QUARTERLY OF APPLIED MATHEMATICS

VOLUME LXI MARCH · 2003 NUMBER 1

MARCH 2003, PAGES 1–16

A MATHEMATICAL REPRESENTATION OF BIOLOGICAL VARIABILITY IN MEDICAL IMAGES

BY

LARISA MATEJIC

Division of Applied Mathematics, Brown University, Box F, Providence, RI 02912

Abstract. Medical image ensembles exhibit variability and it is the aim of computational anatomy to represent such variabilities mathematically and to exploit these knowledge representations by inference algorithms implemented through code. The variability is caused by several factors and our pattern theoretic approach rests on the assumption that they can be understood in terms of groups of transformations and probability measures on such groups. We shall arrange the similarity groups in a cascade, typically starting with the more rigid transformations and continuing with more flexible ones. Most importantly, however, we attach great significance to the physical and biological interpretation of the similarity groups.

I. Introduction. Medical imaging techniques, such as magnetic resonance (MR), x-ray computed tomography (CT), positron emission tomography (PET), and single-photon emission tomography (SPECT), have been studied for the past two decades and are increasingly growing in popularity as it is becoming more clear that such images can be used in many medical applications. Perhaps the most common uses are those of automatic segmentation of images into subimages and restoration of noisy images. Given a fully segmented and labeled anatomical image, a typical template, an algorithm could produce a correspondence, or a mapping, from that image onto the image that is being segmented and thereby transfer the complete segmentation along with all the labeling and knowledge. This approach reduces the time it takes to obtain a segmentation and may well improve accuracy of the end result.

Still other applications can be found in monitoring and diagnosing diseased anatomies. The kinds of algorithms that produce mappings from one image to another can be utilized to statistically formulate shape variability. Information concerning shape variability in anatomies may, among other things, turn out to be extremely useful in diagnosing diseases

Received May 18, 1998.

2000 *Mathematics Subject Classification.* Primary 60G35, 60D05.

Key words and phrases. Image representations, imaging/mapping, group theory, stochastic differential equations.

Supported by ARO grant DAAL03-92-G-0115.

E-mail address: lma@dam.brown.edu

such as schizophrenia or manic depression. Various shape databases can be built to store information that aids in classifying a particular patient as normal, diseased, or at risk. Furthermore, progress of a disease in a patient can be tracked using similar techniques. If a disease involves deviation in anatomical shape or texture, such a change could be monitored both qualitatively and quantitatively as the patient's image from a previous examination is mapped onto the current image.

II. Literature review. The image representation issue was pioneered in the 1970's in Besag [1], where the use of Markov random fields for analysis of spatial statistics was proposed. When analyzing human anatomies, the method of Markov random fields (MRF) proves insufficient in that it does not address the global complexity inherent in anatomies as well. Out of the need for representing complex human anatomies arose a representation from pattern theory referred to as *global shape models* [2] which is used to characterize structural similarities in terms of templates and variability in terms of probabilistic transformations applied to such templates. The *template* is thought of as a typical representative of a set of images under study. In literature, a template is also referred to as the *textbook* or an *idealized atlas*. It is assumed that features of particular images (*personalized or individual atlases*) can be obtained through some transformation applied to the template. A prior distribution is defined on the space of transformations, not the image space, and the task of the algorithm is to select the appropriate transformation from the posterior distribution. Such a transformation or mapping provides a structural understanding of the image.

Much activity can be observed in constructing such image mapping algorithms. Biological variability of shape was studied first using quantitative morphometric analysis by Bookstein [3], [4]. Bookstein used a method of thin-plate splines to produce a mapping which aligns the landmarks in the template and the data. Terzopoulos' algorithms [5] are examples of *physically based modeling*, a method which employs physically constrained transformations. Bajcsy [6] and Christensen [7] have applied elasticity methods to study brain anatomy. Atlas-based anatomical localization is based on the assumption that anatomy of normal individuals at some level can be thought of as quantitative variations on a common underlying qualitative plan. A given individual's anatomy is a special case and can be derived by a physically based model of deformation as an elastic matching. The reference anatomy or the template is modeled as an elastic object [8]. The research that uses elasticity methods to model the brain was primarily done to utilize the available mathematical tools. Even though the brain possesses elastic properties in a limited sense, it is not clear, in a biological sense, why elasticity as a property is singled out to represent such a wide range of variability of the brain. It seems that a better approach would attempt to understand biological variability and use that knowledge to construct appropriate mathematical methods. Perhaps this is too ambitious a task, but it is our belief that efforts should be put to move the research in that direction.

The method presented here employs group cascades for representing anatomical variability in an attempt to allow biology to drive the mathematics. The goal is to map the image template to a deformed image by computing a diffeomorphism with tools from differential geometry (such as Lie groups and Lie algebras) and probability theory. This

approach is used with the belief that finding a representation based on properties of the subject that is being modeled would prove more versatile and more accurate in the long run.

The primary objective of this undertaking is to mathematically represent biological variability in medical images. The mathematics needed for a more empirical, more biological approach to imaging has been developed to aid in this endeavor. This theory does not depend on particular imaging techniques. Rather, the methodology encompasses many existing modalities, such as MR, CT, PET, and SPECT. A general method for analyzing images by taking biological variation into account is provided. In this manuscript nothing is offered in terms of expertise of sensor physics. We are almost exclusively interested in modeling the biology inherent in the images. More information on sensor physics can be obtained from [9].

III. Variability and invariance in medical images. To represent medical images effectively, their common features must be observed and the differences they exhibit must be understood. The mathematical framework to be introduced embodies the similarity and variability observed in the images. Similarity is represented in terms of mathematical groups S_κ of image transformations, and variability is modeled by defining probability measures on such groups. The variability that the images exhibit can be further classified as belonging to either the set of variabilities introduced by the imaging sensors or the set of biological differences.

A. Variability due to Imaging Sensors. A source of variability stems from the setup of the sensor and observation noise inherent in its physical functioning. For many sensors the properties of the noise are fairly well understood. For some it may be described by a Gaussian measure, whereas for others with quantum-limited functioning, a Poisson measure is a more adequate description.

In addition to the random noise, systematic differences could be due to differences in radiation level. The same sensor type used on two different occasions on the same patient can result in different outputs, so that, in order to compensate for this, some calibration is required. This will be done in terms of a calibration group of transformations. Applying the calibration group to two images normalizes pixel intensities, so that they can be compared in further calculations.

Yet another source of variability arises from the positioning of the sensor relative to the patient. In the framework of this research this will be dealt with through a registration group of rigid transformations. More specifically, to compensate for the different positions of the sensors during scanning of the patients, the images of the two subjects will be translated and rotated.

B. Biological Variability. The real challenge, however, appears in trying to understand *biological variability.* In an extreme reductionist approach this would be explained in terms of molecular biology, genetics and developmental factors. This may be possible sometime in the future, but one has to be more pragmatic and aim for less reductionist, but hopefully useful, descriptions with empirical backing.

Medical image ensembles exhibit variability and it is the aim of *computational anatomy* to represent such variabilities mathematically and to exploit these knowledge representations by inference algorithms implemented through code. The pattern theoretic approach rests on the assumption that the biological variability in shape and texture can be understood in terms of groups S_κ of transformations and probability measures on such groups. Groups of interest include those with physical and biological meaning, including calibration (strength of the magnetic field), translation and rotation (subject positioning in the sensor), scale (subject size), landmark (points of biological significance that drive large scale variation), and local groups (small scale variations). Many such groups will be introduced and organized into a *group cascade* that plays a key role in building inference engines.

C. Topological invariances. Although anatomies exhibit large variability, they also possess characteristic invariances, if not in the details, at least on a topological level. The mathematical formulation of invariances is best expressed through groups of transformations, which is why the similarity groups S_κ referred to earlier have been introduced. For example, the registration groups represent invariances due to the arbitrariness of the choice of coordinate system. Another example, at the other extreme, of a flexible rather than a rigid transformation, is the local group, a group of diffeomorphic mappings $X \leftrightarrow X$. However, it should be clear that whatever the similarity group, the choice between its elements should not be considered as arbitrary; instead, some should be thought of as more likely than others. For example, when dealing with a registration group, the positioning of the sensor relative to the body of the patient is known to some extent beforehand and this prior knowledge should be represented by a probabilistic statement, a *prior probability measure* P_κ on S_κ. The values of the subscript κ can be, among others, "calibration", "registration", "scale", and "local".

REMARK. It has been tacitly assumed that the anatomies analyzed in this manuscript are normal. Some pathologies can be represented by similarity groups but with P_κ's that differ from what is used in the case of normal anatomy. Other more severe pathologies cannot be described by groups of diffeomorphic mappings; they lead to more radical transformations.

IV. The concept of group cascades. The similarity groups are arranged in a *cascade*, typically starting with the more rigid transformations and continuing with more flexible ones. This idea was introduced in [10] in the special case of the anatomy of human hands, and it will now be studied more systematically and applied in a more general setting.

This approach could also be looked at as an application of the traditional coarse-to-fine strategy, familiar from the numerical solution of partial differential equations and simulation of Markov random fields [11], [12], [13], but that is only one aspect of the cascade strategy. The coarse-to-fine approach is motivated by the search for fast algorithms. In the present context speed is certainly important, but greater significance is attached to the *physical and biological interpretation* of the similarity groups. For example, translation and rotation groups obviously imply a statement about the positioning of the sensor, the scale group says something about the size of anatomies under development, and so

on. At present, sufficient empirical evidence to accurately estimate the (S_κ, P_κ) from data is not available, but that is the long range goal.

A cascade strategy is applied with a cascade of groups that combine in an *inference engine* which propagates the original image iteratively through groups for translation, scaling, landmark, local, and others. The engine loops through a probabilistic algorithm over each similarity group with the groups ordered in the cascade. Probabilistic tools include the prior, which represents general knowledge about anatomical variability, likelihood, which carries information about the randomness of sensor variability, and posterior, which combines the information from both the prior and the likelihood in the Bayesian sense.

V. Mathematical formulation of the mapping problem. This section presents a *method which employs group cascades* to map the image template to a deformed image by computing a diffeomorphism with tools of group theory and probability. Details are given concerning synthesizing group cascades, both in a general sense and for a particular case. The presentation of the material is further subdivided into a section on inducing the group prior, *pattern synthesis*, and a section on computing the posterior, *pattern analysis*.

The output of an imaging sensor is denoted by

$$I^{\mathcal{D}} : X \to Y$$

where the *background space* X is a rectangular lattice with the lattice points

$$x = (x_1, x_2); \quad 1 \leq x_i \leq l_i, \ i = 1, 2$$

and the *contrast space* Y is the real line or a subset of it. In the continuum approach we shall consider a lattice embedded in a 2D Euclidean *extended space* $X \subset \Xi = \mathbf{R}^2$.

In our analysis of medical images we operate on the premise that some physical truth exists. By that is meant that the image seen through imaging sensors is a noisy variation of the true image. Since the true image is always seen through some imaging sensor, unless it is artificially synthesized, we can never observe it directly. Instead, the observed image $I^{\mathcal{D}}$ is what we use in the algorithm. The observed image is thought of as a *deformed image* because it is the result of a deformation mechanism $\mathcal{D}$ acting upon a typical representative, the *template* I_{temp}, so that

$$I^{\mathcal{D}} = \mathcal{D}sI_{temp},$$

where $\mathcal{D}$ represents the physics of the sensors and the similarity group transformation $s \in S$ represents the biology of the subject, the formulation of which is the real challenge of this undertaking [2]. We speak of *the* template, but we really mean *a* template, since we could have several different templates. For instance, a different template could be associated with each stratum in the population.

A. Cascade Strategy. Once the similarity groups have been chosen, the *cascade strategy* has to be decided, namely the order in which the groups are executed with possible repetitions.

A flow diagram represents the cascade strategy as a series of modules. One such flow diagram is shown in Fig. 1. Fig. 1 shows how the image template is propagated

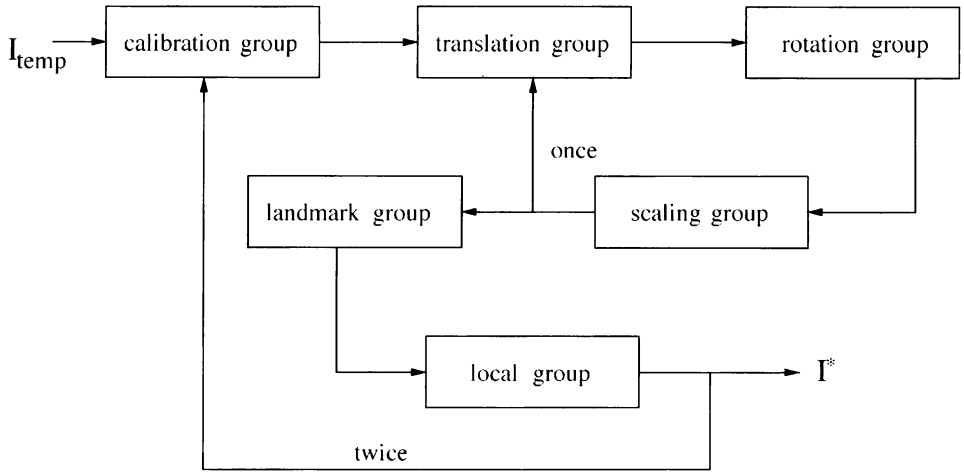

FIG. 1. An example of cascade strategy

through a series of modules, some with repetition. Rigid transformations, translation and rotation, along with non-rigid scale, are performed first to align the global features. The details are transformed using higher dimensional spatial groups such as landmark and local. After executing each group (S_κ, P_κ), the template is updated. In other words, the template does not remain fixed: it is *propagated* through the group cascade to the value of the deformed template after the execution of each group. The choice and the order of groups in the rigid category, as well as the choice and the order of groups in the higher dimensional spatial category, is determined by images that are input into the algorithm. A performance tracking mechanism is incorporated into the algorithm to determine when the template should move from one module to another, as well as to determine when the algorithm should terminate. In addition to the performance tracking mechanism, a look-ahead mechanism is present to decide which modules the image template should enter at a given instance. Upon conclusion of the algorithm, what we call an *inferred image*, I^*, is obtained. The image output by the inference engine I^* should resemble the deformed image I^D.

B. Pattern Synthesis: Inducing the Group Prior. Once the groups to be incorporated into the cascade are established, a probability measure needs to be defined on the similarity groups to model normal biological variation observed in the images. Such a measure is parameterized by a parameter t that is a vector in the Euclidean space $\mathbf{R}^m$. The t-vector represents a local coordinate vector in the Lie group under consideration. The elements of each group (in a 2D domain) in the cascade are transformations characterized by

$$x \to f(x,t), \quad x \in \mathbf{R}^2, t \in \mathbf{R}^m.$$

Here x represents a gridpoint and $f(x,t)$ is a function that determines the amount a particular gridpoint x is displaced given the value of the parameter t. To obtain an

estimate for t, the standard S.D.E. (Langevin's equation) is used:

$$\frac{dt}{d\tau} = -\nabla E(t)\, d\tau + \sqrt{2}\, dN(\tau). \tag{V.1}$$

The reason Langevin's equation is used is that it provides a general mathematical framework for this problem to incorporate any type of prior. For instance, if in our study of anatomy we make observations that point to a different prior, that fact is easily substituted into the equation to make subsequent simulations more accurate. In [14] it is discussed how under certain conditions, if τ is big enough, $t(\tau)$ has the density

$$p(t) = \frac{1}{Z} \exp[-E(t)] \propto \exp[-E(t)],$$

where Z is the normalizing constant or partition function and E is the energy. The energy term specifies the amount of deformation to be applied to the image template.

The next step is to discretize the S.D.E. in (V.1) to obtain

$$t_{new} = t_{old} - h\nabla E(t_{old}) + \sqrt{2h}\, dN. \tag{V.2}$$

Here h stands for the time step (algorithmic time) and N is a 2D random variable with $N(0,1)$ i.i.d. components. Also, $t \in \mathbf{R}^m$, $t = (t_1, t_2, \ldots, t_m)$; t_i's are i.i.d. $N(0, \sigma^2_{group})$. The parameter σ^2_{group} depicts the variance associated with a particular group transformation. The variance places a likelihood on the amount of deformation to be performed by the group on the image template. The value of this parameter is different for each group and is obtained empirically from studying shapes of anatomies and their positioning with respect to different imaging sensors. For instance, we can establish that in taking the MR scan, most brain images are within a certain number of pixels to the right or left of each other. That would lead us to establish a standard deviation, and consequently, a variance on the translation group in use with MR brain images. It should be pointed out that direct simulation rather than an iterative method using S.D.E.'s can be used for pattern synthesis where the prior is defined on the groups. However, for pattern analysis, which involves the posterior, an iterative method is necessary, and thus pattern synthesis was approached in a similar fashion to prepare for what is to come when pattern analysis is performed.

C. Pattern Analysis: Computing the Posterior. So far, only one part of the posterior has been addressed, namely the prior. The likelihood component will now be examined in greater detail and ultimately used to determine the posterior. The prior represents general beforehand knowledge we have about the anatomy under study. The likelihood provides information from observing a particular instance of that anatomy. The posterior synthesizes the two types of information. The exact relationship between the posterior, prior, and likelihood can be summarized with equations that follow. The density is

$$p_{post}(t) \propto \exp[-E_{post}(t)], \quad t \in \mathbf{R}^m,$$

which leads to an expression for the energy gradient

$$\nabla E_{post}(t) = \nabla E_{prior}(t) + \nabla E_{likeli}(t).$$

We now profit from deriving the S.D.E.'s for synthesis in section V-B. Even though the prior did not have to be synthesized in such an indirect way, it makes the work in this

section easier to handle. This is due to the nature of the likelihood. The likelihood has to be computed each time the estimate for t is refined to guide the estimation procedure. Since the prior incorporates *a priori* knowledge, it is needed to make the estimate of t more accurate. Thus, if the nature of the algorithm as a whole is iterative, the prior has to be coded that way, too. Since the expression for $\nabla E_{prior}(t)$ is already available from section V-B, only estimates for $\nabla E_{likeli}(t)$ need to be obtained. In the rest of this manuscript, the image template is referred to as I_{temp}. The symbol $I^{\mathcal{D}}$ describes the deformed image. For notational purposes, $I(x)$ is defined as $I(x) = I_{temp}(f(x,t))$, where f is a function of x and $t (x \in \mathbf{R}^2, t \in \mathbf{R}^m)$ that is module dependent and will be elaborated on in the subsections that follow. Then,

$$E_{likeli}(t) = \sum_x [I(x) - I^{\mathcal{D}}(x)]^2,$$

and the gradient,

$$\nabla E_{likeli}(t) = \left(\frac{\partial E_{likeli}(t)}{\partial t_1}, \frac{\partial E_{likeli}(t)}{\partial t_2}, \ldots, \frac{\partial E_{likeli}(t)}{\partial t_m} \right) \equiv (g_1, g_2, \ldots, g_m).$$

Let us derive an expression for the g_i's with $i = 1, \ldots, m$.

$$\begin{aligned} \frac{\partial E_{likeli}(t)}{\partial t_i} &= \frac{\partial t}{\partial t_i} \sum_x [I_{temp}(f(x,t)) - I^{\mathcal{D}}(x)]^2 \\ &= 2 \sum_x [I_{temp}(f(x,t)) - I^{\mathcal{D}}(x)] \cdot \frac{\partial}{\partial t_i} I_{temp}(f(x,t)), \end{aligned}$$

where

$$\frac{\partial}{\partial t_i} I_{temp}(f(x,t)) = \frac{\partial I_{temp}(f(x,t))}{\partial x_1} \frac{\partial x_1}{\partial t_i} + \frac{\partial I_{temp}(f(x,t))}{\partial x_2} \frac{\partial x_2}{\partial t_i}.$$

Expressions for the quantities $\frac{\partial x_i}{\partial t_i}$ and ∇I_{temp} are needed now. If f is thought of as being a vector function comprised of two components $f(x,t) = (f_1(x,t), f_2(x,t))$, then

$$\frac{\partial x_1}{\partial t_1} = \frac{\partial f_1(x,t)}{\partial t_i}$$

and

$$\frac{\partial x_2}{\partial t_i} = \frac{\partial f_2(x,t)}{\partial t_i}.$$

The gradient of I_{temp} is approximated with differences

$$\begin{aligned} \frac{\partial I_{temp}(f(x,t))}{\partial x_1} &\approx \frac{1}{2}[I_{temp}(f(x_1+1, x_2), t) - I_{temp}(f(x_1-1, x_2), t)], \\ \frac{\partial I_{temp}(f(x,t))}{\partial x_2} &\approx \frac{1}{2}[I_{temp}(f(x_1, x_2+1), t) - I_{temp}(f(x_1, x_2-1), t)]. \end{aligned}$$

Then, to obtain an estimate for t, now with the addition of the likelihood component, the standard S.D.E. is used in discretized form

$$t_i^{n+1} = t_i^n - h \frac{\partial E(t)}{\partial t_i} + \sqrt{2h} N_i$$

for $1 \leq i \leq m$. Here h stands for the time step (algorithmic time) and N_i is a random variable with $N(0,1)$ i.i.d. components.

In the next section, each high-dimensional group is discussed by providing expressions for the gradient of E_{likeli} followed by an expression for $\frac{\partial}{\partial t_1} I_{temp}(f(x,t))$. The expressions for $\frac{\partial}{\partial t_i} I_{temp}(f(x,t))$ with $i > 1$ are derived using the same technique, so they will not be provided in the text.

VI. An example of a group cascade. A particular group cascade that is relevant in the digital anatomy undertaking will be considered in this section, and each group in the cascade will be characterized in 2D.

Typically, the cascade starts off with global transformation groups, such as translation, rotation, scaling, shearing, and bending. These have been studied in depth in [15]. However, for the sake of brevity in this exposition, we choose to concentrate on higher dimensional spatial groups such as the landmark group and the local group that come into play after the image contents have been aligned by the global transformation groups.

Some attention must be paid to the following technicalities. When a similarity transformation s is applied to a template, the parameter t will usually be a non-integer and since we are operating on a discretized integer grid, some interpolation will be necessary to obtain appropriate values at the gridpoints. The algorithm described here uses bilinear interpolation in 2D to resolve this issue. Another possibility would involve taking the median of four neighbors in 2D to obtain the appropriate value at a given gridpoint. Also, when a similarity transformation s is applied to a template contained in the background space X, it can and will usually happen that sX is partially outside of X. Since the template is only contained inside, some sort of extrapolation is needed (see [15]).

A. Specification of Groups.

A.1. Landmark Group. The purpose of the landmark group is to drive alignment of the two images by aligning first points of biological significance called *landmarks* ([3], [16]), followed by aligning of the rest of the image, where the appropriate displacement field is calculated from the landmark trajectories. Using this method a low number of parameters relative to the image size is needed to ensure that the resulting mapping will better account for biological features because the important sections of the image are guaranteed to be aligned. Mathematically, this means that for each image a set of points, Υ of size M, that play a significant role in the anatomy is enumerated. These points can range in number. The purpose of selecting the landmarks is to mark points of biological interest in the image to which impulses are applied which cause the rest of the points in the image to be displaced along with the landmarks.

The landmark group S_{land} is of dimension $2 \cdot M$, or its number of parameters is twice as many as the number of landmarks. The landmarks are represented by a $2 \times M$ matrix

$$x = \begin{pmatrix} x_1^1 & x_2^1 \\ x_1^2 & x_2^2 \\ \cdots & \cdots \\ x_1^M & x_2^M \end{pmatrix}.$$

Consequently, the t matrix is of the same size and contains the elements

$$t = \begin{pmatrix} t_1^1 & t_2^1 \\ t_1^2 & t_2^2 \\ \cdots & \cdots \\ t_1^M & t_2^M \end{pmatrix},$$

where t's come from the standard S.D.E. that defines the prior on the landmark group.

$$t^{n+1} = \left(1 - \frac{h}{\sigma_{land}^2}\right) t^n + \sqrt{2h} N.$$

The symbol h stands for the time step (for algorithmic time) and N is a $2 \times M$ random matrix with $N(0,1)$ i.i.d. components. Note that in this specification the same standard deviation of the amount of deformation, σ_{land}, is used for all landmarks. It is quite possible that due to the differences in density of various anatomical components, landmarks belonging to different anatomical components will have different variances, so that different values for σ_{land}^2 will have to be used for each component of t. Unfortunately, at present no analytical method exists for determining the value of such parameters. We have generally employed the approach of educated guessing supported by limited empirical evidence. It is our belief that substantial empirical evidence will determine values of these parameters for various anatomical regions.

The entire image will be displaced by a vector field, which is derived from applying impulses to the landmark points. Each impulse has a sphere of influence, where the points close to the landmark to which the impulse is applied tend to be displaced more than the points further away from the landmark. The displacement matrices, referred to as s_1 and s_2, are generated from the landmark points, the t's. The matrix s_1 represents displacement in the horizontal direction at each gridpoint of the image. Similarly, the matrix s_2 represents displacement in the vertical direction at each gridpoint of the image. Thus, each of these two matrices has to be the same size as the image. To transform the image, the corresponding point is displaced horizontally by the amount stored in s_1 and vertically by the amount stored in s_2.

Let us now examine more closely the procedure by which matrices s_1 and s_2 are generated from the t's. They are related by the equation

$$Ls_k(x) = \sum_{m \in \Upsilon} t_k^m \delta(x, x^m), \quad k = 1, 2, \tag{VI.1}$$

where L is defined to be a modified Laplacian $L = (-\Delta + a)$; $a \in \mathbf{R}$, and $\delta(\cdot, x^m)$ is the Dirac delta function centered at x^m. The reason a modified Laplacian is used in the definition of L is to avoid the singularity of the Laplacian operator at zero. The Dirac delta function is used to place impulses at the location of landmarks. Equation VI.1 simplifies further to

$$s_k(x) = \sum_{m \in \Upsilon} t_k^m G(x, x^m), \quad k = 1, 2, \tag{VI.2}$$

where $G(x, x^m)$ is the Green's function evaluated at (x, x^m). For an introduction to Green's functions, see [17]. By definition,

$$G(x, x^m) = L^{-1} \delta(x - x^m).$$

In discrete space, the Green's function for each landmark is obtained by solving

$$Lu = (-\Delta + a)u = e,$$

where e is defined to be a Gaussian field centered at each landmark. In other words, $G(x, x^m)$ is obtained by solving, for example by relaxation,

$$(-\Delta + a)u = e,$$

where

$$e(i,j) = \begin{cases} 1 & \text{if } (i,j) = (x_1^m, x_2^m); \quad i < l_1,\ j < l_2, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, the landmark group S_{land} has the elements $x \to x + s(x,t)$, where x is an image of size $l_1 \cdot l_2$ and $s(x,t) = (s_1(x,t), s_2(x,t))$, where $s_1(x,t)$ and $s_2(x,t)$ are the displacement matrices obtained by the method outlined above.

As far as the posterior is concerned, there are $2 \cdot |\Upsilon|$ gradients of the energy function or g's, introduced in section V-C, in the landmark group. The g's can be visualized as a matrix of size $2 \times |\Upsilon|$ of the following form:

$$g = \begin{pmatrix} g_1^1 & g_2^1 \\ g_1^2 & g_2^2 \\ \cdots & \cdots \\ g_1^{|\Upsilon|} & g_2^{|\Upsilon|} \end{pmatrix},$$

where

$$g_1^k = 2 \sum_x [I_{temp}(x + s(x,t)) - I^{\mathcal{D}}(x)] \left(\frac{\partial I_{temp}}{\partial t_1^k} \right) \Bigg|_{x+s(x,t)}$$

and

$$g_2^k = 2 \sum_x [I_{temp}(x + s(x,t)) - I^{\mathcal{D}}(x)] \left(\frac{\partial I_{temp}}{\partial t_2^k} \right) \Bigg|_{x+s(x,t)}.$$

Furthermore,

$$\begin{aligned} \left(\frac{\partial I_{temp}}{\partial t_1^k} \right) \Bigg|_{x+s(x,t)} &= \left(\frac{\partial I}{\partial x_1} \right) \Bigg|_{f(x,t)} \cdot \frac{\partial f_1(x,t)}{\partial t_1^k} \\ &= \frac{1}{2} [I_{temp}(x_1 + s_1 + 1, x_2 + s_2) - I_{temp}(x_1 + s_1 - 1, x_2 + s_2)] \\ &\quad \cdot G(x + s(x,t), x^k). \end{aligned}$$

A.2. Local Group. The local group models smaller scale biological variations in the images. After the lower dimensional groups have done most of the work in finding a mapping between the two images, the local group is applied for fine tuning. One can think of the local group as a generalization of the landmark group, where every gridpoint is a landmark.

The local group S_{local} has the elements $x \to x + t$, where x is an image of size $l_1 \cdot l_2$ and $t = (t_1, t_2)$, where t_1 and t_2 are both matrices of size $l_1 \cdot l_2$. The matrix t_1 represents the displacement in the horizontal direction of each point x. Similarly, the matrix t_2 represents the displacement in the vertical direction of each point of the image.

If the operator L is defined to be the modified Laplacian again, namely

$$L = (-\Delta + a), \tag{VI.3}$$

the local group prior is proportional to

$$\exp - \left[\frac{1}{2\sigma_{local}^2} (\langle Lt_1, Lt_1 \rangle + \langle Lt_2, Lt_2 \rangle) \right],$$

which means that the energy, E_{prior}, can be expressed as

$$\frac{1}{2\sigma_{local}^2} (\langle Lt_1, Lt_1 \rangle + \langle Lt_2, Lt_2 \rangle),$$

and the partial derivative with respect to the component t_1 is

$$\frac{\partial E_{prior}}{\partial t_1(x)} = \frac{1}{\sigma_{local}^2} L^2 t_1(x).$$

Similarly,

$$\frac{\partial E_{prior}}{\partial t_2(x)} = \frac{1}{\sigma_{local}^2} L^2 t_2(x).$$

This completely specifies the S.D.E. for synthesis of the local group measure.

The gradients needed to specify the posterior of the local group have the following form:

$$g_1 = 2 \sum_x [I_{temp}(x + t(x)) - I^{\mathcal{D}}(x)] \left(\frac{\partial I_{temp}}{\partial t_1} \right) \Big|_{x+t(x)}$$

and

$$g_2 = 2 \sum_x [I_{temp}(x + t(x)) - I^{\mathcal{D}}(x)] \left(\frac{\partial I_{temp}}{\partial t_2} \right) \Big|_{x+t(x)}.$$

Furthermore,

$$\begin{aligned} \left(\frac{\partial I_{temp}}{\partial t_1} \right) \Big|_{x+t(x)} &= \left(\frac{\partial I}{\partial x_1} \right) \Big|_{f(x,t)} \cdot \frac{\partial f(x,t)}{\partial t_1} \\ &= \frac{1}{2} [I_{temp}(x_1 + t_1 + 1, x_2 + t_2) - I_{temp}(x_1 + t_1 - 1, x_2 + t_2)]. \end{aligned}$$

B. Results

B.1. Synthesized Images. Pictorial results are displayed in this section to illustrate concepts introduced regarding group cascades. The image template in each case is a synthetically generated image. The deformed image is obtained by applying an arbitrary sequence of modules with random parameters to the template and introducing random noise into the image. The exact sequence used to obtain the deformed image is specified in the figure. The inferred image is what results out of the cascade engine when the template is propagated through it with the goal of finding a mapping between the template and a given deformed image. The closer the inferred image is to the target, which is the deformed image, the better the algorithm has performed. The true image is the unperturbed version, and it is provided to shed more light on the performance of the algorithm. In addition to introducing more noise, the values of the l_2-norms are provided in these two examples. By l_2-norm is meant the sum of the square differences in values of the corresponding pixels in two images. Before the algorithm begins, the

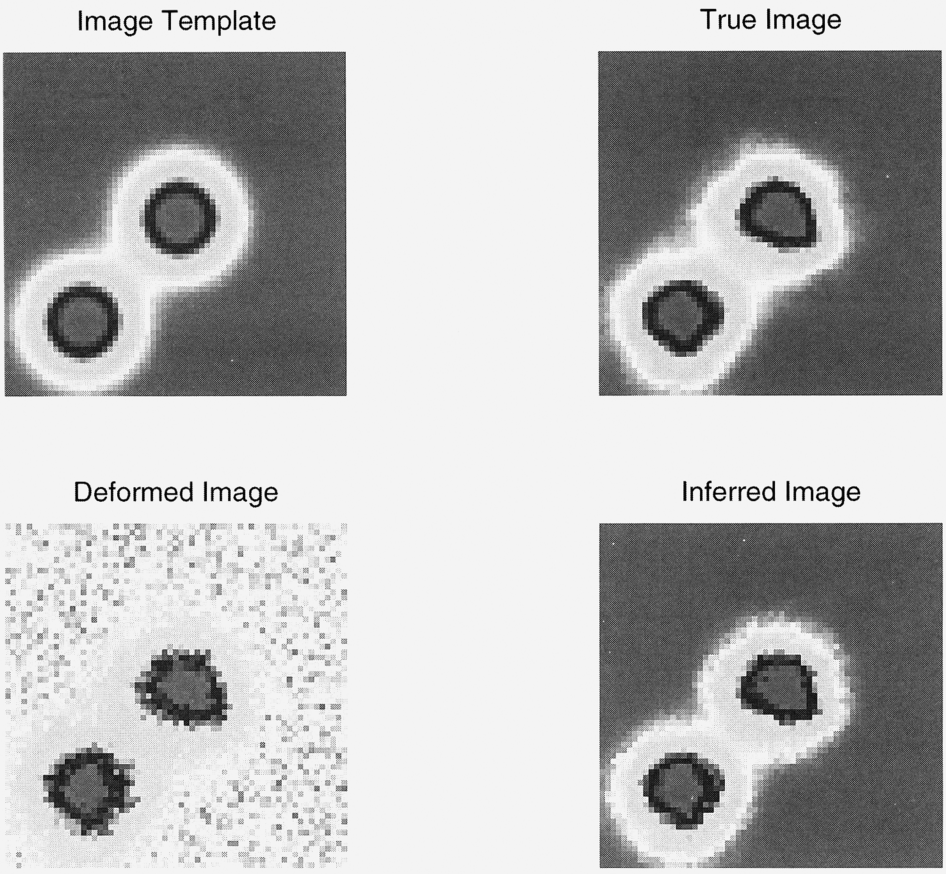

FIG. 2. A sample group cascade. The deformed image was obtained by applying the following sequence of modules to the image template: scale, translate, bend, landmark, local. The initial value of the norm $\sum_x [I_{temp}(x) - I^D(x)]^2$ was 53.637, and the final value of the norm $\sum_x [I^*(x) - I^D(x)]^2$ was 3.628

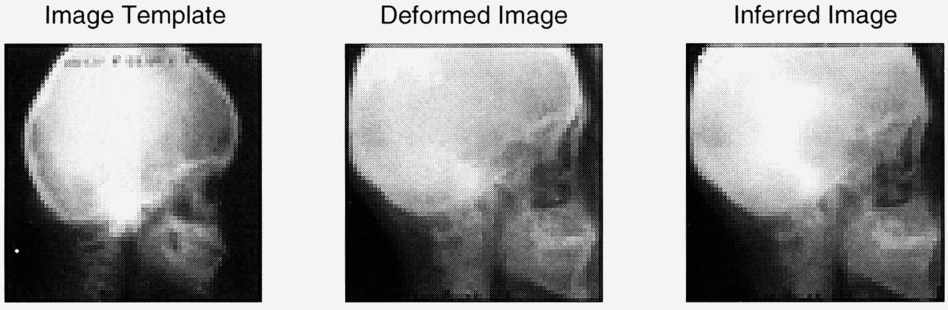

FIG. 3. Human skull: The template represents a child's skull, and the deformed or target image represents an adolescent skull

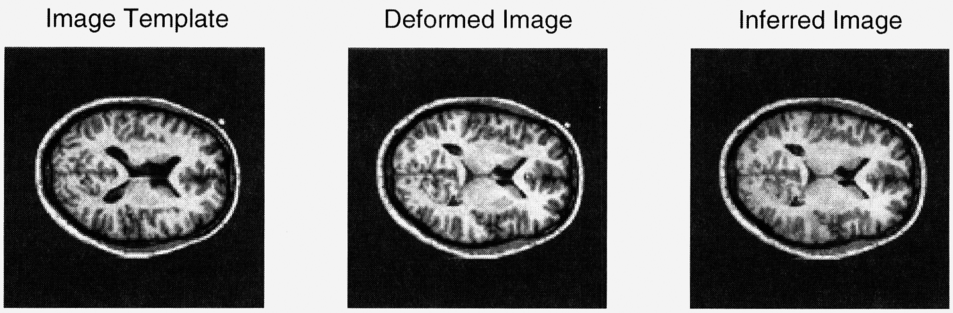

FIG. 4. Brain slice: The two slices are that of two different patients at around the same height; image courtesy of Sarang Joshi, Washington University

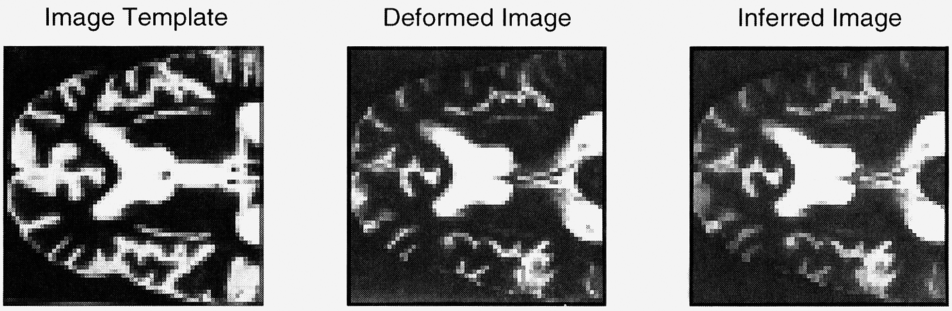

FIG. 5. Another brain slice

norm, or $\sum_x [I_{temp}(x) - I^D(x)]^2$, is computed, and upon conclusion of the algorithm, $\sum_x [I^*(x) - I^D(x)]^2$ is evaluated.

B.2. Medical Images. Analytical reasoning presented in this manuscript is now illustrated with medical images that were used to test the computer implementation of the algorithm. The first panel in a row of three panels that comprise each figure shows the image template I_{temp} that initializes the algorithm. The second panel shows the deformed image $I^{\mathcal{D}}$ or a patient target image, to which the image template is mapped. Finally, the third panel shows I^*, the inferred image, or what the template I_{temp} looks like upon the conclusion of the algorithm. The closer the template resembles the deformed image, the better the algorithm has performed. Sometimes, due to a wide range of intensities present in medical images, slight discrepancies appear in the mapping of intensity values. The algorithm is designed to handle mostly mapping of geometrical shapes, not intensities. A better intensity handling is something to be developed in future work.

VII. Concluding remarks. Using group theory, algebra and probability theory, we have successfully constructed a 2D inference engine. We have shown a way to approach medical image representation with a more biological flavor. Each group in the cascade engine has a biological meaning in a sense that it represents biological properties of the subject or physical properties of the sensors. We are currently making the algorithms

for handling human anatomies more robust. Such an endeavor involves fine tuning parameters to account for the vast variability that occurs in nature and to eliminate noise generated by the sensors.

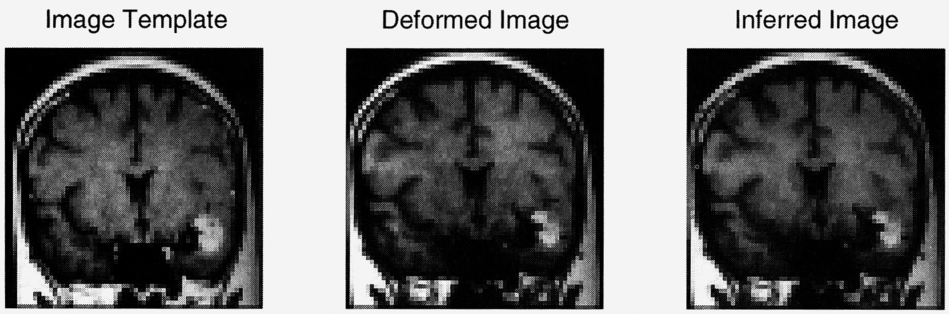

FIG. 6. A head cross-sectional

Acknowledgments. I am grateful to Ulf Grenander for inspiring and influencing this work and for his helpful advice. I would also like to thank Sarang Joshi for supplying some of the brain data. This work was done under the support of the ARO grant DAAL03-92-G-0115.

REFERENCES

[1] J. Besag, *Spatial interaction and the statistical analysis of lattice systems*, J. R. Statist. Soc. B, vol. 36, pp. 192–326, 1974.

[2] U. Grenander, *Pattern Theory: A Mathematical Study of Regular Structures*, Oxford University Press, Oxford, 1993.

[3] F. L. Bookstein, *The Measurement of Biological Shape and Shape Change*, vol. 24, Springer-Verlag: Lecture Notes in Biomathematics, New York, 1978.

[4] F. L. Bookstein, *Biometrics, biomathematics and the morphometric synthesis*, Bulletin of Mathematical Biology, vol. 58, no. 2, pp. 313–365, 1996.

[5] D. Terzopoulos and K. Waters, *Analysis and synthesis of facial image sequences using physical and anatomical models*, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 15, no. 6, pp. 569–579, 1993.

[6] R. Bajcsy, R. Lieberson, and M. Reivich, *A computerized system for the elastic matching of deformed radiographic images to idealized atlas images*, Journal of Computer Assisted Tomography, vol. 7, no. 4, pp. 618–625, 1983.

[7] G. Christensen, *Deformable Shape Models for Anatomy*, Ph.D. Dissertation, Department of Electrical Engineering, Sever Institute of Technology, Washington University, St. Louis, MO, Aug. 1994.

[8] C. Broit, *Optimal registration of deformed images*, Dissertation, University of Pennsylvania, Philadelphia, Pennsylvania, 1983.

[9] D. L. Snyder, *Random Point Processes*, John Wiley and Sons, 1975.

[10] U. Grenander, Y. Chow, and D. M. Keenan, *Hands: A Pattern Theoretic Study of Biological Shapes*, Springer-Verlag, New York, 1991.

[11] M. V. Ranganath, A. P. Dhawan, and N. Mullani, *A multigrid expectation maximization algorithm for positron emission tomography*, IEEE Trans. on Medical Imaging, vol. 7, pp. 273–278, 1988.

[12] B. Gidas, *A renormalization group approach to image processing problems*, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. PAMI-11, pp. 164–180, 1989.

[13] A. D. Sokal, *Monte Carlo methods in statistical mechanics: Foundations and new algorithms*, Cours de Troisieme cycle de la physique en suisse romande, June 1989.

[14] 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.
[15] L. Matejic, *Group cascades for representing biological variability in medical images*, Brown University Technical Report, 1996.
[16] F. L. Bookstein and W. D. K. Green, *Edge information at landmarks in medical images*, in Visualization in Biomedical Computing 1992, Richard A. Robb, Ed., 1992, pp. 242–258, SPIE 1808.
[17] E. Butkov, *Mathematical Physics*, Addison-Wesley, 1968.