A Review of Geometric Transformations for Nonrigid Body Registration

Mark Holden

***Abstract*—This paper provides a comprehensive and quantitative review of spatial transformations models for nonrigid image registration. It explains the theoretical foundation of the models and classifies them according to this basis. This results in two categories, physically based models described by partial differential equations of continuum mechanics (e.g., linear elasticity and fluid flow) and basis function expansions derived from interpolation and approximation theory (e.g., radial basis functions, *B*-splines and wavelets). Recent work on constraining the transformation so that it preserves the topology or is diffeomorphic is also described. The final section reviews some recent evaluation studies. The paper concludes by explaining under what conditions a particular transformation model is appropriate.**

***Index Terms*—Fluid flow registration, linear elastic registration, nonrigid image registration, parametric transformation models, spatial transformations, spline based registration, wavelet based registration.**

Mathematical Notation

$\mathbf{A}(\mathbf{x})$	Target image, $\mathbf{x} \in \mathbf{\Omega}$ is an image location.
$\mathbf{B}(\mathbf{x})$	Source image to be aligned with $\mathbf{A}(\mathbf{x})$.
D_{ij}	The rate of deformation tensor.
ϵ_{ij}	Strain tensor.
η	Mass source term.
$\mathbf{f}(\mathbf{x})$	Body force per unit volume acting at $\mathbf{x}$.
i, j	Cartesian components.
$\mathbf{I}$	Identity matrix.
$\langle f(x), g(x) \rangle$	Inner product of functions $f(x)$ and $g(x)$, i.e., $\int \bar{f}(x) g(x) dx$ where $\bar{f}(x)$ is the complex conjugate of $f(x)$.
λ, μ	Lamé constants describing the mechanical properties of an elastic material.
λ_f, μ_f	Viscosity coefficients of a viscous fluid.
∇	Vector differential operator: $\hat{\mathbf{i}}(\partial/\partial x) + \hat{\mathbf{j}}(\partial/\partial y) + \hat{\mathbf{k}}(\partial/\partial z)$ for $\mathbb{R}^3$.
$\mathbf{\Omega}$	Domain of the image.
$\mathbf{p}_i, \mathbf{q}_i$	ith landmark locations in the source and target images, respectively.
ρ	Fluid density.
σ_{ij}	Cauchy stress tensor.
$\mathbf{T}$	Spatial transformation, refers to the mapping from the space of the source image to the space of the target.
$\oplus$	Direct sum.
$\mathbf{u}(\mathbf{x})$	Displacement vector of point $\mathbf{x}$ in the space of the source image with components (u, v, w).
$\mathbf{v}(\mathbf{x})$	Velocity vector.
V_{ij}	Vorticity tensor.

Definitions

Capture range	The range of attraction of the registration function. The spatial extent of the set of misregistrations from which convergence to the optimal transformation is possible.
CSF	Cerebrospinal fluid.
Diffeomorphism	A differentiable homeomorphism with a differentiable inverse (nonzero determinant of the Jacobian matrix). It is a homeomorphism that maps one differentiable manifold to another.
Eulerian coordinates	These describe the motion of a body of particles relative to a set of fixed points in space $\mathbf{x}$ through which particles pass. The Eulerian coordinate frame refers to the current state of the system.
Homeomorphism	A continuous bijective mapping with a continuous inverse. Intuitively this is achieved by stretching, bending or compressing an elastic material without any cutting.
Lagrangian coordinates	These describe the motion of a body of particles relative to its initial configuration. Given the position of a particle $\mathbf{X}$ at time $t = 0$, its position $\mathbf{x}$ at time t is given by a mapping from the initial to current configuration, i.e., $\mathbf{x} = \mathbf{u}(\mathbf{X}, t) + \mathbf{X}$ where $\mathbf{u}(\mathbf{X}, t)$ is the displacement vector. Since deformations are assumed to be homeomorphisms there exists a unique mapping $\mathbf{X} = \mathbf{x} - \mathbf{u}(\mathbf{x}, t)$ of the current location $\mathbf{x}$ of a particle to its original one $\mathbf{X}$ at $t = 0$. Lagrangian coordinates are also referred to as material or referential coordinates.

Manuscript received January 16, 2007; revised June 22, 2007.

The author is with the CSIRO-ICT Centre, North Ryde, New South Wales, NSW 2113, Australia (e-mail: email@research.mholden.org).

Digital Object Identifier 10.1109/TMI.2007.904691

Positive definite function	A function $f : \mathbb{R} \mapsto \mathbb{C}$ is positive definite if the associated matrix $\mathbf{A}$ with elements $a_{ij} = f(x_i - x_j)$ is positive semi-definite $\forall\, x_i, x_j \in \mathbb{R}$. A matrix $\mathbf{A}$ is positive semi-definite if $\mathbf{z}^*\mathbf{A}\mathbf{z} \geq 0$ for all vectors $\mathbf{z} \in \mathbb{C}$.
Support of a function	Let $f(x)$ be a real-valued function on some set X. The support of f is the smallest closed subset $Y \subseteq X$ outside of which f is zero, i.e., $f(x) = 0\ \forall\, x \in X \cap \bar{Y}$. A function is said to have compact support if its support region is a compact subset of X and global support if $Y = X$.
EBS	Elastic body spline. The solution of the Navier-Cauchy PDE of linear elasticity.
FFD	Free-form deformation.
GEBS	Gaussian elastic body spline.
GM	Grey matter.
MQ	A multiquadric is a type of radial basis function of form $\sqrt{r^2 + c^2}$.
MRA	Multiresolution analysis.
OF	Optical flow based transformation.
PA	Piecewise affine transformation.
PDE	Partial differential equation.
ROI	Region of interest.
TPS	Thin-plate spline.
VS	Volume spline.
WM	White matter.
WMN	Weighted mean radial basis function.

I. Introduction

IMAGE registration is the process of determining the correspondence between objects in two images, by convention between the source and the target image. To determine correspondences it is necessary to find the geometrical or spatial mapping (or spatial transformation) applied to the source image so that it aligns with the target. The mapping is from the image domain Ω to a subset or a superset of Ω. For medical imaging, the mapping is usually 3-D to 3-D. Transformations that preserve the distance between all points in the image are referred to as rigid-body transformations. They are equivalent to a change from one Cartesian system of coordinates to another one which differs by shift and rotation. Transformations that allow for a global change of scale and shear are referred to as affine transformations. Affine transformations map parallel lines to parallel lines. Affine and rigid-body transformations can be conveniently represented using homogeneous matrices, these are 4×4 matrices for 3-D to 3-D mappings.

In contrast, nonrigid transformations map straight lines to curves. Nonrigid registration is the process of determining such transformations given two images of an object. In certain situations the deformation model is known, e.g., the geometrical distortion of the imaging system, but in most cases it is unknown. There are many different nonrigid transformation models. In general, they can be divided into two categories: physical based models and function representations. The physical models in general, are derived from the theory of continuum mechanics and can be divided into two main subcategories: elasticity and fluid flow. Function representations originate from interpolation and approximation theory. They use basis function expansions to model the deformation. There are many different types of basis functions, e.g., radial basis functions, *B*-splines and wavelets.

There are a few reviews of rigid-body registration methods [1]–[3] and a few that consider nonrigid transformations [4]–[8]. Also there are reviews of related image warping methods [9]. Lester *et al.* [4] reviewed a number of transformation models, including linear elasticity, fluid flow, function expansions, and splines. They focussed on hierarchical strategies which can be applied to both the transformation and data. Rohr [5] focussed on landmark based methods, particularly the TPS model and extensions to it. Zitova and Flusser [6] provided a general review of image registration which includes a section on transformation models. Their review describes radial basis functions, elastic and fluid models. Modersitzki [7] concentrated on numerical solutions to registration problems, including nonrigid ones. This included elastic and fluid models and also radial basis functions. Goshtasby [8, Ch. 5] focussed on radial basis functions and compared them to piecewise affine models. None of these provide a comprehensive review of commonly used transformations such as *B*-splines and wavelets and few of them explain the underlying physical models or discuss the comparative evaluation of transformation models. This paper provides a comprehensive quantitative review including an explanation of theoretical basis of the models. It compares models and describes their limitations.

This paper is organized by grouping transformations according to their theoretical basis. This results in two main categories: those that originate from physical models of materials and those that originate from interpolation and approximation theory. In addition to this, there are methods that constrain the transformation according to some desirable mathematical property. Accordingly the paper is organized with the following structure:

- Physical models
  - — linear elasticity
  - — viscous fluid flow
  - — optical flow.
- Basis function expansions
  - — radial basis functions
  - — *B*-splines
  - — wavelets.
- Constraints on the transformation
  - — inverse consistency
  - — topology preservation
  - — diffeomorphic transformations.

The final section describes recent comparative evaluation studies of some of these models. In order to make the paper more self-contained and provide motivation, a brief description of registration metrics and applications of nonrigid registration follows on from the introduction.

A. Registration Metrics

A registration metric takes two images as input and returns a real value that indicates how well the images are aligned. One of the simplest ones is based on the distance between corresponding pairs of landmarks that are extracted from images. The landmarks can be anatomical features or fiducial markers that are rigidly attached to bone. For rigid-body registration the theory is well developed [1], [2], [10]–[12]. For nonrigid registration, landmarks are often used with thin-plate splines, see Section V-D. An advantage of landmarks is that they enable the transformation to be determined in closed form. Disadvantages are that a large number of them are needed to densely sample the deformation field and also the localization process introduces error. Another possibility is to use the distance between corresponding segmented surfaces. But this provides a registration metric only at the surfaces and not throughout the image volume as is often required, also the segmentation process introduces error. A more modern approach is to use an image similarity measure which gives a quantitative measure of image alignment. For intramodality registration, the sum of square differences or cross correlation of the corresponding voxel intensities can be used. For intermodality registration measures based on information theory such as mutual information perform well in the rigid case, see [3]. These measures have the advantage of providing fully automatic algorithms and are suitable for determining dense deformation fields.

II. Applications of Nonrigid Registration

In general, application domains are: medical imaging, remote sensing, and industrial imaging [8]. Crum *et al.* [13] survey a variety of medical applications. Essentially, there are two categories: intrasubject and intersubject. Intrasubject registration refers to the registration of scans of the same subject at different times, while intersubject refers to the registration of scans of different subjects, usually with the same imaging modality. A nonexhaustive list of the main applications is given below.

- Motion correction: To correct for the deformation of a patient's anatomy over time. For example, to correct for motion between preoperative, intraoperative, and postoperative scans in neurosurgery. For instance to correct for brain-shift [14], [15] and facilitate navigation. These deformations are mostly physical in nature and are caused by: changes in the direction of gravity, changes in fluid pressure, physiological motion associated with the heart, respiration, peristalsis, or other muscle groups.
- Motion determination: To quantify the physiological motion of the organ, e.g., heart [16], [17], lungs [18], [19], or joints [20] and use the measurements for diagnosis or therapy monitoring.
- Cross modality image fusion: To combine information from multiple scans of the same patient with different imaging modalities e.g., X-ray-MR [21], PET-CT [22], [23]. This is analogous to the rigid-body case except that the tissue is deformable, the deformations involved are similar to motion correction.
- Change detection: To detect and measure structural change over time. For instance for monitoring disease processes (e.g., longitudinal studies) to aid either diagnosis or therapy. Typically measures of volume and shape are used that are derived from the transformations. Examples include multiple sclerosis [24], rheumatoid arthritis [25], Alzheimer's disease [26], [27], hormone therapy [28], and morphological changes resulting from surgical intervention [29].
- Distortion correction: To measure and correct for geometrical distortion of the imaging system. Possible approaches are to register to scans of other imaging modalities that exhibit less distortion [30], [31] or to use phantoms [32].
- Atlas construction: To produce a representation of the average or variation in anatomy for a patient group. Atlases can be either probabilistic [33]–[35], intensity based [36], label based [37], [38], or deformation based [39].
- Atlas registration: This allows information from a group of subjects to be combined and analysed in the standard space of the atlas, cf. Talairach space [40].
- Segmentation: Given an image containing a set of delineated structures this can be registered to the subject images and the transformations used to propagate the delineations into the space of the subject images so providing a segmentation [41]–[43]. Accurate segmentations of tissue can be obtained from optical images of histologically stained tissue samples. These segmentations can be mapped into anatomical images like MR [44], [45].

III. Linear Elastic Transformations

A. Theory

The theory of linear elasticity is based on notions of stress and strain. The stress at a given location is the contact force per unit area acting on orthogonal planes that intersect the location. Stress can be analysed mathematically using the Cauchy stress tensor. This is a second rank tensor denoted by σ_{ij}, the subscripts i and j denote the three Cartesian directions (x, y, and z). Stress components are either normal to the plane σ_{ii} or within it $\sigma_{ij}, i \neq j$. This tensor has nine components and can be represented as a 3×3 matrix. Strain is a measure of the amount of deformation. It is treated in an analogous way to stress as a second rank tensor ϵ_{ij} with normal ϵ_{ii} and shear $\epsilon_{ij}, i \neq j$ components.

When a body is subject to an external force this induces internal forces within the body which cause it to deform. The internal forces are grouped into body and surface forces. Body forces are distributed throughout the volume and are specified as force per unit volume. When a linear elastic material is in an equilibrium state the body forces $\mathbf{f}$ balance with the surface stresses σ_{ij}. So the integral of the surface (stress) forces and body forces must be zero. Assuming that the stress components vary linearly across an infinitesimal element it is possible to determine the following set of equilibrium equations [46]:

$$\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{zx}}{\partial z} + f_x = 0 \quad (x, y, z) \tag{1}$$

where (x, y, z) indicates that the other two equations are obtainable through cyclic permutation of x, y, and z. By applying Gauss's divergence theorem to the force integral it can be shown that the stress tensor σ_{ij} is symmetric, this

reduces the number of independent stress components to six $(\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{xy}, \sigma_{yz}, \sigma_{zx})$. The normal and shear infinitesimal strain can then be expressed in terms of the spatial derivative of the displacement, as follows:

$$\epsilon_{xx} = \frac{\partial u}{\partial x}$$
$$\epsilon_{xy} = \frac{1}{2}\left[\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right] \quad (x, y, z; u, v, w). \tag{2}$$

The constitutive equations for elasticity relate stress and strain tensors see [46], [47]. This relationship is expressed in the generalised Hooke's law $\sigma_{ij} = C_{ijkm}\epsilon_{km}$. The quantity C_{ijkm} is a fourth rank tensor referred to as the stiffness tensor. Since there are six independent components for both the stress and strain tensors the tensor of elastic constants C_{ijkm} has 36 distinct elastic constants. For an homogeneous isotropic material there are an infinite number of planes of symmetry. Hence, the constitutive equations are independent of the coordinate system. By considering rotational invariance it is possible to reduce the number of independent constants to just two [46]. These are the Lamé constants, λ and μ. μ is also referred to as the shear modulus. Thus, for an isotropic material the stress-strain relation simplifies to the following Piola-Kirchoff form:

$$\sigma_{ij} = \lambda\delta_{ij}\epsilon_{kk} + 2\mu\epsilon_{ij}. \tag{3}$$

Equations (1)–(3) form a system of 15 equations with 15 unknowns (stress, strain, and displacement) and so the unknowns can be determined. Substituting (2) into (3) and then substituting the result into (1) gives the Navier-Cauchy linear elastic PDE

$$\mu\nabla^2\mathbf{u}(\mathbf{x}) + (\mu + \lambda)\nabla(\nabla.\mathbf{u}(\mathbf{x})) + \mathbf{f}(\mathbf{x}) = 0 \tag{4}$$

where $\mathbf{u}(\mathbf{x})$ is the displacement vector at position $\mathbf{x}$, and $\mathbf{f}(\mathbf{x})$ denotes the body force per unit volume, which drives registration. It is possible to determine eigenfunctions of (4). These are products of univariate sinusoidal functions [48], [7, Ch. 9] and have the following form

$$\boldsymbol{\phi} = \begin{pmatrix} \sin(a\pi x)\cos(b\pi y)\cos(c\pi z) \\ \cos(a\pi x)\sin(b\pi y)\,cos(c\pi z) \\ \cos(a\pi x)\cos(b\pi y)\sin(c\pi z) \end{pmatrix}. \tag{5}$$

The corresponding eigenvalues are given by

$$\lambda_{abc,j} = \pi^2(a^2 + b^2 + c^2)\begin{cases} (2\mu + \lambda), & j = 1 \\ \mu, & j = 2, 3. \end{cases} \tag{6}$$

The second-order terms of the displacement gradient are ignored in (2). This leads to error for large deformations. Furthermore, many biological materials have a nonlinear stress-strain relationship which also leads to error. Consequently, the linear model (4) is only really accurate for small deformations.

B. Linear Elastic Algorithms

The Navier-Cauchy PDE (4) is essentially an optimisation problem that involves balancing the external forces (image similarity) with the internal stresses that impose smoothness constraints [4]. It can be solved using variational [7], finite difference [49], [50], FEM models [51], basis function expansion [48], and Fourier transform methods [7].

Broit [49] was the first to propose a linear elastic model for nonrigid image registration. In [49, Ch. 6], an iterative algorithm is described that determines $\mathbf{u}$ for which the internal stresses and external forces of (4) are in equilibrium. The PDE is solved by the finite difference method on a rectangular lattice. The first and second derivatives, $(\partial u/\partial x)$ and $(\partial^2 u/\partial x^2)$, are approximated using discrete derivatives. This results in three linear equations, one for each Cartesian direction (i.e., f_i, f_j, f_k). These linear equations can be solved iteratively from the initial and previously calculated displacements determined using Gauss-Seidel or Jacobi methods. This gives a value of $\mathbf{u}$ for each lattice point.

Bajcsy *et al.* [50] improved this approach. Prior to elastic registration they corrected for global differences using a transformation consisting of translation, rotation and scaling. This was determined by aligning the centres of mass, ellipsoid axes, etc. They used a multiresolution version [50], [52] of Broit's [49] elastic model. The external force was based on the cross correlation of image features. These consisted of the local mean intensity, horizontal, and vertical edges that were extracted from the images.

Davatzikos and Bryan [53] designed an elastic algorithm for intersubject registration of cortical grey matter. They modelled the brain cortex as a thin spherical shell of constant thickness and described the central layer parametrically by $\boldsymbol{\alpha}(u, v)$, with surface parameters u and v and a surface normal $\mathbf{N}(u, v)$. Deformation was modelled as a uniform dilation or contraction with bending (homothetic mapping) between the two surfaces. Their model is also based on a balance between internal and external forces. The external force has two components, the first one deforms a point $\mathbf{x}$ towards the shell. This attractive force is simply the distance between the point $\mathbf{x}$ and the centre of mass function $\mathbf{c}(\mathbf{x})$, i.e., $\mathbf{f}(\mathbf{x}) = \mathbf{c}(\mathbf{x}) - \mathbf{x}$. The second external force acts normally to the shell's surface and has magnitude γ, this either expands or dilates the shell depending on its sign. This leads to the following PDE:

$$\beta(u, v)\left[\mathbf{x}_{uu}(u, v) + \mathbf{x}_{vv}(u, v)\right] + \left[1 - \gamma(u, v)\right] \times \{\mathbf{c}(\mathbf{x}) - \mathbf{x}\} + \gamma(u, v)\mathbf{N}(u, v) = 0 \tag{7}$$

where the repeated subscript refers to partial differentiation. The first term in (7) refers to the elastic force (Laplacian) and the second and third terms refer to the external forces. Davatzikos and Bryan solve (7) iteratively. The image is partitioned into $2N \times N$ square subimages and the partial derivatives are approximated with finite differences. This gives a set of discrete variables and functions of the form $\mathbf{x}_{i,j} = \mathbf{x}((i/2N), (j/N))$, and leads to the following discretised equation:

$$\beta_{i,j}(\mathbf{x}_{i+1,j} + \mathbf{x}_{i-1,j} + \mathbf{x}_{i,j+1} + \mathbf{x}_{i,j-1} - 4\mathbf{x}_{i,j}) + (1 - \gamma_{i,j})(\mathbf{c}_{i,j} - \mathbf{x}_{i,j}) + \gamma_{i,j}\mathbf{N}_{i,j} = 0. \tag{8}$$

Equation (8) is solved iteratively using successive over-relaxation [54].

In [55], Davatzikos further develops this approach to match both the cortical and ventricular surfaces which are pre-segmented from the images. A homothetic mapping is first used to achieve a coarse match then this is further refined. The refinement step is based on curvature and landmark matching. Curvature matching involves determining the minimum $(\kappa_{\min})$, max-

imum ($\kappa_{\max}$) and Gaussian (κ_G) curvatures of the segmented surfaces. The matching criterion involves determining the optimum displacement field such that

$$\arg\min_{\mathbf{u}} \int_S \sum_{\alpha\in\{\min,\max,G\}} [b_s^\alpha(\mathbf{u}(\mathbf{x})) - b_t^\alpha(\mathbf{x})]^2\, dS \quad (9)$$

where b_t^α are binary values of voxels corresponding to target (or source b_s) surfaces such that $b_t^\alpha = 1$ if $\kappa_\alpha >$ threshold T and $b_t^\alpha = 0$ otherwise. The curvature matching results in an external force $\mathbf{f}_1$ of the form

$$\mathbf{f}_1 = -\sum_{\alpha\in\{\min,\max,G\}} (\nabla b_s^\alpha)\,(b_s^\alpha - b_t^\alpha)\,. \quad (10)$$

The curved outlines of N corresponding sulci $\mathbf{s}_s^i(l)$ and $\mathbf{s}_t^i(l)$, l parametrises the sulcal curve, are obtained manually from the source and target images, respectively. The displacement function $\mathbf{u}$ can be constrained by minimising the squared distance between corresponding landmarks as follows:

$$\arg\min_{\mathbf{u}} \sum_{i=1}^{N} \int_L \left\|\mathbf{u}(\mathbf{s}_s^i(l)) - \mathbf{s}_t^i(l)\right\|^2 dl. \quad (11)$$

This results in a second force $\mathbf{f}_2$ that is proportional to the sum of residual distances between the two sets of landmarks:

$$\mathbf{f}_2 = -\sum_{i=1}^{N} \left(\mathbf{u}\left(\mathbf{s}_s^i(l)\right) - \mathbf{s}_t^i(l)\right). \quad (12)$$

The forces $\mathbf{f}_1$ and $\mathbf{f}_2$ are then incorporated as external forces in the Navier-Cauchy linear elastic PDE (4) as follows:

$$\lambda_r \nabla^2 \mathbf{u} + (\mu_r + \lambda_r)\nabla(\nabla.\mathbf{u}) + \mathbf{f}_1 + \mathbf{f}_2 = 0. \quad (13)$$

Equation (13) defines a 3-D elastic deformation that brings the cortical surfaces into registration. Davatzikos points out that there are large shape differences of certain brain structures across a population, e.g., the lateral ventricles. To accommodate this they propose to model the brain as an inhomogeneous body with nonzero initial strain (ϵ) at the ventricular surface. This results in a modified Navier-Cauchy PDE, of the form

$$\begin{aligned}&\{\mathbf{f} + \lambda\nabla^2\mathbf{u} + (\mu+\lambda)\nabla(\nabla.\mathbf{u})\}\\ &\quad +\{(\nabla\mathbf{u} + \nabla\mathbf{u}^T - 2\mathbf{I})\nabla\lambda + (\nabla\cdot\mathbf{u} - 3)\nabla\mu\}\\ &\quad +\{\epsilon(2\nabla\lambda + 3\nabla\mu) + (2\lambda + 3\mu)\nabla\epsilon\} = 0. \end{aligned} \quad (14)$$

The first term of (14) is the standard Navier-Cauchy PDE (4), the second term allows for material inhomogeneity, and the third term allows for the pre-strained ventricular surface.

Validation of Bajcsy's Algorithm: Bajcsy *et al.* [50] validated their elastic algorithm by registering a segmented brain atlas to patient CT scans. They manually segmented the ventricles in patient images and compared this to the corresponding region propagated from the atlas. Results indicated a maximum error of three to four pixels. Later Gee *et al.* [56] validated the algorithm for atlas to MR registration. This time the atlas was derived from myelin-stained sections to provide a segmentation of GM, WM, and CSF. The tissue types were manually delineated in the patient images. The voxel overlap was 66% for the region bound by the brain and ventricular surfaces and 78% for the region bounded by the GM/WM interface.

Validation of Davatzikos Algorithm: Davatzikos [55] evaluated his algorithm using six T1 weighted 3-D MR brain images of volunteers. Thirty six anatomical landmarks, corresponding to sulcal roots, ventricular horns, etc., were manually identified and used to measure the registration error. The mean, maximum (std) registration error was 3.4, 10.4 (2.1) mm.

IV. Fluid Flow Transformations

A. Theory

It is often useful to register images where there are large deformations. Large deformations are typically needed for intersubject registration because of anatomical variation over a population. A major limitation of the linear elasticity approach, using the Navier-Cauchy PDE (4) is that it is based on the assumption of an infinitesimally small deformation. Furthermore, for the regularization strategy used in linear elasticity (and TPS), the restoring force increases monotonically with strain [57] which penalizes large deformations. Christensen *et al.* [48], [57], [58] proposed a viscous fluid flow model to recover large deformations. This was applied after linear elastic registration.

Continuum mechanics provides the theoretical foundation for fluid flow. There are many standard texts on continuum mechanics, e.g., see [47] and [59]. Fluid flow models are based on idealised physical properties of fluids, e.g., they behave as a collection of particles that conform to Newtonian mechanics. Fluid models must satisfy physical laws such as the conservation of mass, energy, and linear and angular momentum. When a fluid is stationary there is no shear stress and so the Cauchy stress tensor $\boldsymbol{\sigma}_{ij}$ consists entirely of normal stresses (the hydrostatic pressure p). When a fluid flows the shear stresses are no longer negligible. They are represented by a viscous (shear) stress tensor. The Cauchy stress tensor then becomes the sum of a hydrostatic pressure term and the viscous stress tensor. The viscous stress tensor is usually considered to be a function of the rate of deformation tensor. If this relationship is linear then the fluid is known as Newtonian otherwise it is considered Stokesian [47]. Fluid flow can be explained in terms of the following notions from continuum mechanics.

- **Fluid velocity:** In the Eulerian frame[1], the velocity of an element of mass passing through $\mathbf{x}$ at time t is given by the material derivative of the displacement $\mathbf{u}(\mathbf{x})$ as follows:

$$\mathbf{v}(\mathbf{x}, t) = \frac{\partial \mathbf{u}}{\partial t} + \mathbf{v}.\nabla\mathbf{u}. \quad (15)$$

- **Rate of deformation:** The velocity gradient tensor $(\partial v_i/\partial x_j)$ can be considered as the sum of a symmetric tensor $\mathbf{D}$ and an anti-symmetric one $\mathbf{V}$ such that $(\partial v_i/\partial x_j) = D_{ij} + V_{ij}$. D_{ij} is referred to as the rate of deformation tensor and V_{ij} the vorticity tensor. In tensor notation, $D_{ij} = (1/2)((\partial v_i/\partial x_j) + (\partial v_j/\partial x_i))$ and

[1]The Eulerian frame is thought by some authors [57] to be the most suitable for tracking large deformations.

$V_{ij} = (1/2)((\partial v_i/\partial x_j) - (\partial v_j/\partial x_i))$. In vector notation, the rate of deformation tensor $\mathbf{D}$ can be expressed as follows:

$$\mathbf{D} = \frac{1}{2}(\nabla \mathbf{v} + (\nabla \mathbf{v})^T) \tag{16}$$

where $\mathbf{v}$ is the velocity vector and T denotes transpose.

- **Conservation of mass:** Leads to the following continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla.(\rho \mathbf{v}) = \eta \tag{17}$$

where ρ denotes the density of the fluid and the mass source term η allows for arbitrary creation or destruction of mass.[2]

- **Conservation of linear momentum:** Leads to the equation of motion:

$$\nabla \boldsymbol{\sigma} + \mathbf{f} = \rho \frac{d\mathbf{v}}{dt} + \eta \mathbf{v} \tag{18}$$

where $\mathbf{f}$, sometimes written $\mathbf{b}$, is the body force per unit volume.

- **Constitutive equations:** For a Newtonian fluid the viscous stress tensor $\boldsymbol{\sigma}$ is linearly related to the rate of deformation tensor $\mathbf{D}$ as follows:

$$\boldsymbol{\sigma} = -p\mathbf{I} + \lambda_f tr(\mathbf{D})\mathbf{I} + 2\mu_f \mathbf{D} \tag{19}$$

where p is the hydrostatic pressure and λ_f and μ_f are the viscosity coefficients of the fluid, tr is the trace.

Substituting (19) into (18) and then substituting for $\mathbf{D}$ in (16) allows us to derive the Navier-Stokes-Duhem equation

$$\rho \frac{d\mathbf{v}}{dt} = \mathbf{f} - \nabla p + (\mu_f + \lambda_f)\nabla(\nabla.\mathbf{v}) + \mu_f \nabla^2 \mathbf{v} - \eta \mathbf{v}. \tag{20}$$

For very slow flow rates (low Reynolds number) it is possible to neglect the inertial terms $\rho(d\mathbf{v}/dt)$ and $\eta\mathbf{v}$.

Assuming there is only a small spatial variation in the hydrostatic pressure then ∇p can also be neglected and (20) simplifies to the Navier-Stokes equation for a compressible viscous fluid

$$\mu_f \nabla^2 \mathbf{v} + (\mu_f + \lambda_f)\nabla(\nabla.\mathbf{v}) + \mathbf{f} = 0. \tag{21}$$

Essentially the Navier-Stokes PDE describes the balance of forces acting in a given region of the fluid. It characterizes an equilibrium state where changes in momentum of the fluid balance with changes in pressure and dissipative viscous forces. The $\mu_f \nabla^2 \mathbf{v}$ term is associated with constant volume or incompressible viscous flow whereas the $(\mu_f + \lambda_f)\nabla(\nabla.\mathbf{v})$ term allows for the expansion or contraction of the fluid. Remarkably, the Navier-Stokes PDE (21) is identical to the Navier-Cauchy PDE of linear elasticity (4) except that the PDE operates on velocity $\mathbf{v}$ rather than displacement $\mathbf{u}$.

B. Fluid Flow Algorithms

These are based on the viscous fluid flow model defined by the Navier-Stokes PDE (21). Because of the similarity to Navier-Cauchy PDE (4) solutions of linear elasticity can be transferred to fluid flow. Differential operators are applied to a velocity field that describes pixel motion. Fluid flow allows large localized deformations to be modelled, but has the disadvantage of sometimes increasing registration error [4] and high computational cost.

The most well-known fluid flow algorithm is due to Christensen *et al.* [48], [57], [58]. The overall registration strategy is based on a transformation hierarchy of successively increasing numbers of degrees-of-freedom [58] starting with affine then linear elasticity and finally a fluid flow algorithm [48]. The fluid flow algorithm iteratively solves the Navier-Stokes PDE. It evolves velocity fields that describe the motion of voxels over time (iteration). The body force $\mathbf{f}$ is determined from image similarity like elastic algorithms. It is assumed to take the form of a Gaussian sensor model [57]

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = -\alpha[A(\mathbf{x}) - B(\mathbf{x} - \mathbf{u})]\nabla B(\mathbf{x} - \mathbf{u}) \tag{22}$$

where $[A(\mathbf{x}) - B(\mathbf{x} - \mathbf{u})]$ is the difference of intensities between the target and deformed source images. The gradient of the source image $\nabla B(\mathbf{x} - \mathbf{u})$ gives the direction of the local forces applied to B. Given the current estimate of $\mathbf{u}$, the velocity and body force can be estimated using (15) and (22). This provides initial values for the Navier-Stokes PDE which is subsequently solved in discrete time steps by successive over-relaxation [60]. The updated velocity field then is used to update the displacement field. For large deformations, the numerical solution of the Navier-Stokes PDE can produce displacement fields that become singular [57]. To avoid this the determinant of the Jacobian of the transformation (J) is tracked. Each time it falls below 0.5 a new source image is generated by interpolation using the current displacement field and the algorithm is restarted using the new source image.

Solving the Navier-Stokes PDE is particularly computationally intensive which is a major disadvantage. To address this Christensen's algorithm was implemented on parallel hardware so that results could be obtained in a few hours [57]. Other authors have proposed faster solutions. Bro-Nielsen *et al.* [61] used a filter (convolution with Green's functions) in scale space. Freeborough *et al.* [27] solved the PDE hierarchically using the full multigrid method [62].

C. Validation of the Fluid Flow Algorithm

Christensen *et al.* [57] experimentally compared the fluid flow and linear elasticity algorithms. They used a synthetic image pair consisting of a small rectangular patch (source image) and a "C" shape with an area about an order of magnitude larger. They reported that the fluid algorithm was able to produce deformations to achieve an overlap $> 90\%$ whereas the linear elastic algorithm only achieved about 25%.

D. Optical Flow

Optical flow [63] has been widely used to track small scale motion in time sequences of images. It is based on the principle of intensity conservation between image frames. There is a similarity to fluid flow. The equation of motion for optical flow can

[2]Christensen *et al.* [57] argue that from an image registration perspective it is often desirable to allow local mass creation or destruction, however they do not seem to implement this in [57].

be derived by retaining the first-order terms of the Taylor expansion of the intensity function in the target frame. It is possible to relate the displacement[3] $\mathbf{u}$ to the change in intensity between frames $b(\mathbf{x}) - a(\mathbf{x})$ and the spatial derivative of intensity in the target frame $\nabla a(\mathbf{x})$, as follows:

$$\mathbf{u}.\nabla a(\mathbf{x}) = b(\mathbf{x}) - a(\mathbf{x}). \tag{23}$$

Demons Algorithm: The demons algorithm [64] uses the optical flow model (23). First (23) is approximated to give a numerically stable expression for $\mathbf{u}$

$$\mathbf{u}(\mathbf{x}) = \frac{[b(\mathbf{x}) - a(\mathbf{x})]\nabla a}{(\nabla a)^2 + [b(\mathbf{x}) - a(\mathbf{x})]^2}. \tag{24}$$

The displacement (force) on the source image is in the direction of ∇a and its orientation is $+\nabla a$ if $b(\mathbf{x}) > a(\mathbf{x})$ and $-\nabla a$ otherwise. A disadvantage of this model is that there are no constraints on the displacement and it does not necessarily preserve the topology. To reduce the effects of noise the displacement field is smoothed by Gaussian convolution. The algorithm iterates over time, during each iteration an incremental displacement field is determined and the source image is resampled for the next iteration.

E. Continuum Biomechanics

Although continuum mechanics explains well the behaviour of simple rubber-like materials it does not explain well the more complex behaviour of biological materials like soft tissue. Continuum biomechanics is concerned with extending the theory, particularly non-linear continuum mechanics [65], to deal with this. Humphrey [66] provides a detailed review of the subject. An important observation here is that although soft biological tissue has many different forms it is composed of only two basic components: cells and an extra-cellular matrix [66]. So mechanical models at a cellular level should be able to explain tissue behavior. Tissues could be modelled as mixture-composites that exhibit anisotropy, visco-elasticity and inelastic behavior using the theory of elasticity or visco-elasticity. The constitutive relations should be used to describe the material under certain conditions and not the material itself. In conclusion, new models are anticipated that better explain: the multiaxial behavior of muscle, growth, remodelling, damage, regeneration, cell mechanics, etc.

V. Transformations Based on Basis Function Expansions

In general, these transformations are not derived from physical models, but instead model the deformation using a set of basis functions. The coefficients are adjusted so that the combination of basis functions fit the displacement field (cf. interpolation). Much of the mathematical framework arises from the theory of function interpolation [67] and approximation theory [68], [69]. In approximation theory it is assumed that there is error in the samples, so the standard interpolation requirement that the function intersects samples is relaxed. As a result, the approximating function is usually much smoother than its interpolating counterpart.

Polynomial functions might seem an intuitive choice, however, global polynomials of degree larger than two can be unstable [1, Ch. 8]. Radial basis functions and piecewise polynomials (splines) are more stable and are widely used.

In general, these functions do not preserve the topology, however, recent work [70], [71] has sought to design functions that are diffeomorphisms, see Section VI-C.

A. Radial Basis Functions

Radial basis functions [72]–[74] are functions of the distance $||\mathbf{x} - \mathbf{x}_i||$ between the interpolation point $\mathbf{x}$ and basis function centre or landmark position $\mathbf{x}_i$. They can be defined as follows:

$$f(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i R(||\mathbf{x} - \mathbf{x}_i||) \tag{25}$$

where i indexes the landmarks, e.g., landmark pairs, N is the total number of landmarks, and α_i are weights which are determined by solving a set of linear equations. Examples are the Gaussian [75] and the inverse multiquadric, IMQ [76] defined in (29). These functions asymptotically tend to zero, but have global support. RBFs are positive definite functions which allow an optimal set of coefficients to be determined in closed-form. This is a particularly useful property for landmark based registration. Rohr *et al.* [77]–[79] and Fornefett *et al.* [80] have extensively investigated RBFs for the landmark-based registration of medical images. In [80], a general form is given which consists of a sum of polynomials and RBFs

$$\mathbf{u}(\mathbf{x}) = \sum_{j=1}^{M} \boldsymbol{\beta}_j \phi_j(\mathbf{x}) + \sum_{i=1}^{N} \boldsymbol{\alpha}_i R(||\mathbf{x} - \mathbf{p}_i||). \tag{26}$$

This requires solving a set of linear equations of the form

$$\begin{pmatrix} \mathbf{K} & \mathbf{P} \\ \mathbf{P} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{q}_k \\ 0 \end{pmatrix} \tag{27}$$

where the matrix $\mathbf{K}$ has elements $R_{ij} = R(||\mathbf{p}_i - \mathbf{p}_j||)$ and $\mathbf{P}$ has elements $P_{ij} = \phi_j(\mathbf{p}_i)$ and the column vectors $\mathbf{q}_k$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ have elements consisting of landmark positions and the coefficients $\{\alpha_i\}$ and $\{\beta_j\}$.

Others have compared the performance of TPS to polynomials and multiquadrics [81]–[83]. Arad [75] suggested that the TPS had favorable properties for image registration.

B. Multiquadrics

The multiquadric (MQ) is a type of radial basis function (25) with $R(||\mathbf{x} - \mathbf{x}_i||)$ is defined as follows [74]

$$R(||\mathbf{x} - \mathbf{x}_i||) = \sqrt{r_i^2 + d^2} \tag{28}$$

where r_i is the Euclidean distance $|\mathbf{x} - \mathbf{x}_i|$, $\mathbf{x}$ is an interpolated point, and $\mathbf{x}_i$ is the location of the ith landmark. The parameter d controls the amount of smoothing, larger d results in more smoothing. The inverse multiquadric (IMQ) is defined as the reciprocal of (28)

$$R(||\mathbf{x} - \mathbf{x}_i||) = [r_i^2 + d^2]^{-1/2}. \tag{29}$$

[3]This is actually velocity—displacement over the time interval of the two images. However, it can be considered as a displacement without loss of generality.

C. Weighted Mean

The WMN is defined by: $R(\|\mathbf{x}-\mathbf{x}_i\|) = G_i(\|\mathbf{x}-\mathbf{x}_i\|)/\sum_i^N G_i(\|\mathbf{x}-\mathbf{x}_i\|)$ chosen such that $G_i(\|\mathbf{x}-\mathbf{x}_i\|)$ is a monotonically decreasing RBF such as a Gaussian or cubic $G_i(\|\mathbf{x}-\mathbf{x}_i\|) = 1-3d_i^2+2d_i^3$ where $d_i = \|\mathbf{x}-\mathbf{x}_i\|_2/\|\mathbf{x}-\mathbf{x}_n\|_2$ [84]. The weighted sum makes it an approximating rather interpolating function. The width of R is tuned to the density of landmarks and the function becomes interpolating as R decreases [84].

D. Thin-Plate Splines

Historically, the TPS was used to design structures such as aircraft wings [85]. Later it was applied to function [86], [87] and spatial [88], [89] interpolation. Grimson [90] and Terzopoulos [91] described the TPS function mathematically as a variational Euler-Lagrange equation which minimises the bending energy. Essentially the TPS is the solution of a square Laplacian $\nabla^4\mathbf{u} = c\delta(0,0)$[7], [92]. Goshtasby [93] applied the TPS to the registration of remote sensing images. Bookstein *et al.* [92] introduced it into the modelling shape deformation in medical image analysis. According to [6] it is the most commonly used RBF.

The TPS is applicable to multidimensional interpolation problems and has useful smoothing properties. It is usually used with sets of homologous features, anatomical landmarks, which are typically manually located in the images. The TPS can be used even if the landmarks are irregularly spaced. Given a set of corresponding sets of features, the spline coefficients can be determined by the method of least squares [94]. In 2-D, the TPS has a logarithmic basis function $r^2\log(r)$, in 3-D this simplifies to r [80]. So the TPS displacement $\mathbf{u}(\mathbf{x})$ can be determined as follows [8]:

$$\mathbf{u}(\mathbf{x}) = \mathbf{A}\mathbf{x}+\mathbf{B}+\mathbf{I}\sum_{i=1}^{N} F_i r_i^2\log(r_i) \tag{30}$$

$$\mathbf{u}(\mathbf{x}) = \mathbf{A}\mathbf{x}+\mathbf{B}+\mathbf{I}\sum_{i=1}^{N} F_i r_i \tag{31}$$

where (30) and (31) refer to 2-D and 3-D space, respectively. The matrices $\mathbf{A}$ and $\mathbf{B}$ define an affine transformation and $\mathbf{I}$ is the identity matrix. Goshtasby [8] includes an additional stiffness parameter d in r_i such that $r_i^2 = (x-x_i)^2+(y-y_i)^2+d^2$. The coefficients of the linear transformation defined by $\mathbf{A}$ and $\mathbf{B}$ and the TPS coefficients $\{F_i\}$ are determined by solving the set of linear equations at the locations of landmarks in the source image $\mathbf{u}_i = \mathbf{u}(\mathbf{x}_i)$.

The TPS is a globally supported function and so it cannot accurately model localized deformation. Furthermore, outliers have a global impact and also large deformations can lead to singularities in the sets of equations that need to be solved which can result in the topology not being preserved. The global extent also leads to high computational complexity when large numbers of landmarks are used. Hence, some authors have improved its computational efficiency [95]–[98].

E. Approximating Thin-Plate Splines

Rohr *et al.* [78], [79] proposed an approximating rather than interpolating TPS that is more robust to outliers which occur because of errors in feature (landmark) localization. Landmark errors are considered as anisotropic and are measured using a quadratic approximation term. The registration functional $J_\lambda(\mathbf{u})$ consists of a landmark registration measure term and a TPS term $J_m^d(\mathbf{u})$ that regularizes the transformation

$$J_\lambda(\mathbf{u}) = \frac{1}{N}\sum_{i=1}^{N}[\mathbf{q}_i-\mathbf{u}(\mathbf{p}_i)]^T\mathbf{S}_n^{-1}[\mathbf{q}_i-\mathbf{u}(\mathbf{p}_i)]+\lambda J_m^d(\mathbf{u}) \tag{32}$$

where $\mathbf{p}_i$ and $\mathbf{q}_i$ denote landmarks in the source and target images. The covariance matrix $\mathbf{S}_n$ is a 3×3 matrix and represents anisotropic landmark localization errors, d refers to the dimension of the image and m to the chosen derivative order of the functional. The term $J_m^d(\mathbf{u})$ is defined by $J_m^d(\mathbf{u}) = \sum_{k=1}^d J_m^d(u_k)$ and $J_m^d(u_k)$ is defined as [9]

$$J_m^d(u_k) = \sum_{\alpha_1+\cdots+\alpha_d=m}\frac{m!}{\alpha_1!\ldots\alpha_d!} \times\int_{\mathbb{R}^d}\left(\frac{\partial^m u_k}{\partial x_1^{\alpha_1}\ldots\partial x_d^{\alpha_d}}\right)^2 d\mathbf{x}. \tag{33}$$

The $J_m^d(\mathbf{u})$ term defines the TPS and controls the smoothness of the transformation. Hence, the minimization of (32) results in a smooth transformation that approximates the distance between the landmark sets. The λ parameter controls the weighting between the two terms, the transformation becomes smoother as λ increases.

F. Wendland ψ-Function

Fornefett *et al.* [80] required a transformation function that could be used to model brain deformation resulting from neurosurgery. These deformations tend to be highly localized so standard RBFs such as TPS, MQ are unsuitable since they are global supported. They formulated a number of criteria that the transformation function should fulfill for landmark-based brain registration.

(a) Locality: $\mathbf{u}$ should have compact support and the extent of support region should be controllable.
(b) Solvability: (27) describing the mapping of landmark locations in the source image to the target image must be solvable. This amounts to the function $\mathbf{u}$ being positive definite.
(c) Preservation of topology: The transformation function $\mathbf{u}$ must be continuous and locally 1-to-1 and the determinant of the Jacobian of the transformation must be positive, i.e., $\det(\nabla u_k) > 0$ [100].
(d) The numerical solution should be computationally efficient.

They selected the $\psi_{d,k}$ function of Wendland [101] which is a locally supported RBF. Local support is desirable because it also reduces complexity and speeds up optimization. The $\psi_{d,k}$ function is multivariate in $\mathbb{R}^d$ and is $\mathcal{C}^{2k}$ continuous. It has a similar shape to a Gaussian, but it has finite extent, furthermore it is smooth unlike a truncated Gaussian. Like other RBFs, $\psi_{d,k}$

is positive definite in $\mathbb{R}^d$ and has a minimal polynomial degree of $\lfloor d/2 \rfloor + 3k + 1$ where $\lfloor \cdot \rfloor$ denotes the floor operator[4]

$$\psi_{d,k}(r) = I^k (1-r)_+^{\lfloor d/2 \rfloor + k + 1} \tag{34}$$

where

$$(1-r)_+^{\nu} = \begin{cases} (1-r)^{\nu} & 0 \le r < 1 \\ 0 & 1 \le r \end{cases} \tag{35}$$

and $I(f(r))^k$ denotes k applications of the integral operator $I(f(r))$ defined by

$$I(f(r)) \triangleq \int_r^{\infty} x f(x) dx \quad r \ge 0. \tag{36}$$

They selected $\psi_{3,1} = (1-r)_+^4 (4r+1)$ because it provides smooth $\mathcal{C}^2$ continuity with minimal degree and therefore computational expense for 3-D image registration. The locality criteria (a) is satisfied by using a multiplicative scaling factor a on the distance r, such that $\psi_a(r) = \psi(r/a)$. They demonstrated experimentally that for large $a >\approx 1000$ the function resembled a TPS, whereas for $a = 50$ it was much more spatially constrained. They proved that $\psi_{3,1}$ satisfies the aforementioned criteria (a)–(d).

G. Elastic Body Splines

Davis *et al.* [102] proposed an elastic body spline (EBS) for landmark based registration. The EBS is the solution to the Navier-Cauchy PDE of linear elasticity (4). In general, they assume a polynomial radially symmetric force $\mathbf{f}(\mathbf{x}) = \mathbf{c} r^{2k+1}$. They considered $\mathbf{f}(\mathbf{x}) = \mathbf{c} r$ and $\mathbf{f}(\mathbf{x}) = (\mathbf{c}/r)$ where $r = \|\mathbf{x}\|$ (Euclidean norm) and $\mathbf{c}$ is a constant vector. This gives an elastic displacement field of the form

$$\mathbf{u}(\mathbf{x}) = \sum_{i=1}^{N} \mathbf{R}(\mathbf{x} - \mathbf{p}_i) \mathbf{c}_i \tag{37}$$

where $\mathbf{p}_i$ is the location of the ith of N target landmarks and $\mathbf{c}_i$ are the coefficients associated with the force. Essentially, this is a linear combination of translated basis functions $\mathbf{R}$, represented by a 3×3 matrix. Consequently, the three components of the PDE are coupled. They solve these three coupled equations using the Galerkin vector method, see [46]. This transforms the three coupled PDEs into three independent radially symmetric, biharmonic ones and results in the following solutions for forces $\mathbf{f}(\mathbf{x}) = \mathbf{c} r$ and $\mathbf{f}(\mathbf{x}) = (\mathbf{c}/r)$, respectively

$$\mathbf{R}(\mathbf{x}) = \left[(12(1-\nu) - 1) r^2 \mathbf{I} - 3 \mathbf{x}\mathbf{x}^T \right] \tag{38}$$

$$\mathbf{R}(\mathbf{x}) = \left[(8(1-\nu) - 1) r \mathbf{I} - \frac{1}{r} \mathbf{x}\mathbf{x}^T \right] \tag{39}$$

where $\nu = (\lambda / [2(\lambda + \mu)])$ is the Poisson ratio, $\mathbf{I}$ is the 3×3 identity matrix and $\mathbf{x}\mathbf{x}^T$ is an outer product. The coefficients of the force $\mathbf{W} = \left[\mathbf{c}_1^T \cdots \mathbf{c}_N^T \right]^T$ are determined from the system of linear equations $\mathbf{W} = \mathbf{K}^{-1} \mathbf{U}$ where $\mathbf{K}$ is a $3N \times 3N$ matrix of $\mathbf{R}(\mathbf{p}_i - \mathbf{p}_j)$ elements and $\mathbf{U} = \left[\mathbf{u}_1^T \cdots \mathbf{u}_N^T \right]^T$ a vector of displacements. They solve for $\mathbf{W}$ using singular value decomposition [62]. In comparison the volume spline (VS) and 3-D TPS can be written in the form $\mathbf{R}(\mathbf{x}) = r^3 \mathbf{I}$ and $\mathbf{R}(\mathbf{x}) = r \mathbf{I}$. So apart from the multiplicative constants in (38) and (39), the terms including $\mathbf{x}\mathbf{x}^T$ can be considered as modifications to the VS and TPS so that they conform to the Navier-Cauchy physical model [102].

Kohlrausch *et al.* [103] argue that the force model $\mathbf{f}(\mathbf{x}) = \mathbf{c} r^{2k+1}$ in [102] does not decrease sufficiently fast and therefore models global rather than local deformations. So instead they propose a Gaussian force model $\mathbf{f}(\mathbf{x}) = (\mathbf{c}/(\sqrt{2\pi}\sigma)^3) e^{-(r^2/2\sigma^2)}$ hence they refer to it as GEBS. The Gaussian model has the advantage that σ can be used to control the localization of the deformation. Following a similar approach to [102] they show that a Gaussian force leads to $\mathbf{u}(\mathbf{x}) = (1/16\pi\mu)(1/1-\nu)\mathbf{R}(\mathbf{x})\mathbf{c}$, μ is the shear modulus and ν the Poisson ratio, and the new basis function has the form

$$\begin{aligned} \mathbf{R}(\mathbf{x}) = \Bigg[& (4(1-\nu) - 1) \frac{\operatorname{erf}(\hat{r})}{r} \\ & - \sqrt{\frac{2}{\pi}} \sigma \frac{e^{-\hat{r}^2}}{r^2} + \sigma^2 \frac{\operatorname{erf}(\hat{r})}{r^3} \Bigg] \mathbf{I} \\ & + \left[\frac{\operatorname{erf}(\hat{r})}{r^3} + 3\sqrt{\frac{2}{\pi}} \sigma \frac{e^{-\hat{r}^2}}{r^4} - 3\sigma^2 \frac{\operatorname{erf}(\hat{r})}{r^5} \right] \mathbf{x}\mathbf{x}^T \end{aligned} \tag{40}$$

where $\hat{r} = (\|\mathbf{x}\| / \sqrt{2}\sigma)$ and $\operatorname{erf}(x) \triangleq (2/\sqrt{\pi}) \int_0^x e^{-t^2} dt$ denotes the error function. The Poisson ratio ν depends on the material and is limited physically such that $0 \le \nu < 0.5$. They argue that because the Gaussian asymptotically falls rapidly to near zero, the affine and elastic registration can be determined separately. This results in a similar system of linear equations as in [102]. They solve $\mathbf{W}$ coefficients using a Tikhonov regularization scheme [62]. In a continuation of this work, Wörz *et al.* [104] have extended the method to deal with anisotropic landmark localization errors.

H. Quantitative Comparison of EBS and GEBS

Kohlrausch *et al.* [103] created a simple brain model in which the tumour can deform within a rigid cranium. The symmetry of the model allows the Navier-Cauchy PDE to be solved in a cylindrical coordinate system. They evaluated the GEBS model by comparing it and the EBS model, with forces $\mathbf{f}(\mathbf{x}) = \mathbf{c} r$ and $\mathbf{f}(\mathbf{x}) = (\mathbf{c}/r)$, to an analytical model for both $\nu = 0$ and $\nu = 0.49$. In all cases, they reported that GEBS outperformed EBS, in some cases by an order of magnitude. However, a disadvantage of the GEBS is that its computational complexity is several times larger than EBS.

I. B-Splines

B-splines were originally proposed for interpolation by Schoenberg [105] in the 1940s. Since then they have been applied widely, they have been popular for interpolation problems in signal processing since the 1990s [106]–[108]. More recently some authors [107], [109] have argued that *B*-splines are optimal as approximating functions. These basis functions can be extended to multivariate ones using tensor products. The FFD is an example of this. The mapping function $\mathbf{\Phi} : \Omega \mapsto \Omega$ is modelled using translations of a regularly spaced grid (lattice) of control points $\{\mathbf{P}_{i,j,k}\}$ where $i = \lfloor x/\delta_x \rfloor, \ldots$ is the index

[4]The floor operator $\lfloor x \rfloor$ gives the largest integer $i \in \mathbb{Z}$ not greater than x

of a control point and δ_x is the control point spacing in the x direction. FFDs usually use compact supported basis functions and $\mathbf{\Phi}$ is expressed in terms of local coordinates (u, v, w) with $u = x/\delta_x - \lfloor x/\delta_x \rfloor \ldots$. Given a set of univariate compact supported basis functions $\{\psi_{l,N}(x)\}$ of degree N. The FFD mapping function $\mathbf{\Phi}(u, v, w)$ can be defined as [110]

$$\sum_{l=0,m=0,n=0}^{N,N,N} \psi_{l,N}(u)\psi_{m,N}(v)\psi_{n,N}(w)\mathbf{P}_{i+l,j+m,k+n}. \tag{41}$$

FFDs have been applied in deformable models of the heart [111]. It was soon recognised that *B*-spline basis functions had superior properties [112] and FFDs based on *B*-splines were used for object modelling in 3-D [113] and 2-D [110], [114] and for animations [115]. Declerck [116] applied cubic *B*-spline FFDs to register SPECT cardiac images using the iterative closest point of extracted surfaces as a registration metric. Rueckert *et al.* [117] also used this FFD in combination with a voxel intensity similarity measure to register dynamic contrast enhanced MR breast images. The theory and methods for object modelling with polynomial and spline curves, including *B*-splines is well developed, see [118] and [119].

1) B-Spline Interpolation: B(asis) splines are a type of minimal support spline that was introduced by Schoenberg [105] for interpolation. A *B*-spline basis function of degree zero β_0 is a rect (step) function, the kernel used for nearest neighbor interpolation. A degree one *B*-spline basis function β_1 is constructed by the self-convolution of β_0, i.e., $\beta_1 = \beta_0 * \beta_0$. A *B*-spline basis function of degree N is constructed from N such convolutions, i.e., $\beta_N = \beta_0 * \cdots * \beta_0$ (N times). The support of *B*-spline basis functions depends on its degree we denote $\beta_{i,n}$ as a *B*-spline basis function of degree [5]n that is applied to an interval defined by the knot point x_i. *B*-spline basis functions of an arbitrary degree can be defined using the Cox-de Boor recursion formula, see [118]:

$$\beta_{i,0}(x) = \begin{cases} 1 & \text{if } x_i < x < x_{i+1} \\ 0 & \text{otherwise} \end{cases} \tag{42}$$

$$\beta_{i,n}(x) = \frac{x - x_i}{x_{i+n} - x_i}\beta_{i,n-1}(x) + \frac{x_{i+n+1} - x}{x_{i+n+1} - x_{i+1}}\beta_{i+1,n-1}(x). \tag{43}$$

B-splines can be used as either interpolators or approximators. Approximators do not intersect samples, but instead minimize an error metric such as the L_2 norm. When used as approximators *B*-splines are said to provide an optimal trade-off between a smooth function and a close fit to the data samples [107], [108]. This property is useful because often it is desirable to have a smooth fit to noisy sample points. To construct a *B*-spline interpolator it is necessary to prefilter the image with a high-boost filter. If this is ignored then there is too much low-pass filtering. Unser [106] proposed using a spatial domain recursive filter to prefilter a discrete function prior to *B*-spline interpolation. In summary, the discrete function $f_{\Delta x}(j)$ with sample points $j \in \mathbb{Z}$ and sample interval Δx can be interpolated to a continuous function $f_{\Delta x}(x)$ [67] as follows:

$$f_{\Delta x}(x) = \sum_{j\in\mathbb{Z}} f_{\Delta x}(j) \sum_{k\in\mathbb{Z}} (b^n)_k^{-1}\beta_{k,n}. \tag{44}$$

The functions $\sum_{k\in\mathbb{Z}}(b^n)_k^{-1}\beta_{k,n}$ are infinite impulse response filters known as cardinal splines of degree n.

Curve Fitting Using De Boor Control Points: A fundamental theorem of *B*-splines [105], [120] states that any spline function of degree n, $S_n(x)$ can be represented as a linear combination of *B*-spline basis functions of the same degree over the same partition i.e., $S_n(x) = \sum_{k=-\infty}^{+\infty} c(k)\beta_{k,n}(x)$. Curve fitting can be achieved using a set of de Boor control points $\{\mathbf{P}_i\}$ that are joined to form a de Boor polygon, see Fig. 1. The control points are adjusted so that $S_n(x \in [x_0, x_m]) = \sum_{i=0}^{m} \beta_{i,n}(x)\mathbf{P}_i$ fits the data samples. The number of control points depends on the degree of splines and the number of knots.

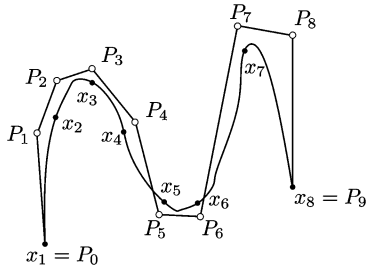

Fig. 1. De Boor Polygon consisting of control points $\{\mathbf{P}_0, \ldots, \mathbf{P}_9\}$ that defines the spline curve $x_1, \ldots, x_8$.

2) Free-form Deformations: FFDs are similar to the idea of tensor products of univariate splines suggested by de Boor [118, Ch. 17] for modelling surfaces. Sederberg and Parry [121] generalised this to volumes. The image domain is partitioned into a lattice of rectangular sub-domains that are aligned with the image axis. They used globally supported Bernstein basis functions $b_{p,q}$ which resulted in the following FFD:

$$\mathbf{\Phi}(x, y, z) = \sum_{i=0,j=0,k=0}^{l,m,n} b_{i,l}(x)b_{j,m}(y)b_{k,n}(z)\mathbf{P}_{i,j,k}. \tag{45}$$

In this case l, m, n denote the number of control points in the three directions. The Bernstein polynomial basis functions are defined by

$$b_{p,q}(x) = \binom{q}{p} x^p(1-x)^{q-p}. \tag{46}$$

They suggested that other basis functions such as *B*-splines would be suitable. FFDs have advantages compared to previous surface based methods. This is because the deformation model is object-independent in the sense that it applies to 3-D space and is independent of the object's surface.

Hsu [112] argued that cubic *B*-splines had superior properties compared to the Bernstein polynomials used in [121]. They provide both local control within a support region and continuity when control points are moved—they join C^2 smoothly

[5]The notation β_n as in [67] is used rather than β_{n+1} to denote a *B*-spline basis function of degree n.

at the knots. For cubic *B*-splines the support region is four control points in each direction and $\beta_{i,n}$ represents the ith *B*-spline basis function of degree n. Their FFD takes the following form within the support region:

$$\mathbf{\Phi}(u,v,w) = \sum_{l=0,m=0,n=0}^{3,3,3} \beta_{l,3}(u)\beta_{m,3}(v)\beta_{n,3}(w) \times \mathbf{P}_{i+l,j+m,k+n} \quad (47)$$

with cubic *B*-spline basis functions

$$\beta_{0,3}(u) = \frac{(1-u)^3}{6}$$
$$\beta_{1,3}(u) = \frac{(3u^3-6u^2+4)}{6}$$
$$\beta_{2,3}(u) = \frac{(-3u^3+3u^2+3u+1)}{6}$$
$$\beta_{3,3}(u) = \frac{u^3}{6}.$$

Lee *et al.* [110], [114] proposed a multilevel cubic *B*-spline FFD for landmark-based matching of 2-D images. They showed that a one-to-one mapping could be achieved by limiting the displacement of control points to less than half the control point spacing. Their FFD had local form

$$\sum_{l=0,m=0}^{3,3} \beta_{l,3}(u)\beta_{m,3}(v)\mathbf{P}_{i+l,j+m}. \quad (48)$$

They chose control point locations based on dyadic subdivision. They regularized their transformation by using a cost function that was based on the TPS bending energy and an image energy term $G_\sigma * |\nabla(I)|^2$, where $*$ denotes convolution and G_σ a Gaussian of standard deviation σ.

Gao and Sederberg [115] used a FFD to generate animations with minimal human interaction. They chose a hierarchy of first order (linear) *B*-splines $\beta_{i,1}$ on a grid of $2^k \times 2^k$ control points $\mathbf{P}_{i,j}$. Their FFD has the form

$$\sum_{l=0,m=0}^{2^{k-1},2^{k-1}} \beta_{l,1}(x)\beta_{m,1}(y)\mathbf{P}_{i+l,j+m} \quad (49)$$

where

$$\beta_{j,1}(x) = \begin{cases} 2^k x - x_j, & \text{if } \frac{x_j}{2^k} < x < \frac{x_{j+1}}{2^k} \\ x_j + 2 - 2^k x, & \text{if } \frac{x_{j+1}}{2^k} < x < \frac{x_{j+2}}{2^k} \\ 0, & \text{otherwise.} \end{cases} \quad (50)$$

They used a regularization strategy based on minimizing a cost function $C = C_i + \alpha C_a + \beta C_s$ that incorporates rudimentary deformation energy terms C_a and C_s as well as color image similarity C_i. The Lagrange multipliers α and β are determined empirically.

Registration Using B-Splines: Rueckert *et al.* [117] proposed using a cubic *B*-spline FFD with a voxel intensity similarity measure. The algorithm searched for the set of control point displacements that minimized the cost function $C(\mathbf{T}(\mathbf{x})) = S(A, B(\mathbf{T}(\mathbf{x}))) + \lambda E(\mathbf{T}(\mathbf{x}))$. Where $S(A, \mathbf{T}(B))$ is the image similarity measure and $E(\mathbf{T})$ is the TPS bending energy. The Lagrange multiplier λ controls the amount of regularization and was chosen empirically. Normalized mutual information was used for similarity metric and $C(\mathbf{T}(\mathbf{x}))$ was minimized using a gradient descent method. The FFD grids can be constructed hierarchically so that deformations can be determined by multiresolution, see [28].

Kybic *et al.* [30], [122] proposed cubic *B*-splines as a deformation model of the distortion in echo planar MR brain images [30] and and for the registration of MR, SPECT, and CT images of the brain and heart [122]. In a continuation of this work, Sorzano *et al.* [123] used cubic *B*-splines to model both image deformation and also to interpolate images. Their 2-D transformation is defined as follows [122]:

$$\mathbf{u}(\mathbf{x}) = \sum_{i,j\in\mathbb{Z}} \mathbf{c}_{i,j}\beta^n\left(\frac{x}{h} - i\right)\beta^n\left(\frac{y}{h} - j\right) \quad (51)$$

where $\beta^n(x) = \sum_{r\in\{-1,1\}}((n+1/2) - rx)\beta^{n-1}(x-(r/2))$, h is a multiresolution scaling parameter and $n = 3$ for cubic *B*-splines. They argued that a function regularization strategy based on a Laplacian or TPS bending energy is deficient because only pure second order derivatives and not cross-terms of first order derivatives are considered. To overcome this, they proposed a regularization scheme based on two terms, the gradients of the divergence $\int_{\mathbb{R}^2} ||\nabla(\nabla\cdot\mathbf{u})||d\mathbf{x}$ and the curl $\int_{\mathbb{R}^2} ||\nabla(||\nabla\times\mathbf{u}||)||d\mathbf{x}$ of the displacement field. They evaluated the accuracy of this approach using 2-D electrophoresis images and simulated barrel and pincushion distortion. According to their results the convergence rate was faster when regularization was used and faster still when both regularization and landmarks were used. However, there was no comparison between their proposed method of regularization and the standard Laplacian one.

J. Validation of Rueckert's B-Spline FFD Algorithm

The cubic *B*-spline FFD was compared to an affine model for the registration dynamic contrast enhanced MR breast images [117]. According to this when there was subject motion, the FFD outperformed both the rigid or affine models, in terms of both the correlation of voxel intensities and visual assessment. In [28], the algorithm was evaluated by comparing atlas propagations of the cerebral lateral ventricles with manual segmentations. Here, it was found that the algorithm was accurate enough to compare cohorts when there was a small deformation. In [124], a bio-mechanical model was used to generate deformations and these were compared with the ones determined using the registration algorithm.

K. Piecewise Affine

This is sometimes also termed block-matching, multiple or poly-affine registration. It is a relatively simple model, the source image is divided into a number of rectangular subimages or blocks and these are individually registered to the target image. Typically an affine or rigid-body model is used. This approach invariably uses a voxel intensity similarity metric as a registration measure. This gives a uniformly spaced displacement field. However, it has a particular problem in that the displacement field is not necessarily continuous. Many authors solve this by regularization, for instance, by low pass filtering.

Another issue is the number of degrees of freedom or, equivalently the size of the blocks. If there are too few voxels in a block then there is insufficient information to drive registration.

The original method is usually credited to Collins *et al.* [125]. They used it for brain segmentation, to determine the geometrical variability of brain structures across a population. The algorithm was named ANIMAL and was intended to be used for intersubject brain registration. It assumes that brain deformation can be modelled as a set of translations at regularly spaced nodes of a dense 3-D cubic lattice. It uses a hierarchical strategy in which the images are convolved with a Gaussian kernel and its first derivative. Large deformations are recovered first at coarse scales and then smaller ones at finer scales. The algorithm iteratively optimizes a cost function consisting of a voxel similarity measure and a regularization term. The similarity measure is based on normalized cross correlation (NCC). The regularization term C is a function of the magnitude of the displacement of each voxel $u(\mathbf{x})$ and the FWHM of the Gaussian convolution kernel, $u_{\max}$, it has the following form:

$$C(\mathbf{x}) = \sum_{\mathbf{x}\in\mathbf{X}} \frac{cu(\mathbf{x})^{2/3}}{u_{\max}^{2/3} - u(\mathbf{x})^{2/3}} \tag{52}$$

where c is a constant. A presegmented atlas is used for the target image. The subject images are registered to this and the transformation is inverted and used to propagate the segmentation into the subject images. The algorithm was evaluated by comparing displacements of 34 manually located landmarks from 3-D MR brain scans of $n = 17$ normal subjects [125]. The average standard deviation of voxel displacements in the three Cartesian directions was 4.2 mm for the algorithm and 3.9 mm for the manual method.

L. Wavelets

In Fourier analysis, a function[6] is decomposed into a set of sinusoidal basis functions. Sinusoidal functions are perfectly localized in frequency, but completely unlocalised in space. In contrast, in wavelet analysis, basis functions are localized in both the frequency and spatial domains [126]. Wavelets are implicitly designed for the multiresolution analysis of signals. A vector function subspace of $L^2(\mathbb{R})$ is constructed from a nested set of subspaces $V^0 \subset V^1 \ldots$, such that $V^{i+1} = V^i \oplus W^i$ where W^i contains the additional detail to generate V^{i+1}. V^i and W^i are, respectively, constructed from usually orthogonal sets of scaling $\{\phi_{a,b}\}$ and wavelet $\{\psi_{a,b}\}$ functions. Defined as $\phi_{a,b} = 2^{-(a/2)}\phi((x/2^a)-b)$ and $\psi_{a,b} = 2^{-(a/2)}\psi((x/2^a)-b)$. A 1-D function can be represented as an expansion of translated and dilated scaling and wavelet basis functions as follows:

$$f(\mathbf{x}) = \sum_{a\in\mathbb{Z},b\in\mathbb{Z}} \langle f(\mathbf{x}), \phi_{a,b}\rangle \phi_{a,b} + \sum_{a\in\mathbb{Z},b\in\mathbb{Z}} \langle f(\mathbf{x}), \psi_{a,b}\rangle \psi_{a,b}. \tag{53}$$

Wavelets have an advantage over the Navier-Cauchy eigenfunctions $\boldsymbol{\phi}$ in (5) because they allow deformations with local support to be modelled from a finite set of basis functions. Furthermore, they provide an implicit multiresolution representation of the transformation.

Amit [127] considered registration as a variational problem where $\mathbf{u}(\mathbf{x})$ is obtained by minimising a cost function consisting of the sum of square intensity differences and a smoothing term. Here, $\mathbf{u}(\mathbf{x})$ is modelled in 2-D by the wavelet expansion $\sum_i \sum_{n,k,l} \left\langle \mathbf{u}(\mathbf{x}), \boldsymbol{\psi}^i_{n,k,l} \right\rangle \psi^i_{n,k,l}$ with basis functions that are tensor products of $\psi_{n,k}$ and $\phi_{n,k}$

$$\begin{aligned}
\psi^0_{n,k,l} &= \phi_{n,k}(x_1)\phi_{n,l}(x_2)\\
\psi^1_{n,k,l} &= \phi_{n,k}(x_1)\psi_{n,l}(x_2)\\
\psi^2_{n,k,l} &= \psi_{n,k}(x_1)\phi_{n,l}(x_2)\\
\psi^3_{n,k,l} &= \psi_{n,k}(x_1)\psi_{n,l}(x_2).
\end{aligned} \tag{54}$$

The smoothing term is

$$u^2_{00} + \sum_{n=1}^{\infty} \sum_{k,l=0}^{2^n-1} (1+4^{2n}) \sum_{i=1}^{3} \left\langle \mathbf{u}(\mathbf{x}), \boldsymbol{\psi}^i_{n,k,l} \right\rangle^2 \tag{55}$$

where $u_{00} = \langle \mathbf{u}(\mathbf{x}), \phi_{0,0}\rangle$, with $\phi_{0,0} = \mathbf{1}$, The optimal value of $\mathbf{u}(\mathbf{x})$ is found by gradient descent and the inner products are determined by the fast discrete wavelet transform of Mallat [128] using Daubechies [129] compactly supported wavelets.

Gefen [130], modelled $\mathbf{u}(\mathbf{x})$ in 3-D with a wavelet expansion

$$u(\mathbf{x}_i) = \sum_{k=0}^{N-1} c^{i1}_{Jk} \Phi^1_{J,k} + \sum_{k=0}^{N-1} \sum_{s=2}^{8} \sum_{j=R}^{J} c^{is}_{jk} \Phi^s_{j,k} \tag{56}$$

where $\Phi^s_{J,k}$ is determined from the tensor product of 1-D order 3 spline wavelet basis functions, cf. (54) and s denotes the subband. In this approach, the Navier-Cauchy PDE is solved by minimising the elastic energy functional separately for each component of the deformation field using the Levenberg-Marquardt algorithm. Gefen compared the wavelet and TPS models for intersubject registration of histology images of rat brains. The mean surface error of both methods decreased with the number of registration parameters. For the same number of parameters, the wavelet method was about 10% more accurate than the TPS, however it is substantially more computationally intensive. For the TPS, the number of parameters depends on the number of corresponding points. However, for the wavelet method it depends on the number of voxels. The wavelet method can therefore be applied at a finer scale, 2800 parameters gave an accuracy of 2.39 voxels.

Wu *et al.* [131], [132] expressed the displacement field as a wavelet expansion and determined the optimal set of wavelet coefficients with a coarse-to-fine optimisation strategy using the Levenberg-Marquardt algorithm. They required a smooth solution and chose the wavelet function of Cai and Wang [133] which are cubic spline basis functions that span a Sobolev rather than L^2 space.

1) Piecewise Affine With Optical Flow Regularization: Hellier *et al.* [134] proposed a multiresolution algorithm based on a piecewise affine transformation regularized with an optical flow model. They used a quadratic form of optical flow, defined by

[6]The function is assumed to be continuous, real-valued and square integrable $L^2(\mathbb{R})$, i.e., $\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$.

(23). They added a second regularization term $||\mathbf{u}(\mathbf{x})-\mathbf{u}(\mathbf{x}_0)||$, where $\mathbf{x}_0$ is in $\nu(\mathbf{x})$ a neighborhood of $\mathbf{x}$, to smooth the displacement field. The combined cost function was

$$C(\mathbf{u}) = \sum_{\mathbf{x}\in\mathbf{X}} [\mathbf{u}(\mathbf{x})\cdot\nabla a(\mathbf{x}) + a(\mathbf{x}) - b(\mathbf{x})]^2 + \alpha \sum_{\mathbf{x_0}\in\nu(\mathbf{x}):\mathbf{x}\in\mathbf{X}} ||\mathbf{u}(\mathbf{x})-\mathbf{u}(\mathbf{x}_0)||^2. \quad (57)$$

They used a multi-resolution strategy to circumvent the small displacement limitation of optical flow. This consisted of both a Gaussian image pyramid and a multi-grid method. The multigrid method involved partitioning the image into rectangular blocks. The displacement field for each block is estimated using the robust M-estimator [135]–[137]. The M-estimator [135] is a robust maximum likelihood estimator that decreases the weighting of values at the tails of the distribution. It is typically used for iterative least squares problems to reduces the impact of outliers. The transformation model used depends on the number of voxels in the block. For blocks with 12 or more voxels a full affine is estimated, for ones with between 6 and 12 voxels a rigid-body and for those less than 6 a simple translation is used. The multigrid method is thought to have an advantage of being robust to MR intensity inhomogeneities which tend to be low spatial frequency.

VI. Constraints on the Transformation

A. Inverse Consistency

Christensen and Johnson [138], [139] used inverse consistency as regularization constraint. Inverse consistency can be explained by considering the transformations obtained by registering image A to B as $\mathbf{T}_{AB}$ and the inverse one from B to A as $\mathbf{T}_{BA}$. If the transformations $\mathbf{T}_{AB}$ and $\mathbf{T}_{BA}$ are consistent their composition is the identity. For almost all registration algorithms $\mathbf{T}_{AB} \neq \mathbf{T}_{BA}^{-1}$. So they introduced a regularization constraint that penalizes the inconsistency. In [138], they consider linear elastic registration as defined by the Navier-Cauchy PDE (4). The cost function C is a linear combination of: image similarity C_{sim}, the consistency of the forward and backward transformations C_{icc}, and a regularization term that is related to the energy of the deformation C_{reg}. $C_{\text{sim}} = S(A, B(\mathbf{T}_{BA}(\mathbf{x})) + S(A(\mathbf{T}_{AB}(\mathbf{x}), B)$ the square difference similarity measure is used so $S(A,B) = \int_\Omega |A(\mathbf{x}) - B(\mathbf{x})|^2 dx$. C_{icc} is determined from the residual between the forward and backward transformation, i.e., $C_{\text{icc}} = C_{\text{icc}}(A,B) + C_{\text{icc}}(B,A)$ with $C_{\text{icc}}(A,B) = \int_{\mathbf{x}\in\Omega}[\mathbf{T}_{AB}(\mathbf{x}) - \mathbf{T}^{-1}{}_{BA}(\mathbf{x})]dx$. C_{reg} relates to energy of the transformation $C_{\text{reg}} = C_{\text{reg}}(A,B) + C_{\text{reg}}(B,A)$ where $C_{\text{reg}}(A,B) = \int_{\mathbf{x}\in\Omega} |L\mathbf{T}_{AB}(\mathbf{x})|^2 dx$.

B. Topology Preservation

The topology can be preserved by ensuring that two conditions are fulfilled [140]: 1) the determinant of the Jacobian of the transformation J is always positive; 2) the transformation is bijective. Continuity is implied by the existence of J.

Noblet *et al.* [141] proposed a method to ensure that $J > 0$ for a hierarchy of linear *B*-splines. For *B*-splines of degree one, the displacement can be represented as a product of three terms of the form $a_v + b_v v$ where v is x or y or z. Expanding the determinant of the derivative terms leads to an expression for J in terms of x, y and z

$$J(x,y,z) = \sum_{i,j,k=(0,0,0)}^{2,2,2} \alpha_{i,j,k} x^i y^j z^k. \quad (58)$$

Given a gradient descent optimization method with step length δ it is possible to express J as a function of δ i.e., $J(x,y,z,\delta)$ in (58). In this way, it is possible to limit δ so that $J > 0$.

C. Diffeomorphic Transformations

Miller *et al.* [142] proposed that the group of diffeomorphic mappings, as proposed by Christensen *et al.* [57] for fluid flow registration, were suitable to generate groups of computational anatomies [143]–[145]. The problem with the fluid flow algorithm [57] is that singularities can arise from the successive overrelaxation method used to solve the Navier-Stokes PDE [146]. These can be avoided by regularising the velocity field [146]. To achieve this a diffeomorphic space-time mapping $\mathbf{\Phi}(\mathbf{x},t)$ where $\mathbf{x}\in\mathbf{\Omega}$ and $t\in[0,1]$ is required. This is mapping is related to the displacement field by $\mathbf{\Phi}(\mathbf{x},1) = \mathbf{x} + \mathbf{u}(\mathbf{x})$ and satisfies the following [142]:

$$\frac{\partial \mathbf{\Phi}(\mathbf{x},t)}{\partial t} = \mathbf{v}(\mathbf{x},t) \quad (59)$$

$$\frac{\partial \mathbf{\Phi}^{-1}(\mathbf{y},t)}{\partial t} = [\nabla\mathbf{\Phi}(\mathbf{y},t)]^{-1}\mathbf{v}(\mathbf{y},t) \quad (60)$$

$$\mathbf{\Phi}(\mathbf{x},0) = \mathbf{\Phi}^{-1}(\mathbf{x},0) = \mathbf{I}. \quad (61)$$

The function $\mathbf{\Phi}(\mathbf{x},t)$ describes a diffeomorphic flow through space-time. The vector $\mathbf{y}$ is a location in the target image and the derivative $\nabla\mathbf{\Phi}(\mathbf{y},t)$ is the Jacobian of the transformation. A numerical method for obtaining a diffeomorphism is given in [147]. Joshi and Miller [70] describe how diffeomorphisms can be used for landmark matching. The optimal displacement $\hat{\mathbf{u}}$ is determined by integrating the optimal velocity $\hat{\mathbf{v}}$ over time, i.e., $\hat{\mathbf{u}}(\mathbf{x}) = \int_T \hat{\mathbf{v}}(\mathbf{x},t)dt$. The optimal velocity is determined by minimizing two terms. The first term relates to the energy of the flow. The second term refers to the distance (residual) between the landmarks in the target image $\{\mathbf{q}_i\}$ and the time dependent mappings of the landmarks in the source image $\{\mathbf{\Phi}(\mathbf{p}_i,t)\}$. This leads to the following equation for the velocity $\hat{\mathbf{v}}$:

$$\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \int_T \int_\Omega ||L\mathbf{v}(\mathbf{x},t)||^2 dxdt + D(\mathbf{\Phi}(\mathbf{p},T)). \quad (62)$$

The differential operator L is modelled on the Navier-Stokes PDE (21) with incompressible flow i.e., $\nabla\cdot\mathbf{v} = 0$. Consequently, L is a diagonal operator and takes the form $L_{ii} = -\nabla^2 + c$ where c is a constant. The operator L and its boundary conditions are chosen so that the 3×3 matrix Green's function $\mathbf{G}(\mathbf{x},\mathbf{y})$ is continuous in $\mathbf{x}$ and $\mathbf{y}$ and $\mathbf{K}(\mathbf{x},\mathbf{y}) = \mathbf{G}(\mathbf{x},\mathbf{y})\mathbf{G}(\mathbf{x},\mathbf{y})^\dagger$ is a positive definite operator. The $D(\cdot)$ term is given by

$$D(\mathbf{\Phi}(\mathbf{p},T)) = \sum_{i=1}^{N} [\mathbf{q}_i - \mathbf{\Phi}(\mathbf{p}_i,T)]^T \mathbf{S}_n^{-1} [\mathbf{q}_i - \mathbf{\Phi}(\mathbf{p}_i,T)]. \quad (63)$$

Here, $\mathbf{S}_n$ is a 3×3 covariance matrix that represents anisotropic landmark localization errors. The $\hat{\mathbf{v}}$ minimiser can be rewritten, see [70], in the following form:

$$\hat{\mathbf{v}}(\mathbf{x},t) = \sum_{i=1}^{N} \mathbf{K}(\hat{\mathbf{\Phi}}(\mathbf{p}_i,t),\mathbf{x}) \sum_{j=1}^{N} [\mathbf{C}^{-1}]_{ij} \dot{\hat{\mathbf{\Phi}}}(\mathbf{p}_j,t) \quad (64)$$

with $\mathbf{K}(\mathbf{x},\mathbf{y})$ a diagonal matrix

$$(2(2\pi)^{(5/2)}/\sqrt{c}) \exp(-(1/\sqrt{c})\|\mathbf{x}-\mathbf{y}\|)\mathbf{I}$$

$\mathbf{I}$ is the 3×3 identity matrix. Assuming a normally distributed velocity field then $\mathbf{C}$ is defined as a $3N \times 3N$ covariance matrix with 3×3 submatrices $\mathbf{C}_{ij} = \mathbf{K}(\hat{\mathbf{\Phi}}(\mathbf{p}_i,t), \hat{\mathbf{\Phi}}(\mathbf{p}_j,t))$. The time derivative of the optimal diffeomorphic mapping $\dot{\hat{\mathbf{\Phi}}}$ for the rth landmark is given by

$$\dot{\hat{\mathbf{\Phi}}}(\mathbf{p}_r,T) = \arg\min_{\dot{\mathbf{\Phi}}} \int_T \sum_{i,j=1}^{N} \dot{\mathbf{\Phi}}^T(\mathbf{p}_i,t)[\mathbf{C}^{-1}]_{ij} \dot{\mathbf{\Phi}}(\mathbf{p}_j,t)dt + D(\hat{\mathbf{\Phi}}(\mathbf{p},T)). \quad (65)$$

They implement (64) by discretizing time and determining the perturbation of landmark displacements over each time step, for details see [70]. They use simple test examples to demonstrate that $J > 0$ for their transformation.

VII. Comparative Evaluation of Nonrigid Transformations

A critical aspect in choosing a transformation model is its impact on the accuracy of the nonrigid registration algorithm. However, this error is difficult to evaluate for real-world medical applications. This is because of the difficulty in obtaining ground truth transformations. A possible approach is to use simulation, but this must be realistic to be valuable. Furthermore, there are a wide range of potential applications and for most of these the transformation model is not precisely known. An alternative approach is to measure the geometric error at corresponding landmark locations. A large number of samples would be needed to give a representative estimate over the entire image domain. Despite this, for many applications, only the error at key structures is important so this approach can be viable. A limitation here is the localization error in identifying landmarks.

There is little published work on the comparative evaluation of nonrigid algorithms with different transformation models. Two recent studies have been published by Hellier *et al.* [148] and Zagorchev and Goshtasby [84]. There are new projects such as NIREP [149] that aim to provide systematic evaluation strategies. Hellier *et al.* [148] evaluated six methods for the intersubject registration of MR brain images. These consisted of rigid-body [150], [151] and nonrigid registration algorithms and the proportional squaring method of Talairach [40] which requires manual input. The nonrigid algorithms were based on fluid flow [139], optical flow [64] and piecewise affine [125] transformation models. They used a variety of metrics to evaluate error, both globally over the entire image domain and locally for certain key brain structures. According to the global criteria accuracy appeared to increase with the number of degrees-of-freedom of the transformation. Also, folding was noted for the optical flow algorithm [64]. For the local criteria there were mixed results and there was no significant difference between rigid and nonrigid algorithms. In conclusion, they recommended combining anatomical landmarks with intensity-based registration to increase accuracy.

Zagorchev and Goshtasby [84] compared four landmark-based methods [TPS, MQ, WMN, and piecewise linear (PL)] in terms of accuracy and computational cost. The TPS, MQ, and WMN are globally supported RBFs while PL is locally supported. The WMN is an approximating function while the others are interpolating. The accuracy generally depended on the size of the local geometrical differences between the images and the number and distribution of landmarks. When there were large local geometrical differences, WMN and PL were considered the most accurate. This was thought to be because: 1) when landmarks are irregularly spaced there are large errors for RBFs like TPS and MQ in image regions with a low landmark density; 2) when the landmark spacing is highly variable the system of equations that needs to be solved for each component becomes ill-conditioned. When there are small local geometrical differences and a small set of widely spread landmarks then TPS and MQ were preferred because they are interpolating functions. Generally, PL was the most suitable method for images with local geometrical differences because the local support property ensures that errors are not propagated globally. When there is a large number of landmarks with localization error the WMN method was preferred because 1) it does not require the solution of a system of equations that could be numerically unstable and 2) the averaging process when calculating the mean value reduces noise effects.

VIII. Conclusion

Nonrigid registration has become an important tool in medical image analysis. It can provide automated quantitative measurements for a large range of biological processes *in vivo*.

In principle, physical models have the advantage of providing physically realistic solutions. However, some of these, e.g., linear elasticity can only accurately model small deformations, which is a limitation because often soft tissue exhibits large deformation. Fluid flow is more appropriate for this and can ensure that the topology is preserved. However, it does not model the elastic component of tissue deformation and contains a solution space that cannot be realized in many tissue deformation states. Also, the solutions of the associated PDEs are often highly computationally complex. In reality, tissue exhibits a complex behavior, only in certain conditions can it be considered as an elastic or visco-elastic material and it usually behaves anisotropically. It is expected that better physical models should emerge as our understanding of continuum biomechanics advances [66].

Basis function expansions do not, in general, describe the physical or biological processes that cause the geometrical change. Instead they construct an interpolating or approximating function to model it. Some basis functions are compactly supported which allows highly localized deformation to be modelled. Compact support also has the advantage of reducing complexity and speeds up optimization. Generally, basis function expansions are easier to solve computationally than PDEs. Radial basis functions are used for landmark-based

registration. They simplify the multidimensional representation of the deformation. They provide fast closed-form solutions. *B*-spline and wavelet expansions are similar, but they do not use a simplifying distance metric. Instead multidimensional basis functions are constructed from linear combinations of univariate ones. An optimal set of coefficients is typically determined using a variational approach. *B*-splines and wavelets have desirable properties such as smoothness and can be constructed hierarchically. *B*-splines have less complexity than wavelets and are smooth and compactly supported.

It is always beneficial to use as much information as possible, so using additional landmarks should improve the accuracy of a non-landmark based method, provided the landmarks are accurately localized. It is possible to design basis functions that satisfy physical models, e.g., [102], [103]. Similarly, preserving the topology or ensuring the transformation is diffeomorphic are principled strategies and can be incorporated, for example, as a regularization constraint.

Generally, there is a lack of evaluation studies of nonrigid algorithms with different transformation models. Consequently, it is difficult to draw robust conclusions about which methods are most accurate or best suited to a particular application. New evaluation projects are being planned that can address this. Evaluation could play an important role in providing feedback when new continuum biomechanical models of *in vivo* soft-tissue deformation are proposed.

ACKNOWLEDGMENT

The author would like to thank the anonymous reviewers for providing the detailed comments. The author would also like to thank colleagues D. Popescu, R. Li, and G. Poulton for help with proofreading.

REFERENCES

[1] J. M. Fitzpatrick and M. Sonka, Handbook of Medical Imaging SPIE, Bellingham, WA, 2000, vol. PM80.

[2] 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.

[3] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Phys. Med. Biol.*, vol. 46, pp. R1–R45, 2001.

[4] H. Lester and S. R. Arridge, "A survey of hierarchical non-linear medical image registration," *Pattern Recognition*, vol. 32, no. 1, pp. 129–149, Jan. 1999.

[5] K. Rohr, *Landmark-Based Image Analysis: Using Geometric and Intensity Models*. Boston, MA: Kluwer, 2001.

[6] B. Zitova and J. Flusser, "Image registration methods: A survey," *Image Vision Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.

[7] J. Modersitzki, *Numerical Methods for Image Registration (Numerical Mathematics and Scientific Computation)*. New York: Oxford Univ. Press, 2004.

[8] A. Goshtasby, *2-D and 3-D Image Registration: For Medical, Remote Sensing, and Industrial Applications*. New York: Wiley-Interscience, 2005, ch. 5, Transformation functions.

[9] C. A. Glasbey and K. V. Mardia, "A review of image warping methods," *J. Appl. Stat.*, vol. 25, pp. 155–171, 1998.

[10] C. R. Maurer and J. M. Fitzpatrick, "A review of medical image registration," in *Interactive Image-Guided Neurosurgery*, R. J. Maciunas, Ed. Park Ridge, IL: American Association of Neurological Surgeons, 1993, pp. 17–44.

[11] P. A. van den Elsen, E.-J. D. Pol, and M. A. Viergever, "Medical image matching—A review with classification," *IEEE Eng. Med. Biol. Mag.*, vol. 12, no. 1, pp. 26–39, Mar. 1993.

[12] L. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, no. 4, pp. 325–376, Dec. 1, 1992.

[13] W. Crum, T. Hartkens, and D. Hill, "Non-rigid image registration: Theory and practice," *Br. J. Radiol.*, vol. 77, pp. S140–S153, 2004.

[14] C. R. Maurer, D. L. G. Hill, H. Y. Liu, M. McCue, D. Rueckert, D. Lloret, W. A. Hall, R. E. Maxwell, D. J. Hawkes, and C. L. Truwit, "Investigation of intraoperative brain deformation using a 1.5T interventional MR system: Preliminary results," *IEEE Trans. Med. Imag.*, vol. 17, no. 5, pp. 817–825, Oct. 1998.

[15] M. Ferrant, A. Nabavi, B. Macq, F. A. Jolesz, R. Kikinis, and S. K. Warfield, "Registration of 3D intraoperative MR images of the brain using a finite-element biomechanical model," *IEEE Trans. Med. Imag.*, vol. 20, pp. 1384–1397, 2001.

[16] T. Makela, P. Clarysse, O. Sipila, N. Pauna, Q. C. Pham, T. Katila, and I. E. Magnin, "A review of cardiac image registration methods," *IEEE Trans. Med. Imag.*, vol. 21, no. 9, pp. 1011–1021, Sep. 2002.

[17] A. F. Frangi, D. Rueckert, and J. A. Schnabel, "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling," *IEEE Trans. Med. Imag.*, vol. 21, no. 9, pp. 1151–1166, Sep. 2001.

[18] J. Gee, T. A. Sundaram, I. Hasegawa, H. Uematsu, and H. Hatabu, "Characterization of regional pulmonary mechanics from serial magnetic resonance imaging data," *Academic Radiol.*, vol. 10, no. 10, pp. 1147–1152, 2003.

[19] T. A. Sundaram, B. B. Avants, and J. C. Gee, "Towards a dynamic model of pulmonary parenchymal deformation: Evaluation of methods for temporal reparameterization of lung data," *Med. Imag., Comput. Computer-Assisted Intervention—MICCAI '05*, pp. 328–335, 2005.

[20] X. Papademetris, D. P. Dione, L. W. Dobrucki, L. H. Staib, and A. J. Sinusas, "Articulated rigid registration for serial lower-limb mouse imaging," *Med. Imag., Comput. Computer-Assisted Intervention—MICCAI '05*, vol. 3750, pp. 919–926, 2005.

[21] C. P. Behrenbruch, K. Marias, P. A. Armitage, M. Yam, N. Moore, R. E. English, J. Clarke, and M. Brady, "Fusion of contrast-enhanced breast MR and mammographic imaging data," *Med. Image Anal.*, vol. 7, no. 3, pp. 311–340, Sep. 2003.

[22] D. Mattes, D. R. Haynor, H. Vesselle, T. K. 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.

[23] Y. C. Tai, K. P. Lin, C. K. Hoh, S. C. H. Huang, and E. J. Hoffman, "Utilization of 3D elastic transformation in the registration of chest X-ray CT and whole body PET," *IEEE Trans. Nucl. Sci.*, vol. 44, no. 4, pp. 1606–1612, Aug. 1997.

[24] D. Rey, G. Subsol, H. Delinette, and N. Ayache, "Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis," *Med. Image Anal.*, vol. 6, no. 2, pp. 163–179, Jun. 2002.

[25] K. K. Leung *et al.*, "Automatic quantification of changes in bone in serial MR images of joints," *IEEE Trans. Med. Imag.*, vol. 25, no. 12, pp. 1617–1626, Dec. 2006.

[26] C. Studholme, V. Cardenas, K. Blumenfeld, N. Schuff, H. J. Rosen, B. Miller, and M. Weiner, "Deformation tensor morphometry of sematic dementia with quantitive validation," *NeuroImage*, vol. 21, no. 4, pp. 1387–1398, 2004.

[27] P. A. Freeborough and N. C. Fox, "Modeling brain deformations in alzheimer's disease by fluid registration of serial 3D MR images," *J. Comput. Assist. Tomogr.*, vol. 22, no. 5, pp. 838–843, Oct. 1998.

[28] M. Holden, J. A. Schnabel, and D. L. G. Hill, "Quantification of small cerebral ventricular volume changes in treated growth hormone patients using non-rigid registration," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, pp. 1292–1301, Oct. 2002.

[29] T. Hartkens, D. L. G. Hill, A. D. Castellano-Smith, C. R. Maurer, A. J. Martin, W. A. Hall, H. Liu, and C. L. Truwit, "Measurement and analysis of brain deformation during neurosurgery," *IEEE Trans. Med. Imag.*, vol. 22, no. 1, pp. 82–92, Jan. 2003.

[30] J. Kybic, P. Thévenaz, A. Nirkko, and M. Unser, "Unwarping of unidirectionally distorted EPI images," *IEEE Trans. Med. Imag.*, vol. 19, no. 2, pp. 80–93, Feb. 2000.

[31] C. Studholme, R. T. Constable, and J. S. Duncan, "Accurate alignment of functional EPI data to anatomical MRI using a physics-based distortion model," *IEEE Trans. Med. Imag.*, vol. 19, no. 11, pp. 1115–1127, Nov. 2001.

[32] M. Holden, M. Breeuwer, K. McLeish, D. J. Hawkes, S. F. Keevil, and D. L. G. Hill, "Sources and correction of higher order geometrical distortion for serial MR brain imaging," in *Proc. Soc. Photo-Optical Instrum. Eng. (SPIE)*, Bellingham, WA, 2001, vol. 4322, pp. 69–78.

[33] P. M. Thompson, D. MacDonald, M. S. Mega, C. J. Holmes, A. C. Evans, and A. W. Toga, "Detection and mapping of abnormal brain structure with a probabilistic atlas of cortical surfaces," *J. Comput. Assist. Tomogr.*, vol. 21, no. 4, pp. 567–581, 1997.

[34] P. M. Thompson, R. P. Woods, M. S. Mega, and A. W. Toga, "Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain," *Human Brain Mapp.*, vol. 9, no. 2, pp. 81–92, Feb. 2000.

[35] K. M. Pohl, J. Fisher, E. L. Grimson, R. Kikinis, and W. M. Wells, "A bayesian model for joint segmentation and registration," *NeuroImage*, vol. 31, no. 1, pp. 228–239, May 15, 2006.

[36] J. C. Mazziotta, A. W. Toga, A. Evans, P. Fox, and J. Lancaster, "A probabilistic atlas of the human brain—Theory and rationale for its development," *NeuroImage*, vol. 2, no. 2, pp. 89–101, Jun. 1995.

[37] A. C. Evans, D. L. Collins, P. Neelin, D. MacDonald, M. Kamber, and T. S. Marett, "Three-dimensional correlative imaging: Applications in human brain mapping," in *Functional Neuroimaging: Technical Foundations*, R. W. thatcher, M. Hallett, T. Zeffiro, and E. R. John, Eds. San Diego, CA: Academic, 1994, pp. 145–161.

[38] D. L. Collins, C. J. Holmes, T. M. Peters, and A. C. Evans, "Automatic 3D model-based neuroanatomical segmentation," *Human Brain Mapp.*, vol. 3, no. 3, pp. 190–208, 1995.

[39] P. M. Thompson and A. W. Toga, "Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations," *Med. Image Anal.*, vol. 1, pp. 271–294, 1997.

[40] J. T. Tournoux, *Co-Planar Stereotaxic Atlas of the Human Brain*. New York: Thieme Medical Publishers.

[41] B. M. Dawant, S. L. Hartmann, J. P. Thirion, F. Maes, D. Vandermeulen, and P. Demaerel, "Automatic 3D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformations: Part I, Methodology and validation on normal subjects," *IEEE Trans. Med. Imag.*, vol. 18, pp. 909–916, 1999.

[42] G. Calmon and N. Roberts, "Automatic measurement of changes in brain volume on consecutive 3D MR images by segmentation propagation," *Magn. Reson. Imag.*, vol. 18, no. 4, pp. 439–453, May 2000.

[43] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "A viscous fluid model for multimodal non-rigid image registration using mutual information," *Med. Image Anal.*, vol. 7, no. 4, pp. 565–575, Dec. 2003.

[44] A. W. Toga, K. Ambach, B. Quinn, M. Hutchin, and J. S. Burton, "Post-mortem anatomy from cryosectioned whole human brain," *J. Neurosci. Methods*, vol. 54, no. 2, pp. 239–252, Oct. 1994.

[45] M. A. Jacobs *et al.*, "Registration and warping of magnetic resonance images to histological sections," *Med. Phys.*, vol. 26, no. 8, pp. 1568–1578, Aug. 1999.

[46] P. C. Chou and N. J. Pagano, *Elasticity: Tensor, Dyadic, and Engineering Approaches*. New York: Dover, 1992.

[47] G. Mase, *Schaum's Outline of Continuum Mechanics*. New York: McGraw-Hill, 1970.

[48] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "3D brain mapping using a deformable neuroanatomy," *Phys. Medicine Biol.*, vol. 39, pp. 609–618, 1994.

[49] C. Broit, "Optimal registration of deformed images," Ph.D. dissertation, Dept. Comput. Inf. Sci., Univ. Pennsylvania, Philadelphia, 1981.

[50] R. Bajcsy, R. Lieberson, and M. Reivich, "A computerized system for the elastic matching of deformed radiographic images to idealized atlas images," *J. Comput. Assist. Tomogr.*, vol. 7, no. 4, pp. 618–625, 1983.

[51] J. C. Gee, D. R. Haynor, M. Reivich, and R. Bajcsy, "Finite element approach to warping of brain images," *SPIE Med. Imag.*, pp. 327–337, 1994.

[52] R. Bajcsy and S. Kovacic, "Multiresolution elastic matching," *Comput. Vision Graph. Image Process.*, vol. 46, no. 1, pp. 1–21, Apr. 1989.

[53] C. Davatzikos and R. N. Bryan, "Using a deformable surface model to obtain a shape representation of the cortex," *IEEE Trans. Med. Imag.*, vol. 15, no. 6, pp. 785–795, Dec. 1996.

[54] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Boston, MA: Johns Hopkins Univ. Press, 1996.

[55] C. Davatzikos, "Spatial transformation and registration of brain images using elastically deformable models," *Comput. Vision Image Understand.*, vol. 66, no. 2, pp. 207–222, May 1997.

[56] J. C. Gee, M. Reivich, and R. Bajcsy, "Elastically deforming 3D atlas to match anatomical brain images," *J. Comput. Assist. Tomogr.*, pp. 225–236, 1993.

[57] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Trans. Image Process.*, vol. 5, pp. 1435–1447, Oct. 1996.

[58] G. E. Christensen, S. C. Joshi, and M. I. Miller, "Volumetric transformation of brain anatomy," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 864–877, Dec. 1997.

[59] E. M. Lifshitz and L. D. Landau, *Fluid Mechanics*. New York: Butterworth-Heinemann, 1987.

[60] J. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, 1 ed. Pacific Grove, CA: Wadsworth, 1989.

[61] M. Bro-Nielsen and C. Gramkow, "Fast fluid registration of medical images," in *Proc. Visualization in Biomedical Computing*, R. Kikinis and K. Hoehne, Eds. Berlin, Germany: Springer Verlag, pp. 267–276.

[62] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C the Art of Scientific Computing*. Cambridge, U.K.: Cambridge Univ. Press, 1992.

[63] B. P. K. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, pp. 185–203, 1981.

[64] J.-P. Thirion, "Image matching as a diffusion process: An analogy with maxwell's demons," *Med. Image Anal.*, vol. 2, no. 3, pp. 243–260, Sep. 1998.

[65] C. Truesdell, W. Noll, and S. S. Antman, *The Non-Linear Field Theories of Mechanics*. New York: Springer, 2004.

[66] J. D. Humphrey, "Continuum biomechanics of soft biological tissues," *Proc. R. Stat. Soc.*, vol. 459, no. 2029, pp. 3–46, 2003.

[67] E. Meijering, "A chronology of interpolation: From ancient astronomy to modern signal and image processing," *Proc. IEEE*, vol. 90, pp. 319–342, 2002.

[68] L. L. Schumaker, *Spline Functions: Basic Theory*. Melbourne, FL: Krieger, 1993.

[69] M. J. D. Powell, *Approximation Theory and Methods*. Cambridge, U.K.: Cambridge Univ. Press, 1981.

[70] S. C. Joshi and M. I. Miller, "Landmark matching via large deformation diffeomorphisms," *IEEE Trans. Image Process.*, vol. 9, no. 8, pp. 1357–1370, Aug. 2000.

[71] S. Marsland and C. J. Twining, "Constructing diffeomorphic representations for the groupwise analysis of nonrigid registrations of medical images," *IEEE Trans. Med. Imag.*, vol. 23, no. 8, pp. 1006–1020, Aug. 2004.

[72] M. J. D. Powell, "Radial basis functions for multivariate interpolation: A review," in *Conf. 1st Algorithms For Approximation (Inst. Math. Its Appl. Conf. Series)*, 1987, vol. 1, pp. 143–167.

[73] M. J. D. Powell, "The theory of radial basis function approximation in 1990," in *Advances in Numerical Analysis*. New York: Oxford Sci. Publications, 1992, vol. 2, ch. 2.

[74] M. D. Buhmann, "Radial basis functions," *Acta Numerica*, vol. 9, pp. 1–38, 2000.

[75] N. Arad, N. Dyn, D. Reisfeld, and Y. Yeshurun, "Warping by radial basis functions—Application to facial expressions," *CVGIP-Graphical Models Image Process.*, vol. 56, no. 2, pp. 161–172, 1994.

[76] R. Detlef and M. Heinrich, *Free Form Deformation with Scattered Data Interpolation Methods*, G. Farin, H. Hagen, and H. Noltemeier, Eds. New York: Springer Verlag, 1993, vol. 8, Geometric modelling (computing suppl.), pp. 267–281.

[77] K. Rohr, M. Fornefett, and H. S. Stiehl, "Spline-based elastic image registration: Integration of landmark errors and orientation attributes," *Comput. Vision Image Understand.*, vol. 90, no. 2, pp. 153–168, May 2003.

[78] K. Rohr *et al.*, "Landmark-based elastic registration using approximating thin-plate splines," *IEEE Trans. Med. Imag.*, vol. 20, pp. 526–534, 2001.

[79] K. Rohr, H. S. Stiehl, R. Sprengel, W. Beil, T. M. Buzug, J. Weese, and M. H. Kuhn, "Point-based elastic registration of medical image data using approximating thin-plate splines," in *Visualization Biomed. Comput, Proc. Soc. Photo-Optical Instrum. Eng. (SPIE)*, Bellingham, WA, 1996, vol. 1131, pp. 297–306.

[80] M. Fornefett, K. Rohr, and H. S. Steihl, "Radial basis functions with compact support for elastic registration of medical images," *Image Vis. Comput.*, vol. 19, no. 1–2, pp. 87–96, 2001.

[81] D. N. Fogel, "Image rectification with radial basis functions: Application to RS/GIS data integration," in *Proc. 3rd Int. Conf. Integrating GIS Environ. Modell.*, Santa Fe, NM, 1996, pp. 19–119.

[82] A. Goshtasby, "Piecewise cubic mapping functions for image registration," *Pattern Recognit.*, vol. 20, no. 5, pp. 525–533, 1987.

[83] R. Wiemker, K. Rohr, L. Binder, R. Sprengel, and H. S. Stiehl, "Application of elastic registration to imagery from airborne scanners," in *Int. Archives Photogrammetry Remote Sensing*, 1996, vol. XXXI-B4, pp. 949–954.

[84] L. Zagorchev and A. Goshtasby, "A comparative study of transformation functions for nonrigid image registration," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 529–538, Mar. 2006.

[85] R. L. Harder and R. N. Desmarais, "Interpolation using surface splines," *J. Aircraft*, vol. 9, pp. 189–191, 1972.

[86] J. Duchon, "Interpolation des fonctions de deux variables suivant le principle de la flexion revue des plaques minces," *RAIRO Analyse Numerique*, vol. 10, no. 12, pp. 5–12, 1976.
[87] J. Meinguet, "Multivariate interpolation at arbitrary points made simple," *J. Appl. Math. Phys.*, vol. 30, pp. 292–304, 1979.
[88] M. F. Hutchinson and P. E. Gessler, "Splines—More than just a smooth interpolator," *Geoderma*, vol. 62, no. 1-3, pp. 45–67, 1994.
[89] D. E. Myers, "Spatial interpolation—An overview," *Geoderma*, vol. 62, no. 1-3, pp. 17–28, 1994.
[90] W. E. L. Grimson, *From Images to Surfaces: A Computational Study of the Human Early Visual System*. Cambridge, MA: MIT Press, 1981.
[91] D. Terzopoulos, "Multiresolution computation of visible-surface representations," Ph.D. dissertation, Dept. Electrical Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, 1984.
[92] F. L. Bookstein, "Principal warps: Thin-Plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567–585, Jun. 1989.
[93] A. Goshtasby, "Registration of images with geometric distortions," *IEEE Trans. Geosci. Remote Sens. E*, vol. 26, no. 1, pp. 60–64, Jan. 1988.
[94] F. L. Bookstein, "Shape and the information in medical images: A decade of the morphometric synthesis," *Comput. Vision Image Understand.*, vol. 66, no. 2, pp. 97–118, May 1997.
[95] R. K. Beatson and G. N. Newsam, "Fast evaluation of radial basis functions," *Comput. Math. Appl.*, vol. 24, pp. 7–19, 1992.
[96] I. Barrodale, D. Skea, M. Berkley, R. Kuwahara, and R. Poeckert, "Warping digital images using thin plate splines," *Pattern Recognit.*, vol. 26, pp. 375–376, 1993.
[97] L. Greengard and V. Rokhlin, "A fast algorithm for particle simulations," *J. Comp. Phys.*, vol. 73, pp. 325–348, 1987.
[98] M. J. D. Powell, Tabulation of thin plate splines on a very fine two-dimensional grid, Numerical analysis report of University of Cambridge Cambridge Univ., Cambridge, U.K., 1992.
[99] G. Wahba, *Spline Models for Observational Data*. Philadelphia, PA: SIAM, 1990.
[100] G. E. Christensen, "Deformable shape models for anatomy," Ph.D. dissertation, Dept. Elect. Eng., Sever Institute of Technology, Washington Univ., St. Louis, MO, 1994.
[101] H. Wendland, "Piecewise polynomial, positive definite and compactly supported radial basis functions of minimal degree," *Adv. Comput. Math.*, vol. 4, pp. 389–396, 1995.
[102] M. H. Davis, A. Khotanzad, D. P. Flaming, and S. E. Harms, "A physics-based coordinate transformation for 3D image matching," *IEEE Trans. Med. Imag.*, vol. 16, pp. 317–328, 1997.
[103] J. Kohlrausch, K. Rohr, and H. S. Stiehl, "A new class of elastic body splines for nonrigid registration of medical images," *J. Math. Imag. Vis.*, vol. 23, no. 3, pp. 253–280, 2005.
[104] S. W. K. Rohr, "Physics-based elastic image registration using splines and including landmark localization uncertainties," *Med. Imag., Comput. Computer-Assisted Intervention—MICCAI '06*, vol. 2, pp. 678–685, 2006.
[105] I. J. Schoenberg, "Contributions to the problem of approximation of equidistant data by analytic functions," *Q. Appl. Math.*, vol. 4, pt. A, pp. 45–99, 1946.
[106] M. Unser, A. Aldroubi, and M. Eden, "Fast b-spline transforms for continuous image representation and interpolation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 3, pp. 277–285, Mar. 1991.
[107] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part I—Theory," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 821–833, Feb. 1993.
[108] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part II—Efficient design and applications," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 834–848, Feb. 1993.
[109] P. Thevanaz, T. Blu, and M. Unser, "Interpolation revisited," *IEEE Trans. Med. Imag.*, vol. 19, no. 7, pp. 739–758, Jul. 2000.
[110] S. Lee, G. Wolberg, K. Y. Chwa, and S. Y. Shin, "Image metamorphosis with scattered feature constraints," *IEEE Trans. Visual. Comput. Graph.*, vol. 2, no. 4, pp. 337–354, Dec. 1996.
[111] E. Bardinet, L. D. Cohen, and N. Ayache, "A parametric deformable model to fit unstructured 3D data," *Comput. Vision Image Understand.*, vol. 71, no. 1, pp. 39–54, Jul. 1998.
[112] M. W. Hsu, J. F. Hughes, and H. Kaufman, "Direct manipulation of free-form deformations," in *Proc. SIGGRAPH*, 1992, vol. 26, pp. 177–184.
[113] G. Wolberg, "Image morphing: A survey," *The Visual Comput.*, vol. 14, no. 8-9, pp. 360–372, Dec. 9, 1998.
[114] S. Lee, G. Wolberg, K. Y. Chwa, and S. Y. Shin, "Image metamorphosis using snakes and free-form deformations," in *Proc. ACM SIGGRAPH'95*, Jul. 1995, pp. 439–448.
[115] P. Gao and T. W. Sederberg, "A work minimization approach to image morphing," *Visual Comput.*, vol. 14, no. 8-9, pp. 390–400, Dec. 9, 1998.
[116] J. M. Declerck, J. Feldmar, M. L. Goris, and F. Betting, "Automatic registration and alignment on a template of cardiac stress and rest reoriented SPECT images," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 727–737, Dec. 1997.
[117] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imag.*, vol. 18, no. 8, pp. 712–721, Aug. 1999.
[118] C. de Boor, *A Practical Guide to Splines*. New York: Springer-Verlag, 1978.
[119] J. H. Gallier, *Curves and surfaces in geometric modeling: Theory and algorithms*. San Francisco, CA: Morgan Kaufmann, 1999.
[120] I. Shoenberg, *Cardinal Spline Interpolation*. Philadelphia, PA: SIAM, 1973.
[121] T. Sederberg and S. Parry, "Free-form deformations of solid geometric models," *Comput. Graph.*, vol. 20, pp. 151–160, 1986.
[122] J. Kybic and M. Unser, "Fast parametric elastic image registration," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1427–1442, Nov. 2003.
[123] C. Sorzano, P. Thevenaz, and M. Unser, "Elastic registration of biological images using vector-spline regularization," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 4, pp. 652–663, Apr. 2005.
[124] J. Schnabel, C. Tanner, A. Castellano-Smith, A. Degenhard, M. Leach, D. Hose, D. Hill, and D. Hawkes, "Validation of nonrigid image registration using finite-element methods: Application to breast MR images," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 238–247, Feb. 2003.
[125] D. L. Collins and A. C. Evans, "Animal: Validation and applications of nonlinear registration-based segmentation," *Int. J. Pattern Recogn. Artif. Intell.*, vol. 11, no. 8, pp. 1271–1294, 1997.
[126] A. Cohen and J. Kovacevic, "Wavelets: The mathematical background," *Proc. IEEE*, vol. 84, no. 4, pp. 514–522, Apr. 1996.
[127] Y. Amit, "Nonlinear variational problem for image matching," *SIAM J. Comput.*, vol. 15, no. 1, pp. 207–224, Jan. 1994.
[128] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.
[129] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909–1906, Nov. 1988.
[130] S. Gefen, E. Tretiak, and J. Nissanov, "Elastic 3D alignment of rat brain histological images," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1480–1489, Nov. 2003.
[131] Y. T. Wu, T. Kanade, C. C. Li, and J. Cohn, "Image registration using wavelet-based motion model," *Int. J. Comput. Vis.*, vol. 38, no. 2, pp. 129–152, 2000.
[132] Y. T. Wu, L. F. Chen, P. L. Lee, T. C. Yeh, and J. C. Hsieh, "Discrete signal matching using coarse-to-fine wavelet basis functions," *Pattern Recognit.*, vol. 36, no. 1, pp. 171–192, Jan. 2003.
[133] W. Cai and J. Z. Wang, "Adaptive multiresolution collocation methods for initial boundary value problems of nonlinear PDEs," *SIAM J. Numerical Anal.*, vol. 33, pp. 937–970, 1996.
[134] P. Hellier, C. Barillot, E. Mémin, and P. Pérez, "Hierarchical estimation of a dense deformation field for 3D robust registration," *IEEE Trans. Med. Imag.*, vol. 20, no. 5, pp. 388–402, May. 2001.
[135] P. J. Huber, "Robust estimation of a location parameter," *Ann. Math. Stat.*, vol. 35, pp. 73–101, 1964.
[136] M. Black and A. Rangarajan, "On the unification of line processes, outlier rejection, and robust statistics with applications in early vision," *Int. J. Comput. Vision*, vol. 19, no. 1, pp. 57–91, Jul. 1996.
[137] E. Mémin and P. Pérez, "Dense estimation and object-based segmentation of the optical flow with robust techniques," *IEEE Trans. Image Process.*, vol. 7, no. 5, pp. 703–719, May 1998.
[138] G. E. Christensen and H. J. Johnson, "Consistent image registration," *IEEE Trans. Med. Imag.*, vol. 20, no. 7, pp. 568–582, Jul. 2001.
[139] H. J. Johnson and G. E. Christensen, "Consistent landmark and intensity-based image registration," *IEEE Trans. Med. Imag.*, vol. 21, no. 5, pp. 450–461, May 2002.
[140] O. Musse, F. Heitz, and J. P. Armspach, "Topology preserving deformable image matching using constrained hierarchical parametric models," *IEEE Trans. Image Process.*, vol. 10, no. 7, pp. 1081–1093, Jul. 2001.

[141] V. Noblet, C. Heinrich, F. Heitz, and J. P. Armspach, "3D deformable image registration: A topology preservation scheme based on hierarchical deformation models and interval analysis optimization," *IEEE Trans. Image Process.*, vol. 14, no. 5, pp. 553–566, May 2005.

[142] M. I. Miller, A. Trouve, and L. Younes, "On the metrics and Euler-Lagrange equations of computational anatomy," *Annu. Rev. Biomed. Eng.*, vol. 4, pp. 375–405, 2002.

[143] U. Grenander and M. I. Miller, "Representations of knowledge in complex systems," *J. R. Statist. Soc. B*, vol. 56, no. 3, pp. 549–603, 1994.

[144] U. Grenander, *General pattern theory*. New York: Oxford Univ. Press, 1994.

[145] U. Grenander and M. I. Miller, "Computational anatomy: An emerging discipline," *Q. Appl. Math.*, vol. 56, no. 4, pp. 617–694, 1998.

[146] A. Trouve, "Diffeomorphisms groups and pattern matching in image analysis," *Int. J. Comput. Vis.*, vol. 28, no. 3, pp. 213–221, 1998.

[147] M. F. Beg, M. I. Miller, A. Trouve, 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.

[148] P. Hellier, C. Barillot, I. Corouge, B. Gibaud, G. Le Goualher, D. L. Collins, A. Evans, G. Malandain, N. Ayache, G. E. Christensen, and H. J. Johnson, "Retrospective evaluation of intersubject brain registration," *IEEE Trans. Med. Imag.*, vol. 22, no. 9, pp. 1120–1130, Sep. 2003.

[149] G. E. Christensen, X. Geng, J. G. Kuhl, J. Bruss, T. J. Grabowski, I. A. Pirwani, M. W. Vannier, J. S. Allen, and H. Damasio, "Introduction to the non-rigid image registration evaluation project (NIREP)," in *Biomed. Image Registration, Proc. LNCS*, 2006, vol. 4057, pp. 128–135.

[150] 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.

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