{1}{\sqrt{g}} [\quad \partial_x (a_{x,y} \cdot (u_{x+\frac{1}{2},y} - u_{x-\frac{1}{2},y}) \\ & \qquad + \frac{b_{x,y}}{2} \cdot (u_{x,y+1} - u_{x,y-1})) \\ & + \quad \partial_y (\frac{b_{x,y}}{2} \cdot (u_{x+1,y} - u_{x-1,y}) \\ & \qquad + c_{x,y} \cdot (u_{x,y+\frac{1}{2}} - u_{x,y-\frac{1}{2}})) \,]. \end{aligned} \tag{67}$$

Applying central differences a second time allows estimating the remaining derivatives:

$$\begin{aligned} H^u = \quad & \tfrac{1}{\sqrt{g}} [\quad a_{x+\frac{1}{2},y} \cdot (u_{x+1,y} - u_{x,y}) \\ & + \quad a_{x-\frac{1}{2},y} \cdot (u_{x-1,y} - u_{x,y}) \\ & + \quad c_{x,y+\frac{1}{2}} \cdot (u_{x,y+1} - u_{x,y}) \\ & + \quad c_{x,y-\frac{1}{2}} \cdot (u_{x,y-1} - u_{x,y}) \\ & + \quad 1/4 \, b_{x+1,y} \cdot (u_{x+1,y+1} - u_{x+1,y-1}) \\ & - \quad 1/4 \, b_{x-1,y} \cdot (u_{x-1,y+1} - u_{x-1,y-1}) \\ & + \quad 1/4 \, b_{x,y+1} \cdot (u_{x+1,y+1} - u_{x-1,y+1}) \\ & - \quad 1/4 \, b_{x,y-1} \cdot (u_{x+1,y-1} - u_{x-1,y-1}) \,], \end{aligned} \tag{68}$$

where $a_{x+\frac{1}{2},y}$, etc., are simply interpolated as $\frac{a_{x,y}+a_{x+1,y}}{2}$. The expression of H^v is obtained analogously.

The multiscale case The numeric scheme for the mean curvature vector in the multiscale case is obtained in the same manner. First, let us abbreviate the anisotropy tensor:

$$\begin{pmatrix} a & d & f \\ d & b & e \\ f & e & c \end{pmatrix} = \sqrt{g} g^{\mu\nu}. \tag{69}$$

The mean curvature component corresponding to u now explicitely writes:

$$\begin{aligned} H^u = \quad & \tfrac{1}{\sqrt{g}}[\quad \partial_x(a\cdot u_x + d\cdot u_y + f\cdot u_\sigma) \\ & + \quad \partial_y(d\cdot u_x + b\cdot u_y + e\cdot u_\sigma) \\ & + \quad \partial_y(f\cdot u_x + e\cdot u_y + c\cdot u_\sigma)\ ]. \end{aligned} \tag{70}$$

After subsequently replacing all partial derivatives by central differences, we get the following numerical scheme:

$$\begin{aligned} H^u = \quad & \tfrac{1}{\sqrt{g}}[\quad 1/2\,(a_{x+1,y,\sigma} + a_{x,y,\sigma})\cdot(u_{x+1,y,\sigma} - u_{x,y,\sigma}) \\ & + \quad 1/2\,(a_{x-1,y,\sigma} + a_{x,y,\sigma})\cdot(u_{x-1,y,\sigma} - u_{x,y,\sigma}) \\ & + \quad 1/2\,(b_{x,y+1,\sigma} + b_{x,y,\sigma})\cdot(u_{x,y+1,\sigma} - u_{x,y,\sigma}) \\ & + \quad 1/2\,(b_{x,y-1,\sigma} + b_{x,y,\sigma})\cdot(u_{x,y-1,\sigma} - u_{x,y,\sigma}) \\ & + \quad 1/2\,(c_{x,y,\sigma+1} + c_{x,y,\sigma})\cdot(u_{x,y,\sigma+1} - u_{x,y,\sigma}) \\ & + \quad 1/2\,(c_{x,y,\sigma-1} + c_{x,y,\sigma})\cdot(u_{x,y,\sigma-1} - u_{x,y,\sigma}) \\ & + \quad 1/4\,d_{x+1,y,\sigma}\cdot(u_{x+1,y+1,\sigma} - u_{x+1,y-1,\sigma}) \\ & - \quad 1/4\,d_{x-1,y,\sigma}\cdot(u_{x-1,y+1,\sigma} - u_{x-1,y-1,\sigma}) \\ & + \quad 1/4\,f_{x+1,y,\sigma}\cdot(u_{x+1,y,\sigma+1} - u_{x+1,y,\sigma-1}) \\ & - \quad 1/4\,f_{x-1,y,\sigma}\cdot(u_{x-1,y,\sigma+1} - u_{x-1,y,\sigma-1}) \\ & + \quad 1/4\,d_{x,y+1,\sigma}\cdot(u_{x+1,y+1,\sigma} - u_{x-1,y+1,\sigma}) \\ & - \quad 1/4\,d_{x,y-1,\sigma}\cdot(u_{x+1,y-1,\sigma} - u_{x-1,y-1,\sigma}) \\ & + \quad 1/4\,e_{x,y+1,\sigma}\cdot(u_{x,y+1,\sigma+1} - u_{x,y+1,\sigma-1}) \\ & - \quad 1/4\,e_{x,y-1,\sigma}\cdot(u_{x,y-1,\sigma+1} - u_{x,y-1,\sigma-1}) \\ & + \quad 1/4\,f_{x,y,\sigma+1}\cdot(u_{x+1,y,\sigma+1} - u_{x-1,y,\sigma+1}) \\ & - \quad 1/4\,f_{x,y,\sigma-1}\cdot(u_{x+1,y,\sigma-1} - u_{x-1,y,\sigma-1}) \\ & + \quad 1/4\,e_{x,y,\sigma+1}\cdot(u_{x,y+1,\sigma+1} - u_{x,y-1,\sigma+1}) \\ & - \quad 1/4\,e_{x,y,\sigma-1}\cdot(u_{x,y+1,\sigma-1} - u_{x,y-1,\sigma-1})\ ]. \end{aligned} \tag{71}$$

8 Results

We have implemented the geodesic active fields and ran it on several test problems. Here, the results are presented in order of task complexity. As for all forward schemes, the step length, and thus the speed of the registration, is heavily limited by the stability of the integration. The implementation was done using Matlab®(R2009a) on a standard 2.4GHz Intel®Core™2 Duo desktop machine, running a 64bit Fedora Core 11.

8.1 Additive versus diffeomorphic weighting function

A nice illustrative example is the registration of a star to a disc. These two binary images have been created at resolution 350×350. The 2D registration configuration of section 3.3 was implemented. For comparison, both the direct squared error weighting function (42) and the diffeomorphic squared error weighting function (59) are applied. The setup, the warped images and the resulting deformation fields are all shown in Fig. 6. The warped images of both direct and diffeomorphic deformation field are of comparable quality, that could be further improved by registering

at higher resolution deformation fields. For this example, the main difference is the inverse deformation field that is automatically obtained within the diffeomorphic setup. Indeed, we show that composition of the forward transform T_1 with its co-evolving inverse T_1^{-1} produces a map very close to identity, as expected. The inverse transform allows to deform the star closer to the disc.

8.2 A few words on regularization

It might be useful to illustrate the influence of the scaling factor β on the deformation field smoothness. Therefore, a pair of constant images, all ones and all zeros, are registered, where the one-directional deformation field u is initialized with a single local impulse. This allows studying the impulse response of the regularization, f being constant. After a few iterations, the deformation field has diffused, as illustrated in Fig. 7(a). Next, the deformation field is initialized with a unit step, disturbed by uniformly distributed, additive random noise, as shown in Fig. 7(b). In Fig. 7(c)-(e), the role of the parameter β becomes clear: The regularizer changes from Gaussian filtering for low β, to highly anisotropic, feature preserving filtering at higher values.

8.3 Application to stereo vision

An example of stereo vision depth recovery was performed as shown in Fig. 8. The image pair `tsukuba` is a well known test image, taken from the middlebury benchmark set for stereo vision. The registration was set up according to the embedding and evolution equation described in section 3.2 and using the absolute error weighting function (50).

8.4 Application to medical imaging

The third case deals with registration of a highly misaligned monomodal medical image pair. An axial slice through a T1 MRI volume is heavily deformed by a 2D deformation field. The images have a resolution of 317×317 pixels. Registration is set up with the diffeomorphic squared error weighting function. The deformations are almost completely recovered, as illustrated in Fig. 9.

Another case aims at registering a pair of multimodal medical images at resolution 317×317. The first image is the same T1 brain slice as above. The second image now is a deformed slice in T2 weighting. For multimodal registration, we use the information theory based joint local entropy as weighting function (47), again in the diffeomorphic setting using exponential map compositions. At the fine resolution, the resampling of an entire image takes considerable time, as well as the computation of the joint histogram. Accordingly, the whole registration process takes around 3 minutes. Registration is widely successful, as is indicated by the before and after checkerboard and overlay images provided in Fig. 10. Compared to the robust squared error weighting function, the local joint entropy is much more delicate with respect to the initial condition, but allows to register images of different modalities.

8.5 Registration on non-flat manifolds

To illustrate the model on a non-flat manifold, we have implemented the spherical patch described in (33). First, the purpose of the pullback relation is nicely illustrated in Fig. 11. We picture the impulse response that corresponds to the diffusion of a local non-zero spot in the deformation field without data term, i.e., the fixed and moving image are simply set to 0 an 1, thus $f = 1$. On the spherical patch, the impulse response is isotropic and equal both close to the North pole and close to the equator. Isotropy on the spherical manifold requires a high degree of anisotropy in the rectangular parameter domain, as low-θ regions map denser on the sphere. This required anisotropy is directly obtained thanks to the pullback relation between the metrics h_{ij} on the patch and $g_{\mu\nu}$ in the parameter space. Further, the registration has been tested on an artificially deformed pair of topological maps of the Earth, see Fig. 12. The patch

spans a good part of the northern hemisphere and some of the southern hemisphere of a globe, hence covering parts of both Americas, entire Europe, Africa, the Atlantic Ocean and of western Asia. Thus, the registration framework is shown to work on non-flat manifolds, such as the sphere.

8.6 Multiscale image registration

Finally, the multiscale registration case is tested on a pair of artificially deformed T1 brain images. Images are repeatedly lowpass filtered with a Gaussian to generate a linear diffusion scale-space. The multiscale image stacks prior to and after registration are shown in Fig. 13, as well as the corresponding intensity residues and the estimated deformation field. Registration succeeds quite well, as illustrated by the almost entirely removed intensity errors.

9 Conclusion

In this paper, we have presented a novel method called geodesic active fields to register images. The fundamental idea is to embed deformation fields in a higher dimensional space, and to let them evolve towards minimal surfaces, while being attracted by configurations that bring the images into register. In contrast to classic approaches in variational methods, that make use of purely additive competition between data and regularization term, our method combines the terms in a multiplicative way. In fact, the data term is represented by a local image distance function, that acts as multiplicative weighting on the geometric regularization term, resulting in the weighted Polyakov energy.

Weighting functions We have provided three instances of weighting function, namely squared error, absolute error and local joint entropy. On the one hand, the absolute and squared error weighting functions minimize the global L^2 and L^1-norm between the two images, and are suitable for monomodal image registration. The local joint entropy, on the other hand, maximizes the mutual information between images, and lends itself to multimodal image registration.

Diffeomorphisms and invertibility Introducing the concept of exponential maps has enabled us to define deformations that are diffeomorphic, i.e. that keep the warping free of folds. Indeed, at each iteration, the deformation field is updated through composition with a small diffeomorphic field, obtained as the exponential map of the computed flow, keeping the deformation field diffeomorphic. Simultaneously, the exponential map of the reverse flow is composed with the current estimate of the inverse deformation to update the inverse diffeomorphism. In image registration, existence and computation of the inverse deformation is as important as obtaining the forward transform, because this allows mapping points between fixed and moving image in both directions. We have shown within the limits of precision of numerical computations, that composition of the forward and inverse diffeomorphism actually comes close to identity.

The parameters α and β It is important to emphasize the role of the parameters α and β. First, β tunes the aspect ratio between the deformation field dimensions and the spatial dimensions in the embedding. In the simplest case of stereo matching, it has been shown that this allows interpolating between L^2 and L^1-norm minimization of the deformation field gradient magnitudes, whereas interpretation is more difficult in the general case. Second, note that β only changes the nature of the regularization, but not its relative weight with respect to the data term, which is precisely the role of the balancing parameter α.

Familiar terms in the evolution equations In the evolution equations, both the mean curvature vector of the deformation field used for regularization, and the classic distance metric gradient descent terms can be identified. Interestingly, the mean curvature component in the flow is weighted by the image distance function. As a direct result, the regularization force is controlled by the quality of the current alignment: The better the match, the less the regularization is allowed to interfere. This is in contrast to additive competition in classical schemes, where the regularization force always causes a bias off the optimal data position in the end result.

Generality We have derived the minimizing flow of this weighted Polyakov energy for different image registration configurations. First, the framework applies to standard Euclidean images, defined on Cartesian planes and volumes. Further, our proposed method also directly generalizes to images on Riemannian manifolds, such as non-flat image domains and various scale-spaces, and ultimately the combination of both.

In true multiscale registration and in contrast to hierarchic multiresolution approaches, image pairs are registered at all scales simultaneously. Communication between different scales is bidirectionally achieved by the regularization term, smoothing the deformation field across scales. In this context, our model has the advantage over classical approaches of automatically taking the relation between space and scale into account.

Our next efforts will focus on bringing the evolution equations in a suitable form for more efficient numerical implementations, both in terms of speed and stability.

References

[1] D. W. Thompson, *On Growth and Form*. Cambridge: Cambridge Univ. Press, 1917.

[2] M. A. Audette, F. P. Ferrie, and T. M. Peters, "An algorithmic overview of surface registration techniques for medical imaging," *Med. Image Anal.*, vol. 4, no. 3, pp. 201–17, 2000.

[3] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surveys*, vol. 24, no. 4, pp. 325–376, Dec. 1992.

[4] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Med. Image Anal.*, vol. 2, no. 1, pp. 1–36, Mar. 1998.

[5] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.

[6] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford, U.K.: Clarendon Press, 2000.

[7] G. Hermosillo, C. Chefd'hotel, and O. Faugeras, "Variational methods for multimodal image matching," *Int. J. Comput. Vis.*, vol. 50, no. 3, pp. 329–343, 2002.

[8] N. Sochen, R. Kimmel, and R. Malladi, "A general framework for low level vision," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 310–318, Mar. 1998.

[9] A. M. Polyakov, "Quantum geometry of bosonic strings," *Phys. Lett. B*, vol. 103, no. 3, pp. 207–210, 1981.

[10] N. Sochen and R. Ben-Ari, "A geometric framework and a new criterion in optical flow modeling," *J. Math. Imaging Vis.*, vol. 33, no. 2, pp. 178–194, Feb. 2009.

[11] X. Bresson, P. Vandergheynst, and J.-P. Thiran, "Multiscale active contours," *Int. J. Comput. Vis.*, vol. 70, no. 3, pp. 197–211, Dec. 2006.

[12] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *Int. J. Comput. Vis.*, vol. 22, no. 1, pp. 61–79, 1997.

[13] R. Kimmel, N. Sochen, and R. Malladi, "From high energy physics to low level vision," *Scale-Space Theory in Computer Vision*, vol. 1252, pp. 236–247, 1997.

[14] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear Total Variation based Noise Removal Algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.

[15] L. Rudin and S. Osher, "Total Variation based Image Restoration with Free Local Constraints," *Proc. IEEE Int. Conf. Image Process.*, vol. 1, pp. 31–35, 1994.

[16] R. Kimmel, R. Malladi, and N. Sochen, "Images as embedded maps and minimal surfaces: Movies, color, texture, and volumetric medical images," *Int. J. Comput. Vis.*, vol. 39, no. 2, pp. 111–129, 2000.

[17] N. Sochen, R. Deriche, and L. Lopez-Perez, "The beltrami flow over manifolds," INRIA, Tech. Rep. 4897, 2003.

[18] T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[19] I. Bogdanova, X. Bresson, J.-P. Thiran, and P. Vandergheynst, "Scale Space Analysis and Active Contours for Omnidirectional Images," *IEEE Trans. Image Process.*, vol. 16, no. 7, pp. 1888–1901, Jul. 2007.

[20] A. Spira and R. Kimmel, "Geodesic curvature flow on parametric surfaces," in *Proc. Curve and Surface Design*, 2002, pp. 365–373.

[21] B. S. Morse, "Computation of Object Cores from Grey-Level Images," Ph.D. dissertation, University of North Carolina, Chapel Hill, 1994.

[22] D. Marr and E. Hildreth, "Theory of edge detection," *Proc. Roy. Soc. Lon. Series B, Biol. Sci.*, vol. 207, no. 1167, pp. 187–217, Feb. 1980.

[23] J. J. Koenderink, "The structure of images," *Biol. Cybern.*, vol. 50, no. 5, pp. 363–370, 1984.

[24] D. Zosso and J.-P. Thiran, "A scale-space of cortical feature maps," *IEEE Signal Process. Lett.*, vol. 16, no. 10, pp. 873–876, Oct. 2009.

[25] A. P. Witkin, "Scale-space filtering," *Int. Joint Conf. AI*, vol. 2, pp. 1019–1021, 1983.

[26] D. H. Eberly, "A differential geometric approach to anisotropic diffusion," in *Geometry-Driven Diffusion in Computer Vision*, B. ter Haar Romeny, Ed. Kluwer, 1994, pp. 371–392.

[27] ——, "Geometric methods for analysis of ridges in n-dimensional images," Ph.D. dissertation, Chapel Hill, NC, USA, 1994.

[28] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.

[29] A. W. Toga, *Brain Warping*. Academic Press, 1999.

[30] W. M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Med. Image Anal.*, vol. 1, no. 1, pp. 35–52, Mar. 1996.

[31] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," *Int. J. Comput. Vis.*, vol. 24, no. 2, pp. 137–154, 1997.

[32] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, Apr. 1997.

[33] P. Rogelj, S. Kovacic, and J. C. Gee, "Point similarity measures for non-rigid registration of multi-modal data," *Computer Vision and Image Understanding*, vol. 92, no. 1, pp. 112–140, 2003.

[34] P. Thevenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," *IEEE Trans. Image Process.*, vol. 9, no. 12, pp. 2083–2099, Dec. 2000.

[35] D. Mattes, D. Haynor, H. Vesselle, T. Lewellen, and W. Eubank, "PET-CT image registration in the chest using free-form deformations," *IEEE Trans. Med. Imag.*, vol. 22, no. 1, pp. 120–128, Jan. 2003.

[36] C. Zach, T. Pock, and H. Bischof, "A duality based approach for realtime TV-L1 optical flow," in *Pattern Recognition*, ser. Lecture Notes in Computer Science, vol. 4713. Berlin, Heidelberg: Springer, 2007, pp. 214–223.

[37] G. Christensen, R. Rabbitt, and M. Miller, "Deformable templates using large deformation kinematics," *IEEE Trans. Image Process.*, vol. 5, no. 10, pp. 1435–1447, Oct. 1996.

[38] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *Int. J. Comput. Vis.*, vol. 61, no. 2, pp. 139–157, Feb. 2005.

[39] D. Rueckert, P. Aljabar, R. A. Heckemann, J. V. Hajnal, and A. Hammers, "Diffeomorphic registration using b-splines," in *MICCAI*, ser. Lecture Notes in Computer Science, vol. 4191. Berlin, Heidelberg: Springer, 2006, pp. 702–709.

[40] J. Ashburner, "A fast diffeomorphic image registration algorithm," *Neuroimage*, vol. 38, no. 1, pp. 95–113, Oct. 2007.

[41] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," *Neuroimage*, vol. 45, no. 1, Supplement 1, pp. S61–S72, Mar. 2009.

[42] B. Yeo, M. Sabuncu, T. Vercauteren, N. Ayache, B. Fischl, and P. Golland, "Spherical demons: Fast surface registration," in *MICCAI*, 2008, pp. 745–753.

[43] M. Bossa, E. Zacur, and S. Olmos, "Algorithms for computing the group exponential of diffeomorphisms: Performance evaluation," in *CVPRW*. IEEE, Jun. 2008, pp. 1–8.

[44] V. Arsigny, O. Commowick, X. Pennec, and N. Ayache, "A log-euclidean framework for statistics on diffeomorphisms," in *MICCAI*, 2006, pp. 924–931.

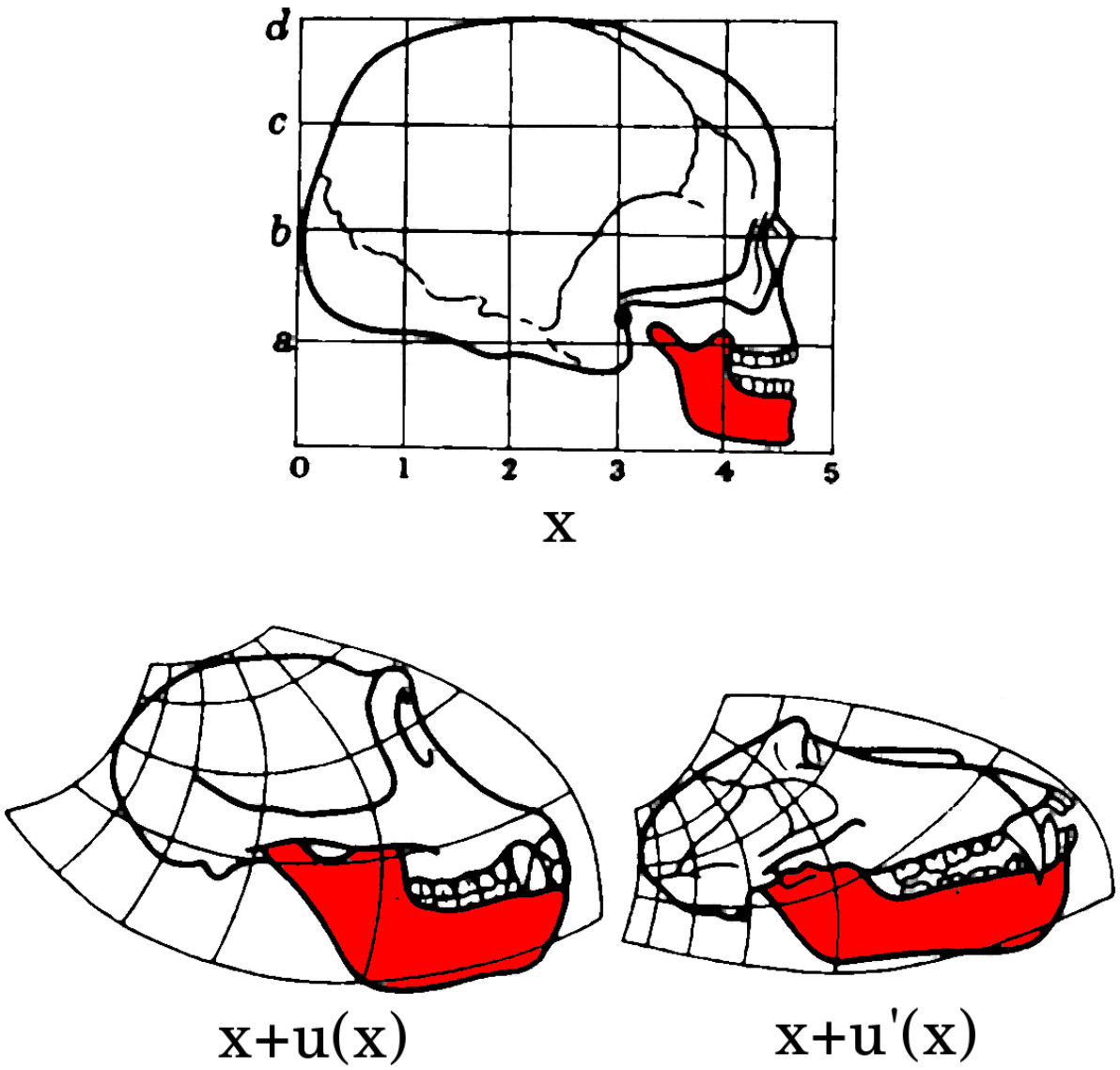
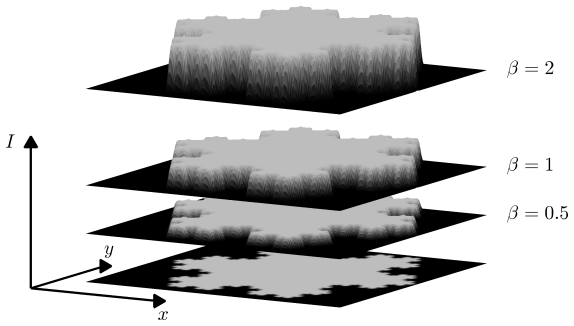
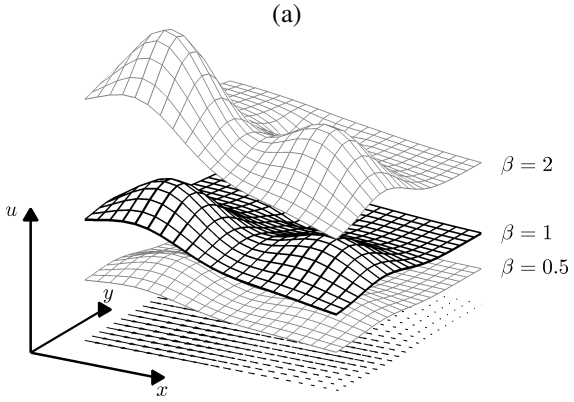


Figure 1: *Image registration is the concept of mapping homologous points of different images*. The skull of a human is registered to shimpanzee and baboon by finding the deformation fields $\mathbf{u}(\mathbf{x})$ and $\mathbf{u}'(\mathbf{x})$, such that human features, e.g. the mandible (red), at $\mathbf{x}$ match those of shimpanzee and baboon at $\mathbf{x}+\mathbf{u}(\mathbf{x})$ and $\mathbf{x}+\mathbf{u}'(\mathbf{x})$. Skull sketches reproduced from [1]



(a)



(b)

Figure 2: Beltrami embedding of images and deformation fields in higher dimensional spaces. (a) A graylevel image is embedded in 3D according to $X : (x,y) \rightarrow (x,y,I)$. The apparent effect of the aspect ratio β is illustrated as a relative scaling of the surface variations with respect to the spatial dimensions x and y. Minimizing the Polyakov energy smoothens the image, where β interpolates between total variation and Gaussian smoothing. (b) The scalar deformation field (lateral shift along x) of a planar image is embedded in $\mathbb{R}^3$. Formally, the embedding writes $X : (x,y) \rightarrow (x,y,u)$. The arrows illustrate the corresponding deformation field in the image plane. The Polyakov energy measures the area of the embedded surface, and is a measure of regularity of the deformation field. Weighting of the energy allows to drive the minimization towards the optimal registration result.

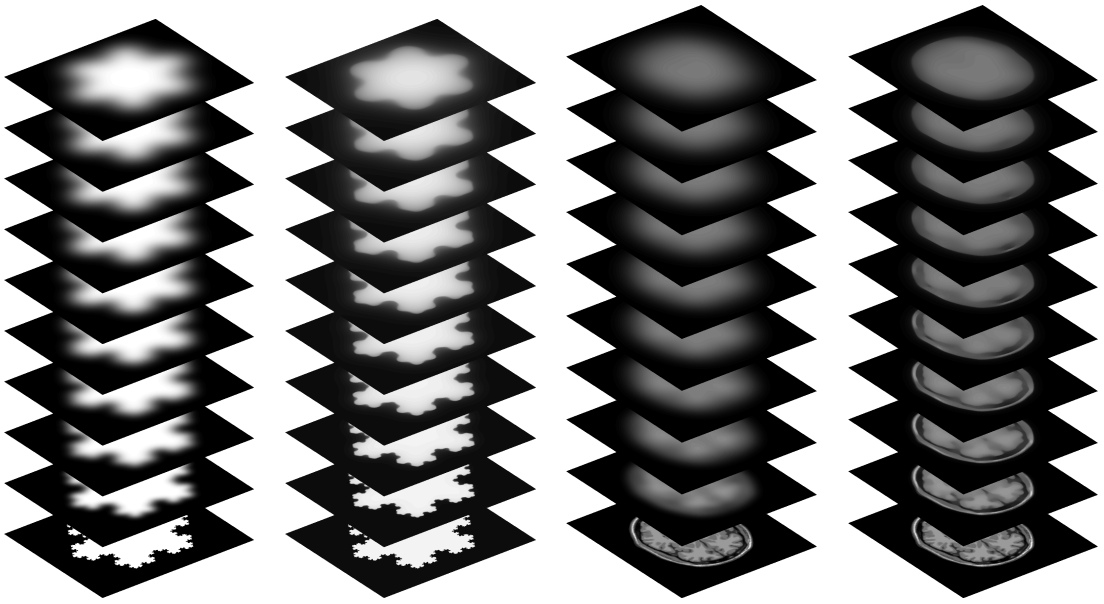


Figure 3: Different scale-spaces and multiscale images. (a)–(b) Linear and Beltrami scale-spaces of the Von Koch snowflake. The scale σ increases linearily from bottom to top of the image stack, thus constituting an additional image dimension. (c)–(d) Multiscale representation of a T1-weighted magnetic resonance image of a human brain, in a linear and Beltrami scale-space.

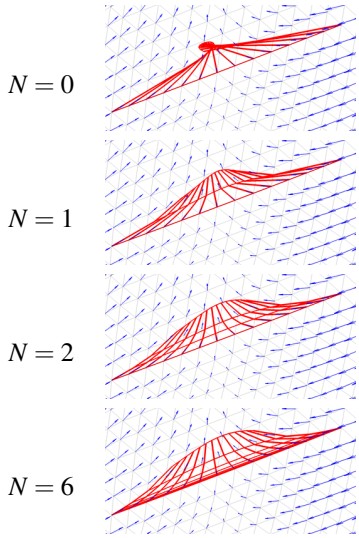


Figure 4: *Convergence of exponential maps with increasing scaling and squaring depths N*. The blue arrows indicate the velocity field. The thick red lines indicate the individual trajectories of a set of vertices during squaring, while the thin red lines show the front propagation. $N = 0$ corresponds to the basic, non-diffeomorphic model and can introduce heavy folding. Indeed, the whole set of points is mangled onto a very small region. For N increasing, the exponential map approximation becomes rapidly more accurate.

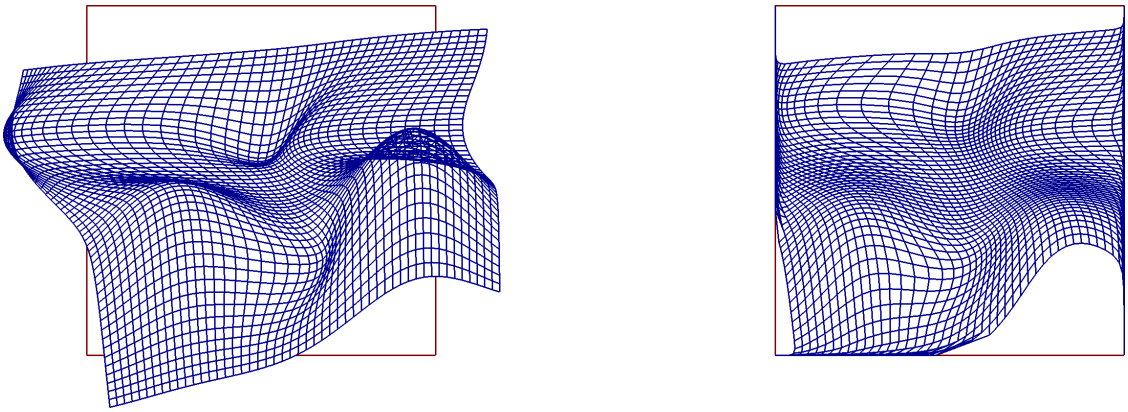
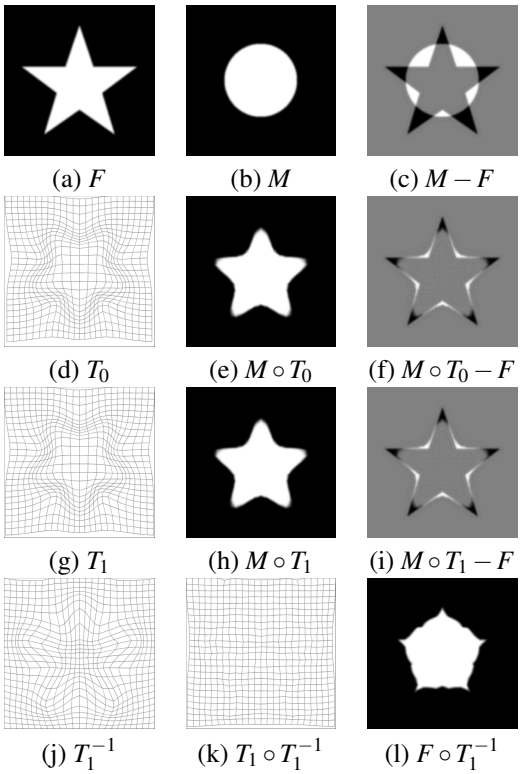


Figure 5: *A square mesh warped by a vector field (left) and its exponential map (right)*. The velocity field is considered zero outside the square, therefore the mesh cannot flow beyond that border. The exponential map produces a diffeomorphic deformation of the mesh, whereas direct deformation by the vector field introduces folding.



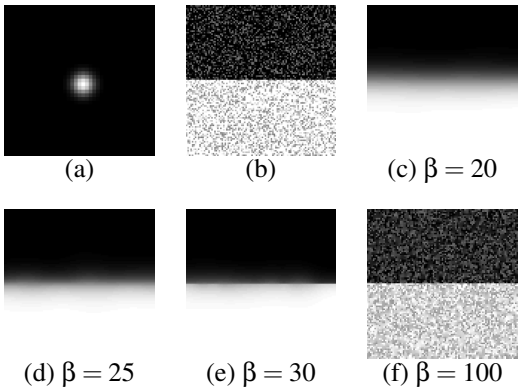
(a) F (b) M (c) $M - F$

(d) T_0 (e) $M \circ T_0$ (f) $M \circ T_0 - F$

(g) T_1 (h) $M \circ T_1$ (i) $M \circ T_1 - F$

(j) T_1^{-1} (k) $T_1 \circ T_1^{-1}$ (l) $F \circ T_1^{-1}$

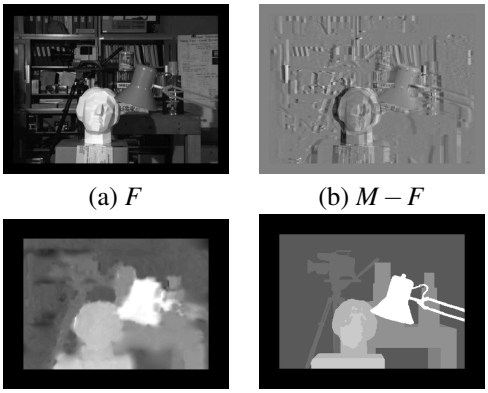
Figure 6: Registration of a star to a disc. (a)-(c) The fixed star and the moving disc, and their difference. (d)-(f) The computed non-diffeomorphic deformation field, T_0, the warped disc and the remaining intensity difference. (g)-(i) Deformation field T_1, warped disc and remaining intensity difference in the diffeomorphic setting. (j) Inverse transformation T_1^{-1} in the diffeomorphic setup. (k) Composition of the forward and backward transform is very close to identity:$T_1 \circ T_1^{-1} \approx Id$. (l) The star warped with the inverse transform maps close to the disc. Both the non-diffeomorphic and diffeomorphic deformation are free of folds. The precision of the registration, i.e. the remaining shape differences, are mainly due to the limited resolution we chose for the deformation field.



(a) (b) (c) $\beta = 20$

(d) $\beta = 25$ (e) $\beta = 30$ (f) $\beta = 100$

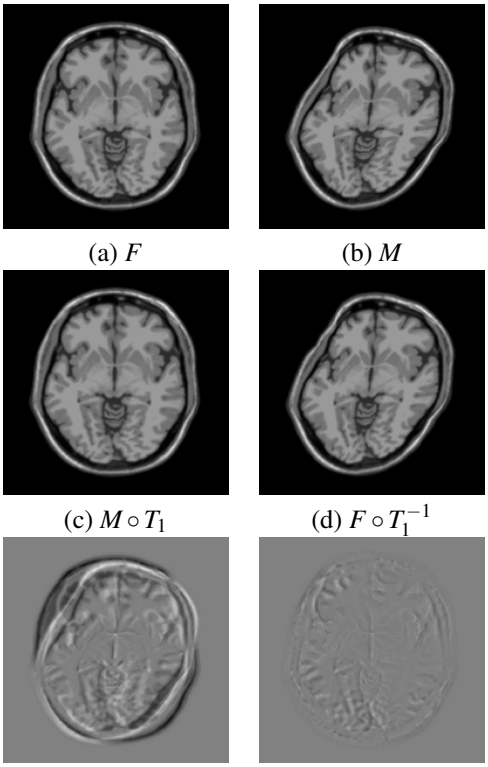
Figure 7: Impulse response and "denoising" of the Laplace-Beltrami regularization under constant data term $f = 1$. (a) Mid-time response, i.e. after some 100 iterations, to a single impulse in the deformation field. (b) Initial unit-step deformation field with uniform additive noise, and (c)-(f) its smoothing by the regularizer at different β.



(a) F (b) $M - F$

(c) (d)

Figure 8: Stereo vision through disparity recovery. (a)-(b) The `tsukuba` test image and the image pair difference. (c) The recovered disparity map, (d) Ground truth.

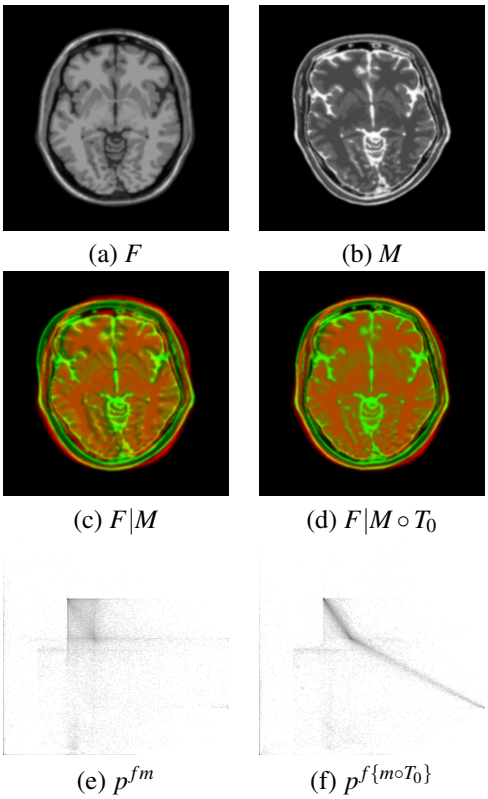


(a) F (b) M

(c) $M \circ T_1$ (d) $F \circ T_1^{-1}$

(e) $M - F$ (f) $M \circ T_1 - F$

Figure 9: Monomodal registration of brain MRI. (a)-(b) Fix and moving image. (c) The moving image warped by the recovered deformation field T_1. (d) The fix image warped with the inverse deformation field T_1^{-1}. (e)-(f) Intensity differences before and after registration.



(a) F (b) M

(c) $F|M$ (d) $F|M \circ T_0$

(e) p^{fm} (f) $p^{f\{m \circ T_0\}}$

Figure 10: Multimodal registration of T1 and T2 brain MRI. (a) The fix T1 image. (b) The artificially deformed T2 weighted image. (c)-(d) Red-green overlay of T1 (red) and T2 (green) images prior to and after registration. (e)-(f) The joint intensity histograms prior to and after registration. While initially, the histogram is widely spread, registration results in important histogram focalization.

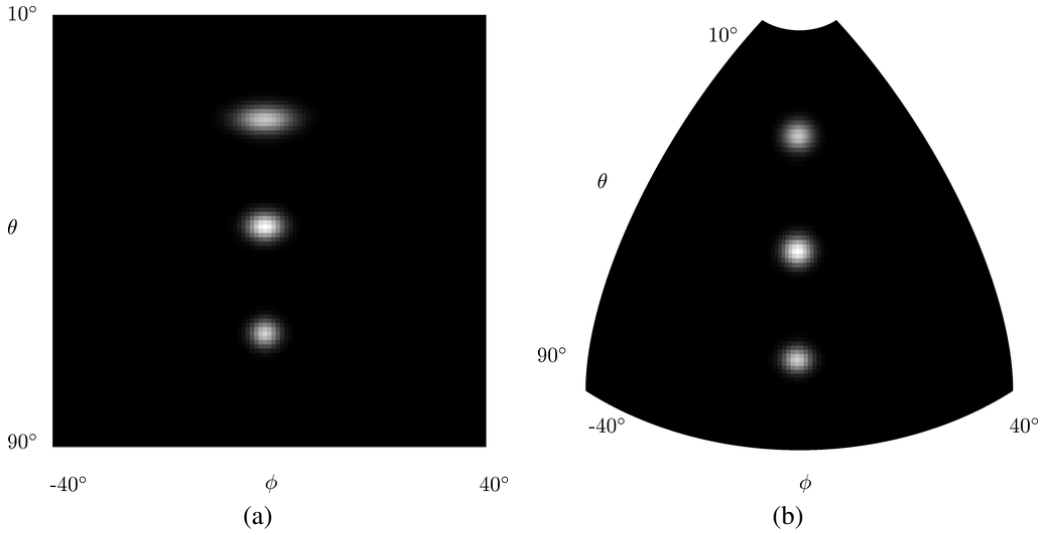


Figure 11: Regularization impulse response on a non-flat manifold. Three isolated non-zero spikes in the deformation field diffuse without data term, i.e. $f = 1$. (a) Diffusion takes place in parameter domain (θ, ϕ), governed by the metric tensor $g_{\mu\nu}$. At low θ, diffusion is highly anisotropic. (b) Diffusion as seen on the embedded spherical patch. Both impulse responses look the same and are isotropic. This is obtained through the pullback relation that links the respective metrics h_{ij} on the patch and $g_{\mu\nu}$ in the parameter space.

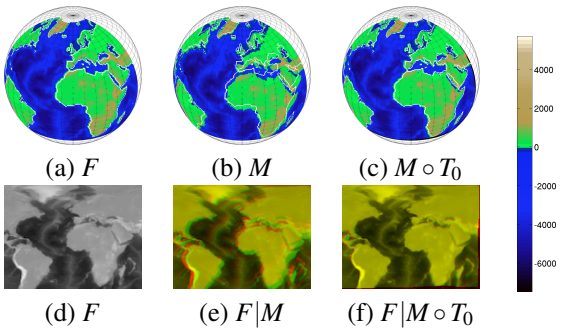


Figure 12: Image registration on a non-flat manifold. A pair of two artificially deformed topological maps of the Earth are registered. (a) The fixed map with its approximate coastlines highlighted as white contours. (b) The moving image, with the coastlines of the fixed image superimposed. (c) The warped moving image after registration. The map fits well with the superimposed fixed coastlines, except at a few locations where small misregistration is observed (e.g. Red Sea). The colormap indicates height in meters above (below) sealevel. (d) The fixed image in rectangular, flat parameter domain. (e)-(f) The red-green overlay before and after registration.

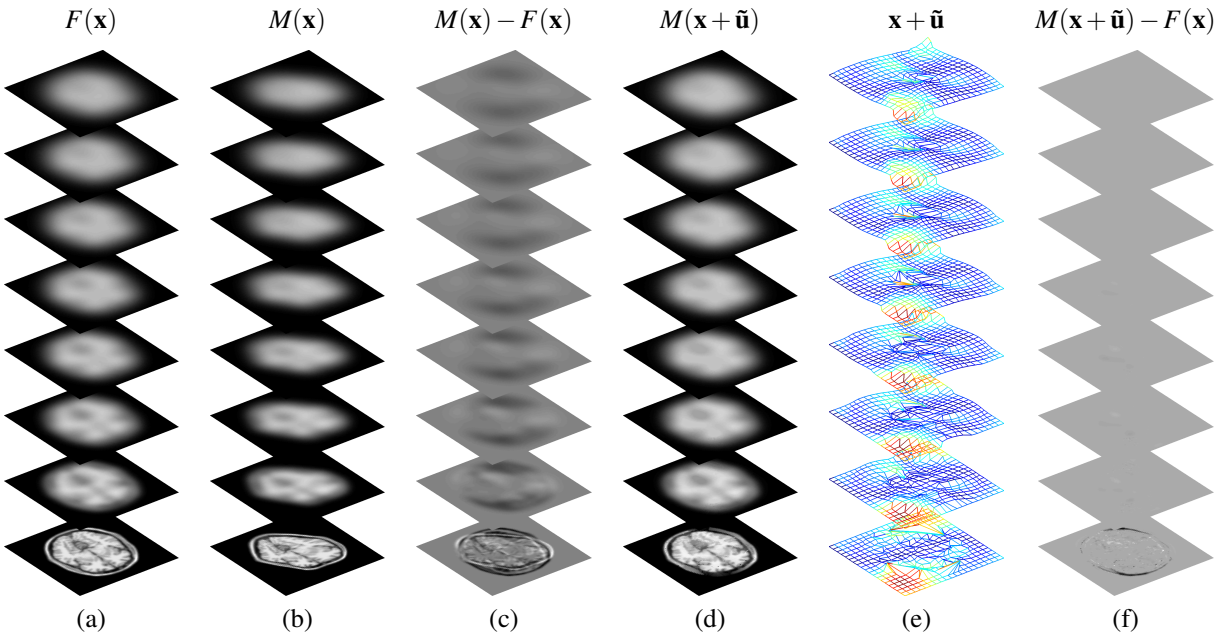


Figure 13: Multiscale registration of a T1 image pair. (a) The multiscale stack of the fixed image of a T1 brain slice. (b) The synthetically deformed moving image of a T1 brain slice. (c) The intensity difference illustrates the misalignment at all scales. (d) The registered moving image. (e) Mesh illustrating the deformation field at all scales. The heat map indicates error from ground truth. (f) The residue after registration is significantly reduced. Some misregistration is observed at the frontal parts of the skull.