Mapping of hyperelastic deformable templates using the finite element method

R. D. Rabbitt
Department of Bioengineering, University of Utah, Salt Lake City, UT

J.A. Weiss
Orthopedic Biomechanics Institute
Department of Bioengineering, University of Utah, Salt Lake City, UT
and
Applied Mechanics Group
Lawrence Livermore National Laboratories, Livermore, CA

G.E. Christensen
Malincrot Inst. of Radiology, Washington University School of Medicine, St. Louis, MO

M.I. Miller
Department of Electrical Engineering, Washington University, St. Louis, MO

June 10, 1995

ABSTRACT

In the current work we integrate well established techniques from finite deformation continuum mechanics with concepts from pattern recognition and image processing to develop a new finite element (FE) tool that combines image-based data with mechanics. Results track the deformation of material continua in the presence of unknown forces and/or material properties by using image-based data to provide the additional required information. The deformation field is determined from a variational problem that combines both the mechanics and models of the imaging sensors. A nonlinear FE approach is used to approximate the solution of the coupled problem. Results can be applied to i) track the motion of deforming material and/or, ii) morphological warping of template images or patterns. Two-dimensional (2D) example results are provided for problems of the second type. One of the present examples was motivated primarily by a problem in medical imaging – mapping intersubject geometrical differences in human anatomical structures – with specific results given for the mapping 2D slices of the human distal femur based on X-ray computed tomographic (CT) images.

keywords: warping, morphing, continuum mechanics, mesh generation, global geometric models

0-8194-1932-X/95/$6.00

1 INTRODUCTION

It has been argued that the human visual cortex is capable of constructing an innumerable variety of imaginary topographical scenes on the basis of warping or deforming template images recalled from memory. Although it has not been proven that this mechanism is actually employed by the brain, the bio-based concept of "deformable templates" has spawned a branch of research in machine-based image processing. Some previous applications in biology address the recognition and tracking of physical matter including plants, hands, cells, cellular organelles and anatomical structures of the brain[1–6] The present work builds upon this idea to describe the deformable template as a three-dimensional material continuum from which numerous images can be queried using a variety of imaging modalities. The set of all images that can be sampled from the deformed template is therefore determined by the material configuration space admissible by the laws of continuum mechanics. To apply the method a geometrical model of the material continuum is constructed in a reference configuration. The deformable template is defined with the addition of material properties, which may be both inhomogeneous and anisotropic In the present work we employ an objective hyperelastic material model valid for large-magnitude nonlinear deformations. The deformed configuration of the template is then determined by driving the deformation using the difference between images, interrogated from the deformed material and the original template configuration. This yields a variational problem which is solved using a three-dimensional finite-element (FE) approach. The method is illustrated through three example problems, two of which are based on simulated image data and one of which is based on CT images of the human distal femur.

Although the motivation for the current work comes primarily from biomechanics and the inter-subject registration of anatomical structures, the concepts are broad and apply equally well to a variety of deforming continua. Applications of deformable templates can generally be divided into two categories: imaginary deformations and real deformations. For physical deformations the goal is to quantify both the kinematics and the kinetics of the deformation, but for imaginary deformations the interest is only in the kinematics of the warping. The present results are derived from the point of view of real deformations but are also relevant to imaginary deformations. For example, we previously applied fluid mechanics to the intersubject registration of human neuroanatomies, based on MRI data, and to the registration of macaque monkey visual cortex based on optical cryosection data[3,4] This imaginary deformation treats the deforming template as a low Reynold's number flow and allows for very large distance deformations to be realized. For some imaginary deformation applications however, the fluid-based warping is somewhat underconstrained and admits to unreasonable mappings mainly because very large volumes of matter may stream through very small areas from one region to another. The hyperelastic model presented here bypasses this problem by introducing hyperelastic solid, which tends to penalize the aforementioned type of deformation.

Incorporation of hyperelasticity is also relevant to real deformations where we would like to quantify the three-dimensional (3D) stress field associated with the reconstructed deformation field. The stress field predicted by the algorithm will be reasonable only if the constitutive law is reasonable and hence it is important to incorporate problem specific material descriptions. Hyperelastic constitutive models have been shown to be appropriate to represent the material behavior of soft tissues such as ligaments and tendons.[7,8] When combined with relaxation spectra, these laws define hyper-viscoelastic materials which provide reasonable descriptions of the behavior of numerous biological fluids and solids.[9] The theory of linear elasticity is not adequate for the present work because it assumes small strains and validity of Hooke's law – both of which are invalid for large magnitude material deformations.[10] The present application requires use of the full nonlinear finite-deformation theory. For the hyperelastic materials considered here the stress is derived from a potential energy density. This energy is written as a function of the strain tensor invariants and its form is subsequently reduced as appropriate for the particular material symmetry.

The following sections begin with the mathematical foundations necessary to combine nonlinear continuum mechanics with hyperelastic material models and with probabilistic imaging sensor models. A description of the finite element implementation is also provided, followed by two-dimensional (2D) example problems.

2 THEORY

2.1 Hyperelastic energy density

Employing notation from finite deformation theory,[11,12] we define a material continuum $\wp$ with admissible configurations $\chi : \wp \rightarrow \Re^3$. The location of an individual point $\mathcal{P}$ within the reference configuration is denoted by the vector $\mathbf{X} = \chi_o(\mathcal{P}) \subset \Re^3$, and the same point in the current configuration is denoted by the vector $\mathbf{x} = \chi(\mathcal{P}) \subset \Re^3$. Lower-case letters denote quantities associated with the spatial or current configuration and upper-case letters denote quantities associated with the reference configuration. The mapping $\varphi = \chi \circ \chi_o^{-1}$ defines the deformation from the reference configuration to the current configuration. In the present work we use the same orthonormal bases to describe both the reference and current configurations such that the mapping has the deformation gradient

$$\mathbf{F}(\mathbf{X}) = \frac{\partial \varphi}{\partial \mathbf{X}}, \tag{2.1}$$

which is a rank-2 tensor. The admissible configuration space of the deformed continuum is prescribed as

$$\mathcal{C} := \{\varphi : \chi_o(\wp) \rightarrow \chi(\wp) | J := \det[\mathbf{F}] > 0\}, \tag{2.2}$$

where the Jacobian J relates differential volumes in the reference and current configuration according to $dv = JdV$. The requirement $\det[\mathbf{F}] > 0$ insures a one-to-one mapping of differential material volumes consistent with the continuity of mass.

To relate the deformation to the stress, we decompose the deformation gradient using the polar decomposition theorem into a proper rotation $\mathbf{R}(\mathbf{X})$ and a right stretching tensor $\mathbf{U}(\mathbf{X})$. This gives $\mathbf{F} = \mathbf{RU}$. The rotation $\mathbf{R}(\mathbf{X})$ is not associated with deformation of the material but rather can be viewed as a local rigid-body motion.

For the hyperelastic constitutive behavior considered here the stress tensor is taken to be independent of path and hence derivable from a potential function W. This potential function depends only on the instantaneous stretch $\mathbf{U}$ and not on the rate or history of the deformation. To define W as a function of $\mathbf{U}$ in a form that avoids actual computation of the polar decomposition, we introduce the right Cauchy-Green strain tensor,

$$\mathbf{C} = \mathbf{F}^T\mathbf{F} = \mathbf{U}^2. \tag{2.3}$$

The potential function, or stored energy density function, is therefore written in the form

$$\hat{W} = W\left(\mathbf{X}, \mathbf{C}\left(\varphi(\mathbf{X})\right)\right). \tag{2.4}$$

This energy density is valid for large-magnitude nonlinear deformations but is limited to the fraction of the total energy stored by deformation of the hyperelastic component of the material. It does not include any other energy densities such as that associated with applied loads or any deviation from the hyperelastic material model. In the present work we are interested in problems where these other stresses and associated energy densities are not known, but rather are implied through the collection of experimental data. This data is assumed to be in the form of images that provide information about the physics of the deforming material.

2.2 Image data energy density

It is assumed that both the reference and current configurations are observed by an imaging sensor(s) that interrogate physical properties or fields associated with the material (i.e. such as hydrogen content, refractive index, density, etc.). These material properties are convected by the deformation and in some cases modified by the deformation. A tensor field $\hat{\mathbf{Q}}$ is defined to contain the values of the spatio-temporal distribution of the interrogated properties. It is further assumed that diffusion of $\hat{\mathbf{Q}}$ (thermal or electro-chemical) is negligible in

the current implementation. Under the mapping φ, the reference field of physical properties is transformed such that $\hat{\mathbf{Q}} = \mathbf{Q}(\mathbf{X}, \varphi, \mathbf{F})$. Experimental observations of the reference and transformed fields, $\mathbf{Q}_o$ and $\mathbf{Q}$ respectively, are employed to provide the additional information necessary to supplement the hyperelastic energy density and to complete the formulation of the problem.

Image-based data is used to query the continuum and obtain estimates of the physical properties contained in $\mathbf{Q}$. It is not possible to obtain exact measurements of the tensor field $\mathbf{Q}$ and hence probablistic models are employed. Such descriptions have been established for numerous imaging sensors including magnetic resonance scanners, X-ray computed tomography, optical sectioning microscopes, and satellite-based sensors.[13–16] Consistent with these models we assume that the probability density distributions are conservative in nature such that they can be derived from a potential function and cast in Gibb's energy density form. The derivation of the potential function follows the classical Boltzman approach.[17] We assume here that the from of the sensor potential function U is known for the particular sensing modality and can be written in the form $U = U(\mathbf{Q}(\mathbf{X}, \varphi, \mathbf{F}))$. In this form, the properties contained in $\mathbf{Q}$ are defined on the continuum. Including the tensor field $\mathbf{Q}$ implicitly, this can be written as

$$U = U(\mathbf{X}, \varphi, \mathbf{F}). \tag{2.5}$$

We shall henceforth refer to this potential function as the image-based energy density or the likelihood energy density. To be consistent, this energy density must be written in the same units as the stored energy function W. As described below, the energy density U may be defined from image-based data.

It is interesting to note that the form of the energy density U is precisely that defined in the field of mechanics from *live conservative body forces*. This arises owing to the fact that the total potential energy is a path independent functional that can be written in terms of the initial configuration and the instantaneous deformation. An example of a live, conservative body force is pressure loading, since the orientation of the applied force depends on the surface normal, which is in term dependent on the deformation.

2.3 The boundary value problem

We construct an energy functional consisting of both the hyperelastic energy density W and the likelihood energy density U. This direct approach facilitates the solution via the FE numerical method as detailed in later sections. The total energy functional is

$$E(\varphi) = \int_{\mathcal{B}} W(\mathbf{X}, \mathbf{C}(\varphi(\mathbf{X})))dV - \int_{\mathcal{B}} U(\mathbf{X}, \varphi(\mathbf{X}), \mathbf{F}(\mathbf{X}))dV, \tag{2.6}$$

where $\varphi(\mathbf{X})$ is the mapping to the current configuration as described above and $\mathcal{B} := \chi_o(\varphi)$ denotes the reference configuration of the material.

The weak form of the boundary value problem follows from the extremum of $E(\varphi)$ as obtained from the first variation with respect to φ. This is denoted $G(\varphi; \eta_0)$ and is computed as the Gâteaux (directional) derivative of $E(\varphi)$ in the direction η_0, denoted DE:

$$G(\varphi; \eta_0) = DE(\varphi) \cdot \eta_0 \tag{2.7}$$

The vector field η_0 contains the variations in the coordinates of each material point $\mathcal{P}$ with respect to the reference configuration. Performing the variation

$$G(\varphi; \eta_0) = \int_{\mathcal{B}} \mathbf{S} : \nabla_0^s \eta_0 \mathbf{F} dV - \int_{\mathcal{B}} \frac{\partial U}{\partial \varphi} \cdot \eta_0 dV, \tag{2.8}$$

where we have introduced the rank-2 2nd Piola-Kirchhoff stress tensor

$$\mathbf{S} = 2\frac{\partial W(\mathbf{X}, \mathbf{C})}{\partial \mathbf{C}}. \tag{2.9}$$

We have also used

$$D(\mathbf{C}) \cdot \boldsymbol{\eta}_0 = 2\nabla_0^s \boldsymbol{\eta}_0 \mathbf{F} \tag{2.10}$$

and

$$\nabla_0^s \boldsymbol{\eta}_0 = \frac{1}{2}(\nabla_0 \boldsymbol{\eta}_0 + (\nabla_0 \boldsymbol{\eta}_0)^T), \tag{2.11}$$

where ∇_0^s denotes the gradient using derivatives taken with respect to the reference, or material, coordinates. Absence of the subscript "0" denotes the spatial gradient operator and is written ∇.

$$G(\boldsymbol{\varphi}; \boldsymbol{\eta}) = \int_{\boldsymbol{\varphi}(\mathcal{B})} \boldsymbol{\sigma} : \nabla^s \boldsymbol{\eta} dv - \int_{\boldsymbol{\varphi}(\mathcal{B})} \frac{\partial U}{\partial \boldsymbol{\varphi}} \cdot \boldsymbol{\eta} \frac{dv}{J}, \tag{2.12}$$

where we have used the definition of the Jacobian given above and introduced the rank 2 Cauchy stress tensor $\boldsymbol{\sigma}$ as

$$\boldsymbol{\sigma} = \frac{1}{J}\mathbf{F}\mathbf{S}\mathbf{F}^T = \frac{2}{J}\mathbf{F}\frac{\partial W}{\partial \mathbf{C}}\mathbf{F}^T. \tag{2.13}$$

When set to zero, equations (2.8) and (2.12) provide the weak form of the Euler-Lagrange equations written in the material and spatial frames respectively. These are nonlinear equations which we solve via the finite element method as outlined below.

2.4 Local linearization

In order to determine the solution of the nonlinear Euler-Lagrange equations we solve a sequence of linear problems, each of which is obtained for small variations in the displacement field from a previously estimated configuration. Expanding (2.12) in a Gâteau-Taylor series and truncating higher order terms gives the linearized Euler-Lagrange equations in the form

$$LG(\boldsymbol{\varphi}; \boldsymbol{\eta}_0) = G(\boldsymbol{\varphi}_0; \boldsymbol{\eta}_0) + DG(\boldsymbol{\varphi}_0; \boldsymbol{\eta}_0) \cdot \mathbf{u}_0, \tag{2.14}$$

where $\mathbf{u}_0 = \boldsymbol{\varphi} - \boldsymbol{\varphi}_0$ is the variation from the configuration $\boldsymbol{\varphi}_0$. The Gâteaux derivative appearing in this expression is

$$D(\boldsymbol{\varphi}; \boldsymbol{\eta}_0) \cdot \mathbf{u}_0 = 0, \tag{2.15}$$

which expands to

$$\int_{\mathcal{B}} \nabla_0 \mathbf{u}_0 : \mathbf{S} : \nabla_0 \boldsymbol{\eta}_0 dV + \int_{\mathcal{B}} \mathbf{F} \cdot \mathsf{C} \cdot \nabla_0^s \mathbf{u}_0 \mathbf{F} : \nabla_0^s \boldsymbol{\eta}_0 dV - \int_{\mathcal{B}} \mathbf{u}_0 : \frac{\partial^2 U}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}} : \boldsymbol{\eta}_0 dV = 0. \tag{2.16}$$

The rank 4 tensor C appearing here is the material version of the elasticity tensor,[11]

$$\mathsf{C} = 4\frac{\partial^2 W}{\partial \mathbf{C} \partial \mathbf{C}} = 2\frac{\partial \mathbf{S}}{\partial \mathbf{C}}. \tag{2.17}$$

For the finite element implementation we write equations (2.16) and (2.17) in the spatial configuration to obtain

$$D(\boldsymbol{\varphi}; \boldsymbol{\eta}) \cdot \mathbf{u} = 0, \tag{2.18}$$

which expands to

$$\int_{\boldsymbol{\varphi}(\mathcal{B})} \nabla^s \mathbf{u} : \boldsymbol{\sigma} : \nabla^s \boldsymbol{\eta}_0 dv + \int_{\boldsymbol{\varphi}(\mathcal{B})} \nabla^s \mathbf{u} : \mathbf{C} : \nabla^s \boldsymbol{\eta} dv - \int_{\boldsymbol{\varphi}(\mathcal{B})} \mathbf{u} : \frac{\partial^2 U}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}} : \boldsymbol{\eta} dv = 0. \tag{2.19}$$

To obtain this we have used

$$\nabla_0 \boldsymbol{\eta}_0 = \nabla \boldsymbol{\eta} \mathbf{F}, \tag{2.20}$$

$$\nabla_0 \mathbf{u}_0 = \nabla \mathbf{u} \mathbf{F}, \tag{2.21}$$

and introduced the spatial version of the material elasticity tensor $\mathbf{C}$,[11]

$$\mathbf{C} := \frac{1}{J} \varphi_* \mathbf{C} \iff \mathsf{c}_{ijkl} = \frac{1}{J} F_{Ii} F_{Jj} F_{Kk} F_{Ll} \mathsf{C}_{IJKL}. \tag{2.22}$$

Combining Eq. (2.19) with Eq. (2.12) we obtain the linearized version of the weak form Euler-Lagrange equations

$$\int_{\varphi(\mathcal{B})} \nabla \mathbf{u} \boldsymbol{\sigma} : \nabla \boldsymbol{\eta} dv + \int_{\varphi(\mathcal{B})} \nabla^s \mathbf{u} : \mathbf{C} : \nabla^s \boldsymbol{\eta} dv - \int_{\varphi(\mathcal{B})} \mathbf{u} : \frac{\partial^2 U}{\partial \varphi \partial \varphi} : \boldsymbol{\eta} \frac{dv}{J}$$
$$= \int_{\varphi(\mathcal{B})} \frac{\partial U}{\partial \varphi} \cdot \boldsymbol{\eta} \frac{dv}{J} - \int_{\varphi(\mathcal{B})} \boldsymbol{\sigma} : \nabla^s \boldsymbol{\eta} dv. \tag{2.23}$$

Three terms in this expression result from the stored energy density function W. These mechanical terms define the initial stress or geometric stiffness (first term), the material stiffness (second term), and the internal force from the material stress (last term). The remaining two terms are associated with the image energy density function U. The fourth term defines external applied loads per unit volume (i.e. live body forces) that arise from the image energy density U which serve to drive the deformation of the material. In addition to providing the applied loads, the energy density U also contributes to the stiffness through the appearance of the third term – the body force stiffness.

2.5 Specialization for a Gaussian sensor model

Example results provided below are for the special case of a Gaussian sensor model such that the likelihood energy density U is a quadratic form. It is further assumed that the sensor data is a function only of the pointwise convected material properties and is independent of the deformation gradient. Under these additional assumptions the image energy density has the form

$$U(\mathbf{X}, \varphi(\mathbf{X})) = \frac{1}{2} \left(\mathbf{T}_o(\mathrm{X}) - \mathbf{T}(\varphi(\mathrm{X}))\right)^{\mathrm{T}} \Lambda \left(\mathbf{T}_o(\mathrm{X}) - \mathbf{T}(\varphi(\mathrm{X}))\right) \tag{2.24}$$

where $\mathbf{T}_o(X)$ and $\mathbf{T}(\varphi(X))$ are vector fields containing image data queried from the reference and deformed configurations respectively. Each element of $\mathbf{T}_o$ and $\mathbf{T}$ are piecewise continuous scalar fields constructed from discrete image data. The length of these vectors is determined by the number and type of imaging modalities employed. The above form assumes that a Gaussian model is appropriate for each imaging modality such that the variance associated with the likelihood of each is assembled into the matrix $\mathbf{\Lambda}$. Present example results assume $\mathbf{\Lambda}$ is diagonal such that

$$U(\mathbf{X}, \varphi(\mathbf{X})) = \sum_{m=1}^{M} U_m = \sum_{m=1}^{M} \frac{\lambda_m}{2} \left(T_{0m}(\mathbf{X}) - T_m(\varphi(X))\right)^2, \tag{2.25}$$

where U_m is the contribution to the image energy density by each modality and M is the total number of image pairs queried from the original and the deformed configurations. For this specific image sensor model the first derivative is

$$\frac{\partial U}{\partial \varphi} = \sum_{m=1}^{M} -\lambda_m \left(\mathrm{T}_{0\mathrm{m}}(\mathbf{X}) - \mathrm{T}_{\mathrm{m}}(\varphi(\mathbf{X}))\right) \frac{\partial \mathrm{T}_{\mathrm{m}}(\varphi(\mathbf{X}))}{\partial \varphi}. \tag{2.26}$$

This term determines the relationship between the image data and the body force driving the deformation. Note that Eq. (2.24) is equal to zero when images queried from the reference configuration $T_{0m}(\mathbf{X})$ are equal to images queried from the deformed configuration $T_m(\varphi(\mathbf{X}))$. The driving body force is also zero if the gradient term

$\partial T_m over \partial \varphi$ in the deformed configuration is 0, if the data are homogeneous. In this case the the image energy density is insensitive to small perturbations in the configuration, and hence the image data does not drive the deformation when the derivative is zero.

The second derivative of the image energy density contributes to the tangent stiffness and is provided as

$$\frac{\partial^2 U}{\partial \varphi \partial \varphi} = \sum_{m=1}^{M} -\lambda_m \left[(T_{0m}(\mathbf{X}) - T_m(\varphi(\mathbf{X}))) \frac{\partial^2 T_m}{\partial \varphi \partial \varphi} - \frac{\partial T_m}{\partial \varphi} \otimes \frac{\partial T_m}{\partial \varphi} \right]. \tag{2.27}$$

2.6 Finite element approximation

We find an approximate solution of Eq. (2.12) by solving a series of linear problems defined by Eq. (2.23). This is achieved via a finite element approximation. Following a procedure that is standard in the field of computational mechanics,[18] the continua is discretized into finite elements. The unknown variations in configuration are interpolated over each finite element using a set of basis functions that specify the polynomial order of the interpolation. We introduce an isoparametric conforming finite element approximation of the space of admissible variations η in (2.23). These variations are defined in the standard manner as:

$$\eta_e \equiv \eta|_{\varphi(\Omega_e)} = \sum_{k=1}^{N_{\text{nodes}}} N_k(\boldsymbol{\xi}) \eta_k, \quad \eta_k \in \Re^3, \tag{2.28}$$

where the subscript e specifies that the variations are restricted to a particular element, and N_{nodes} is the number of nodes composing each element. Here $\boldsymbol{\xi} \in \square$, where $\square := \{(-1,1) \times (-1,1) \times (-1,1)\}$ is the bi-unit cube, N_k are the isoparametric shape functions (having a value of "1" at their specific node and varying to "0" at every other node), and

$$\mathbf{X}_e = \sum_{k=1}^{N_{\text{nodes}}} N_k(\boldsymbol{\xi}) \mathbf{X}_k \tag{2.29}$$

is the isoparametric map. The deformed configuration of an element is then computed using the shape functions as

$$\varphi_e = \mathbf{X}_e + \mathbf{U}_e = \sum_{k=1}^{N_{\text{nodes}}} N_k(\boldsymbol{\xi})(\mathbf{X}_k + \mathbf{U}_k), \tag{2.30}$$

where $\mathbf{U}_e$ is the discrete displacement field restricted to Ω_e. The discrete deformation gradient $\mathbf{F}_e$ then follows:

$$\mathbf{F}_e = \frac{\partial \varphi_e}{\partial \mathbf{X}} = \sum_{k=1}^{N_{nodes}} (\mathbf{X}_k + \mathbf{U}_k) \otimes \nabla_0 N^k \tag{2.31}$$

Using (2.28), the nonsymmetric and symmetric gradient terms in (2.23) can be written in terms of the finite element discretization as

$$\nabla^s \eta = \sum_{k=1}^{N_{nodes}} \mathbf{B}_k^L \eta_k, \tag{2.32}$$

and

$$\nabla \eta = \sum_{k=1}^{N_{nodes}} \mathbf{B}_k^{NL} \eta_k, \tag{2.33}$$

where $\mathbf{B}^L$ and $\mathbf{B}^{NL}$ are the linear and nonlinear strain-displacement matrices, respectively.[18] With these results in hand, all unknowns in (2.23) can be approximated in terms of the finite element approximation for the displacements and configuration. This allows the equations to be represented as a linearized system of equations that are to be solved iteratively for the unknown displacement increments.

Substitution into (2.23) gives the following linearized matrix equation for the incremental displacements $\mathbf{u}$

$$\left[\int_{\varphi(\mathcal{B})} (\mathbf{B}^{NL})^T \boldsymbol{\sigma} \mathbf{B}^{NL} dv + \int_{\varphi(\mathcal{B})} (\mathbf{B}^L)^T \mathbf{C}\, \mathbf{B}^L dv - \int_{\varphi(\mathcal{B})} \mathbf{N}^T \mathbf{k}\, \mathbf{N} \frac{dv}{J}\right] \mathbf{u} \tag{2.34}$$
$$= \int_{\varphi(\mathcal{B})} \mathbf{N}^T \frac{\partial U}{\partial \varphi} \frac{dv}{J} - \int_{\varphi(\mathcal{B})} \mathbf{B}_L^T \boldsymbol{\sigma} dv,$$

where the second rank tensor $\mathbf{k}$ is the finite element approximation of the third term in Eq. (2.23). In the above it is assumed that the appropriate Voight representations are used for $\boldsymbol{\sigma}$, $\mathbf{C}$ and $\mathbf{k}$.[18]

Assuming that the solution at a configuration φ_0 is known, we seek the solution at some increment $\varphi_0 + \mathbf{u}$. The staring point is the linearization of (2.23) at φ_0 to get an initial estimate for $\mathbf{u}$. This estimate for $\mathbf{u}$ is then iteratively improved using a Newton (or quasi-Newton) method. The linearization, on the element level, a system of linear algebraic equations of the form

$$\sum_{i=1}^{N_{\text{nodes}}} \sum_{j=1}^{N_{\text{nodes}}} \left({}^M\mathbf{K}(\varphi_0) + {}^G\mathbf{K}(\varphi_0) + {}^W\mathbf{K}(\varphi_0)\right)_{ij} \mathbf{u} \;=\; \sum_{i=1}^{N_{\text{nodes}}} \left({}^{ext}\mathbf{F} - {}^{int}\mathbf{F}(\varphi_0)\right)_i . \tag{2.35}$$

The term ${}^M\mathbf{K}$ represents the material stiffness matrix, while the term ${}^G\mathbf{K}$ is the geometric (initial stress) stiffness matrix. These terms both arise in a traditional displacement-based nonlinear analysis. The warping stiffness, ${}^W\mathbf{K}$, is a direct result of the deformation dependent body forces arising from the fact that the warping potential changes with configuration (a deformation-independent body force such as gravity would not have a stiffnes contribution. The warping body forces are applied with other external forces via the vector ${}^{ext}\mathbf{F}$. $\mathbf{u}$ is the initial guess at the increment to the unknown configuration φ_{n+1}. By solving for $\mathbf{u}$, the configuration at $n+1$ is approximated as

$$\varphi_{n+1} \;=\; \varphi_0 + s\mathbf{u}. \tag{2.36}$$

Here s is a scalar parameter between 0 and 1 determined by a line search. The determination of an accurate value for φ_{n+1} follows by iterative solution using a Newton (or quasi-Newton) strategy.[19]

The above FE discretization of the variational principle (2.35) was implemented into NIKE3D, a nonlinear, implicit finite element code for problems in solid and structural mechanics. NIKE3D was developed and is maintained by the Methods Development Group at Lawrence Livermore National Laboratory.[20] The code was modified to allow the addition of the warping body force and stiffness. These quantities were computed using the template and study image data, read in at runtime. For speed of computation, a trilinear interpolation was used for the displacements, dictating the use of 8-node hexahedral elements in three dimensions. Because of the linear interpolation for displacement over each element, the first derivative terms in the warping force stiffness were constant on an element. Additionally, the second derivative terms in the warping force stiffness were zero. A number of simple 1-D and 2-D test problems were used to validate the implementation, several of which are described in the following section. All analyses were carried out on a Silicon Graphics.

3 2D EXAMPLE RESULTS

The first example discretizes the rectangular domain of a 2D template hyperelastic sheet using a 64x64 (4k elements) finite element mesh originally coincident with discretization of simulated image data. In the second

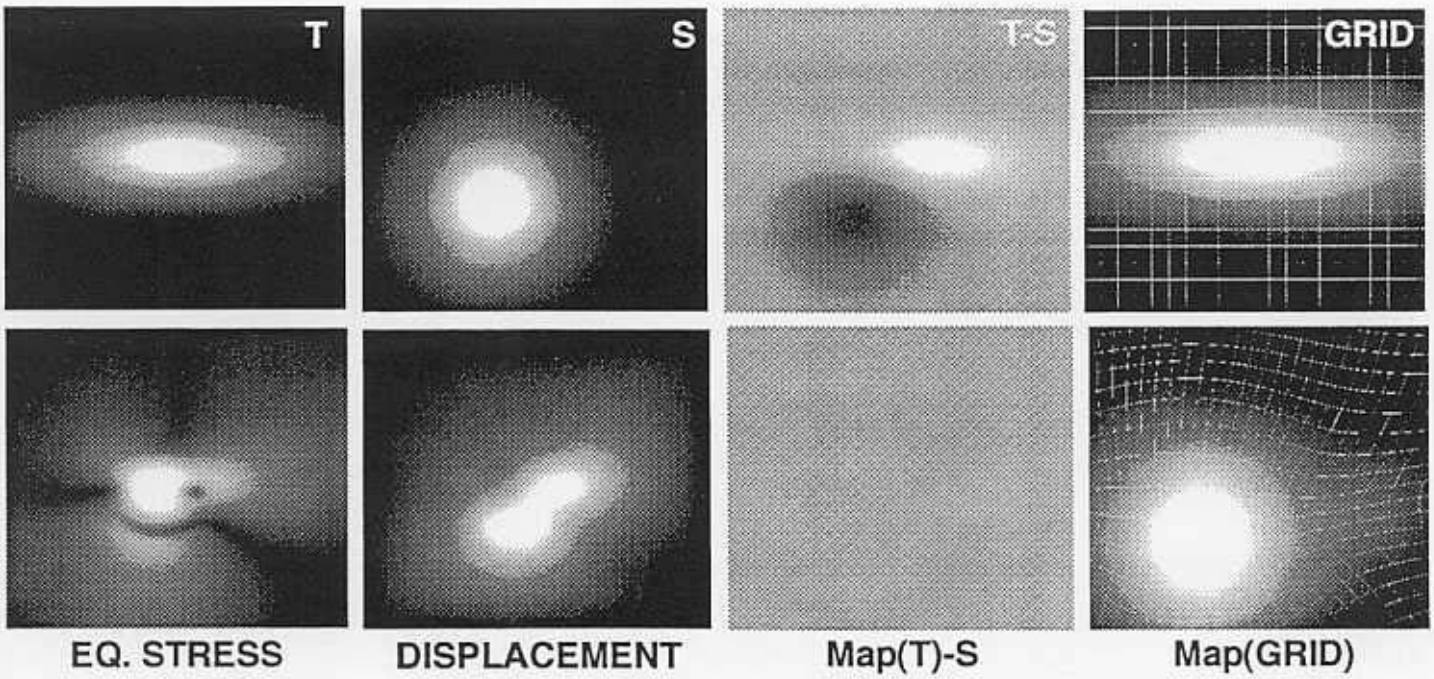

Figure 1: Two-dimensional Gaussian blur. The deformation computed using 64x64, 8bit, grayscale images interpolated using the FE shape functions. The first two columns of the torp row show the template (T) and the specific case (S). The difference image before and after deformation are also shown are shown in the first and second rows of the third column respectively. The bottom row also shows the equivalent stress and the magnitude of the in-plane displacement. A magnified section of the of the FE grid is shown in both the undeformed and deformed configurations in the 4th column.

example, X-ray computed tomographic images of the distal section of the human femur are used to align a template model with a specific subject. using uniform 128x128 (16k elements) rectangular mesh. In the final example a simple rectangular beam is bent using a computaional grid independent from the image discretization.

3.1 Example 1: Gaussian blur

Fig. 1 shows the deformation of a 2D square hyperelastic sheet driven by simulated image data. The intensities shown in the images simulate 8bit, 64x64 pixelized images. The template is associated with a Gaussian intensity blur having unequal spatial variance in the two principle directions. A uniform computational grid is overlayed on this image space. The grid consists of 4096, 3D, 8-node hexahedral elements constrained to deform in a single plane. In this particular example the images have the same spatial resolution as the FE grid, but the present implementation is not limited to this and it is often advantageous to use a FE grid distinct from the image data. The unstressed sheet is modeled as a homogeneous-isotropic material from which the images and are sampled in the reference and deformed configurations respectively. Fixed boundary conditions are employed around the outside periphery of the rectangular domain.

3.2 Example 2: CT distal femur alingment

We also applied the approach to a problem in orthopedic biomechanics – warping a 2D section of the human distal femur of a specific subject into alignment with the anatomy of a template knee. CT slice data were used to drive the deformation of a 128x128 computational grid generating the results summarized in Fig. (2). The

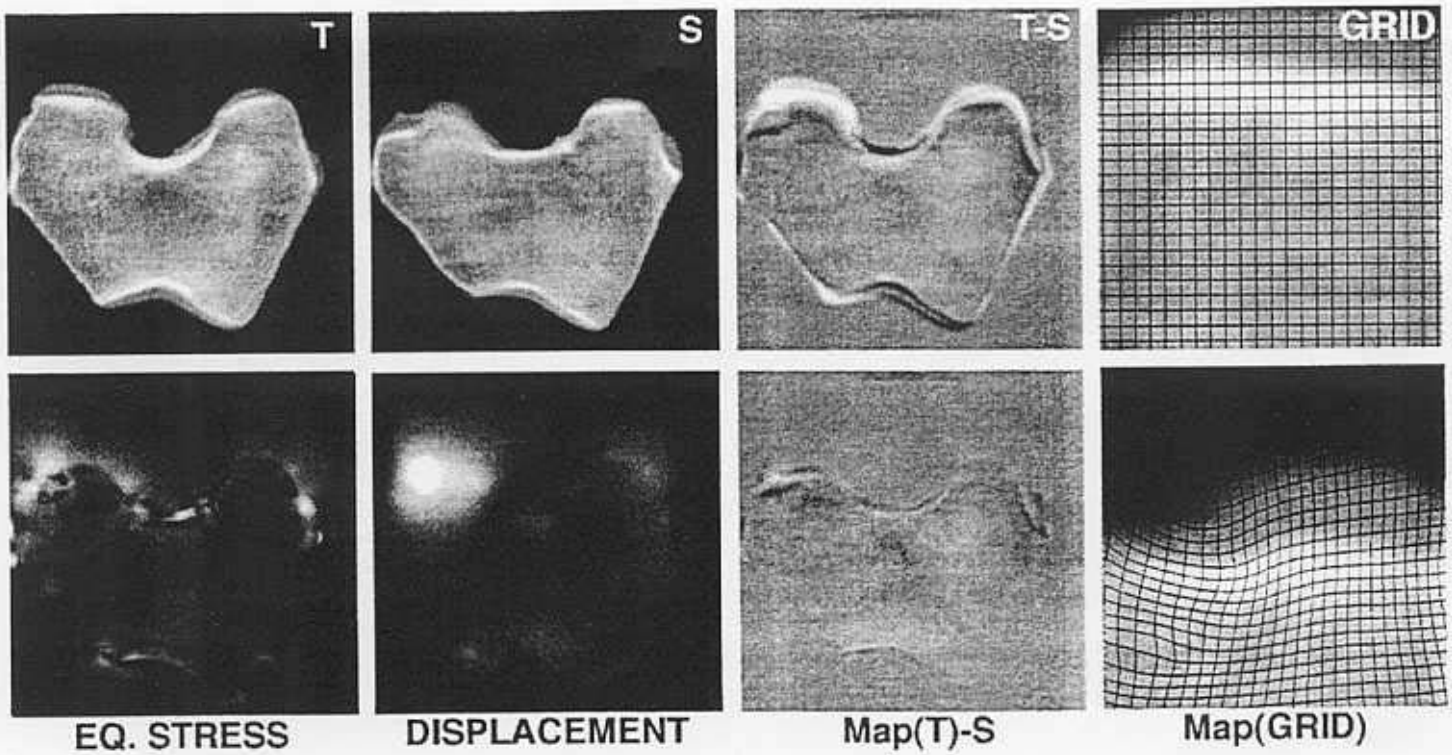

Figure 2: Hyperelastic mapping of the distal femur using CT image data. The top row from left to right shows the template image T, the specific subject S, the initial difference between the template and the subject T-S, and a close-up view of one section of the undeformed 128x128 FE grid. The bottom row shows some results of the image driven deformation with images from left to right denoting the equivalent stress, the magnitude of the in-plane displacement and a magnified section of the deformed FE grid.

images also have a resolution of 128X128. A CT image of the template (T) is shown in the upper-left panel and a CT of the individual is shown as the second panel (S). Prior to warping, significant differences are apparent in the location of the cortical shell, the size and shape of the condyles, the intercondylar notch, collateral ligaments and surrounding cartilage. These differences are reflected in the difference image T-S prior to the deformation. After warping the difference image Map(T)-S is nearly equal to the background level (i.e. zero) . Gross translation and rotation are accommodated in addition to fine-scale distortions. Results obtained here for the hyperelastic constitutive model compare favorably with results for the same test problem reported by Rabbitt *et al.*, computed using a massively parallel implementation of a viscous fluid constitutive model.[21]

3.3 Example 3: Bending of a beam using an object-based FE grid

To illustrate independence of the computational grid and the image discretization, the method was applied to bend a uniform hyperelastic beam into alignment with distorted shape. Results are summarized in Fig. (3). The deformation was driven by the difference between two 128x128 (16k pixels each), 8bit grayscale images but computed on the on the much smaller (2k element) FE mesh associated with the template beam. In both of the previous examples all of the nodal points around the periphery of the mesh were fixed. For the beam example simple springs were attached to the corners of the beam in order to insure a positive definite stiffness matrix at the initial time. The image Map(T)-S in Fig. (3) shows the influence of the springs in bending the corners of the beam. No attempt was made optimize the spring constraints. In this particular example the elements in the reference configuration fall on a rectangular grid which is natural given the initial shape of the beam, but this is not necessary. One of the advantages of FE over finite difference and related techniques is the ability adapt the computational mesh to the geometry of specific structures. This is particularly advantageous when using the present technique to track the deformation of real materials.

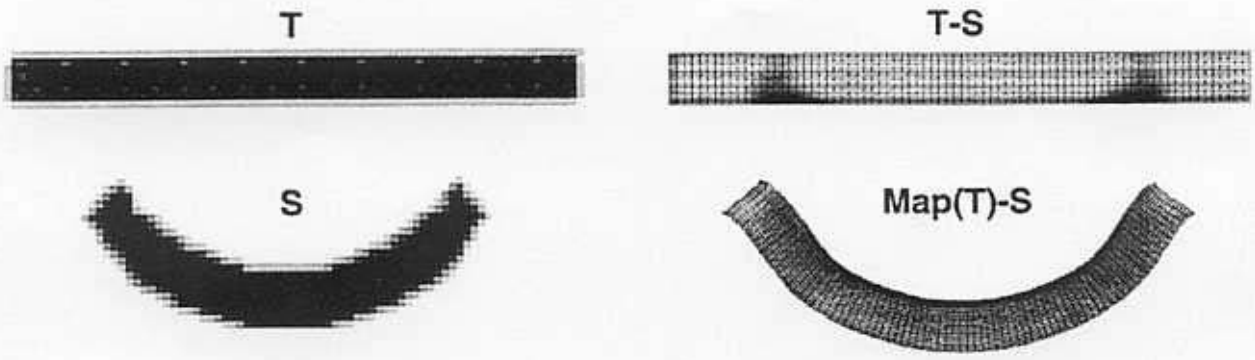

Figure 3: Bending of a beam. The top row shows the template image (T) and the undeformed FE mesh (T-S). The bottom row shows image data for the deformed configuration (S) and the resulting deformed shape of the beam (Map(T)-S). In the right column, the grayscale shading superimposed on the computational mesh indicates the difference between the deformed template and the specific case.

4 DISCUSSION

In the present work we develop a FE tool to combine image-based data with the mechanics of deforming continua. Mathematical results provide the ability to track the deformation of materials in the presence of unknown forces, using image data to provide the necessary additional information. This is achieved using standard techniques from nonlinear continuum mechanics in combination with probabilistic models of the imaging sensors. A detailed derivation is provided for the case of a hyperelastic material and a Gaussian image sensor, but many of the main results are equally valid for numerous other material types and imaging modalities. For example, the form of the governing equations (Eq's 2.23, 2.35, 2.35) apply to numerous other situations.

From a medical image-analysis viewpoint, the approach described here is most closely related to previous work by Christensen *et al.*applied to the intersubject registration of neuroanatomies, but also differs in several respects.[4] The most notable difference is the current mathematical derivation and associated FE implementation. Use of the Lagrangian reference frame is one of the most obvious departures from this previous work. New results are based on an implicit formulation requiring the tangent stiffness in Eq. (2.35) whereas the previous work employed explicit time integration. Since the new mathematical results have been incorporated into a well developed non-linear FE code (NIKE3D[22]) numerous capabilities provided by the code are now accessible to image driven deformations. For example, the new implementation includes the ability to address i) inhomogeneous and anisotropic properties, ii) a nonuniform computational mesh independent of the image discretization, and iii) a wide variety of physical material descriptions. This is achieved for large-magnitude nonlinear deformations. Since NIKE3D employs an implicit solution scheme, its application is most suitable for quasi-static and low to medium-rate in problems in solid mechanics.

In addition to its relationship to the field of deformable image templates, the current work also has roots

in the field of mechanics. This is most obvious for a class in *inverse problems* where *a priori* field equations are combined with sensor data to estimate physical parameters. Examples arise in atmospheric dynamics, ocean dynamics, astronomy, heat transfer, geophysics and other areas.[23–30] In some of these problems the field equations are treated as hard constraints with unknown boundary and/or initial conditions. In this approach it becomes necessary to solve the adjoint system which is typically an ill-posed problem. The present work describes a slightly different approach where the field equations are assumed to contain a spatially distributed unknown body force. If the mechanical model were exact the forward solution of the field equations would predict the configuration of the system and hence the body force associated with the image data would be zero. This is reflected in the resulting momentum equations (Eq's 2.8 and 2.12). For real systems the mechanical model is not exact and the body force arising from the image data serves to correct for the *a priori* assumptions.

Although the method is designed to address real deformations of physical materials many potential applications, including the examples presented here, are analogues to real materials and do not involve any actual physical deformations. Application of the technique to such problems results in a homeomorphic mapping consistent with the laws of continuum mechanics. Indeed, these mapping applications served as an original motivation for the work.[2,31,3] When the present technique is applied to imaginary deformations such as anatomical mapping special care must be taken in the design and/or selection of the template(s). For example, in order to accurately reproduce the specific anatomy the template must contain all of the same structures as the specific subject (i.e. a brain can not be warped into alignment with a knee). Future work in the anatomical mapping area may employ multiple templates in order to account for pathologies or congenital conditions where specific structures are highly irregular or completely missing.

Related to the choice of templates, some initial global alignment is necessary in order for the method to progress automatically. This is easily seen with respect to results for examples 1 and 3 where partial overlap in the initial images is present. If the two Gaussian intensity patches were completely separated then the body force associated with the difference between the two images would be zero, when computed on the continuum using Eq. (2.24), and the deformation would not proceed. Even in such cases when the initial images have no overlap, alignment can often be achieved using the current FE approach by starting with a coarse computational grid and refining the grid as the computation proceeds. In the current implementation this is facilitated through the assignment of an effective image intensity at each computational element equal to the average over the image domain spanned by the element. If there is some image overlap on the coarse mesh then "global" alignment will proceed to within the spatial resolution of the coarse mesh. Once global alignment is achieved a higher resolution mesh is introduced to complete the fine-scale registration.

Results for hyperelastic materials show that the strain energy approaches infinity as the Jacobian becomes singular. Hence, non-homeomorphic deformation fields are disallowed because they are associated with infinite penalty. This is also true for hyper-viscoelastic materials. In our implementation, material properties are associated with the geometric model (template) such that any inhomogeneity or anisotropy moves and rotates locally with the deforming material. Tracking of the nonlinear motion and inhomogeneous properties allows for the embedding of structures such as hyperelastic thick shells or volumes into a soft-elastic or visco-elastic base material. Embedding provides a means to constrain the deformation of sub-structures even in the presence of large global motions of the sub-structures and in the presence of large deformations of the surrounding base media.

Once the displacement field is determined that aligns the image data, the mapping can subsequently be applied to any other structures associated with the template. For biological applications this may involve mapping of physiological data from the specific case to the template configuration or vice versa. For example, the mapping obtained by the alingment of two neruoanatomies using MR data can subsequently be applied to positron emission tomography (PET) results to map the PET images from the individual onto the anatomy of a canonical template. Geometrical examples are shown in the figures where the mapping is applied to the computational grid itself. It is just as simple to apply the deformation field to surface or volumes identified in the template to map them to the individual case. This can also be done for computational meshes in order to map the model grid to the individual subject. Since the algorithm is based on nonlinear finite-deformation theory the transformation is one-to-one and avoids singularities encountered with simpler schemes.

In summary, a new FE implementation to combine image-based data with solid mechanics problems has been developed and presented. Future applications include both real deformations and analog homeomorphic mapping problems.

5 Acknowledgements

Partial support for this work was provided by the NIH NCRR-RR01380 and by joint funding from NIMH, NASA and NSF under the Human Brain Project grant RO1 MH/DA52158.

6 REFERENCES

[1] U. Grenander and M. I. Miller. Representations of knowledge in complex systems. *J. Royal Stitistical Soc. B*, 56(3), 1994.

[2] R. Bajcsy, R. Lieberson, and M. Reivich. A computerized system for the elastic matching of deformed radiographic images to idealized atlas images. *J. Computer Assisted Tomography*, 7(4):618–625, 1983.

[3] G. E. Christensen, R. D. Rabbitt, and M.I. Miller. 3D brain mapping using a deformable neuroanatomy. *Physics in Medicine and Biology*, 39:609–618, 1994.

[4] G. E. Christensen, R. D. Rabbitt, and M.I. Miller. Deformable templates using large deformation kinematics. *(submitted 10/94) IEEE Trans. on Image Processing*, 0, 1994.

[5] A. Knoerr. Global models of natural boundaries. *Pattern Analysis Report*, 148, 1994.

[6] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models – their training and application. *Computer vision and image understanding*, 61(1):38–59, 1995.

[7] J. A. Weiss, J. A. Painter, and E. P. France. Finite element mesh genration of the knee and its soft tissues from ct data. In *Proc. 2nd World Congress of Biomechanics*, page 149, Amsterdam, The Netherlands, July 2-9 1995. IBC.

[8] Y. Lanir. Constitutive equations for fibrous connective tissues. *J. Biomechancis*, 16(1):1–12, 1983.

[9] Y. C. Fung. *Biomechanics: Motion, Flow, Stress and Growth.* Springer-Verlag, 1990.

[10] A. J. M. Spencer. *Continuum Mechanics.* Longman Scientific and Technical, Essex, 1992.

[11] J. E. Marsden and T.J.R. Hughes. *Mathematical Foundations of Elasticity.* Dover, Minneola, New York, 1994.

[12] J.C. Simo. A framework for finite strain elastoplasticity based on maximum dissipation and the multiplicative decomposition: Parti. continuum formulation. *Computational Methods in Applied Mechanics and Engineering*, 66:199–219, 1988.

[13] D. L. Snyder, Jr. L. J. Thomas, and M. M. Ter-Pogossian. A mathematical model fo positron-emission tomography systems having time-of-flight measurements. *IEEE Trans. on Nuclear Sci.*, 28(3):3575–3583, 1981.

[14] J. Zhang, J. W. Modestino, and D. A. Langan. Maximum-liklihood parameter estimation for unsupervised stochasit model-based image segmentation. *IEEE Trans. on Image Processing*, 3(4):1057–71, 1994.

[15] E. R. McVeigh, R. M. Henkelman, and M. J. Bronskill. Noise and filtration in magnetic resonance imaging. *Mecical Physics*, 12(5):586–591, 1985.

[16] T. J. Schaewe and M. I. Miller. A maximum-liklihood reconstruction algorithm for magnetic resonance imaging. In R. H. Schneider, editor, *Medical Imaging IV: Image Formation*, pages 188–194. SPIE 1231, 1990.

[17] E. T. Jaynes. On the rationale of maximum-entropy methods. *Proceedings of the IEEE*, 70(9):939–952, 1982.

[18] K-J. Bathe. *Finite Element Procedures in Engineering Analysis.* Prentice-Hall, New Jersey, 1982.

[19] H. Matthies and G. Strang. The solution of nonlinear finite element equations. *Int J Numer Methods Eng*, 14:1613–1626, 1979.

[20] B.N. Maker, R.M. Ferencz, and Hallquist J.O. Nike3d: A nonlinear, implicit, three-dimensional finite element code for solid and structural mechanics. *Lawrence Livermore National Laboratory Technical Report*, UCRL-MA(105268), 1990.

[21] R. D. Rabbitt, J. A. Weiss, G. E. Christensen, and M. I. Miller. Mapping of hyperelastic deformable templates. In *Proceedings of the International Soc. for Optical Engineering*, San Diego, CA, July 9-14 1995. SPIE.

[22] B.N. Maker, R.M. Ferencz, and Hallquist J.O. Nike3d: A nonlinear, implicit, three-dimensional finite element code for solid and structural mechanics. *Lawrence Livermore National Laboratory Technical Report*, UCRL-MA(105268), 1990.

[23] K. F. Gauss. *Theory of the motion of heavenly bodies moving about the sun in conic sections.* reprinted by Dover Publications (1963), 1809.

[24] J. V. Beck and Jr. C. R. St. Clair. *Inverse Heat Conduction: Ill-posed Problems.* Wiley-Interscience, 1985.

[25] W. Menke. *Geophysical Data Analysis: Discrete Inverse Theory.* Academic Press, 1984.

[26] T. M. McGee. Seismic reverberations and the remote estimation of properties of underwater soils. *Int. J. of Imag. Sys. and Tech.*, 3:40–57, 1991.

[27] J. H. Justice, A. A. Vassiliou, S. Singh, J. D. Logel P. A. Hansen, B.R. Hall, P.R. Hutt, and J. J. Solanki. Tomographic imaging in hydrocarbon reservoirs. *Int. J. of Imag. Sys. and Tech.*, 1:62–72, 1989.

[28] J. Schröter. Driving of non-linear time-depened ocean models by obervation of transient tracers – a problem of constrained optimisation. In D. L. T. Anderson and J. Willebrand, editors, *Ocean Criculation Models: Combineing Data and Dynamics*, pages 257–285, Dordrech, 1989. Kluwer Academic Publishers.

[29] D. J. Webb. Assimilation of data into ocean models. In D. L. T. Anderson and J. Willebrand, editors, *Ocean Criculation Models: Combineing Data and Dynamics*, pages 233–256, Dordrech, 1989. Kluwer Academic Publishers.

[30] D. Olbers and M. Wenzel. Determining diffusivities from hydrographic data by inverse methods with applicatios to the circumpolar current. In D. L. T. Anderson and J. Willebrand, editors, *Ocean Criculation Models: Combineing Data and Dynamics*, pages 95–139, Dordrech, 1989. Kluwer Academic Publishers.

[31] R. Bajcsy and S. Kovacic. Multiresolution elastic matching. *Computer Vision, Graphics, and Image Processing*, 46:1–21, 1989.