Unwarping of Unidirectionally Distorted EPI Images

Jan Kybic*, *Student Member, IEEE*, Philippe Thévenaz, *Member, IEEE*, Arto Nirkko, and Michael Unser, *Fellow, IEEE*

***Abstract*—Echo-planar imaging (EPI) is a fast nuclear magnetic resonance imaging (MRI) method. Unfortunately, local magnetic field inhomogeneities induced mainly by the subject's presence cause significant geometrical distortion, predominantly along the phase-encoding direction, which must be undone to allow for meaningful further processing. So far, this aspect has been too often neglected.**

In this paper, we suggest a new approach using an algorithm specifically developed for the automatic registration of distorted EPI images with corresponding anatomically correct MRI images. We model the deformation field with splines, which gives us a great deal of flexibility, while comprising the affine transform as a special case. The registration criterion is least squares. Interestingly, the complexity of its evaluation does not depend on the resolution of the control grid. The spline model gives us good accuracy thanks to its high approximation order. The short support of splines leads to a fast algorithm. A multiresolution approach yields robustness and additional speedup.

The algorithm was tested on real as well as synthetic data, and the results were compared with a manual method. A wavelet-based Sobolev-type random deformation generator was developed for testing purposes. A blind test indicates that the proposed automatic method is faster, more reliable, and more precise than the manual one.

***Index Terms*—Geometrical distortion, image registration, splines, unwarping.**

I. Introduction

A. EPI Features

ECHO planar imaging (EPI) [1] is a fast magnetic resonance imaging (MRI) technique permitting an acquisition of a two-dimensional (2-D) slice using a single excitation, which leads to very short scan times. It is used mainly for functional imaging (fMRI), the *in vivo* noninvasive study of the temporal, spatial, and behavioral dependencies of brain activations. The basis of fMRI lies in the fact that deoxyhemoglobin (the hemoglobin without a bound oxygen molecule) is paramagnetic. Neural activation in the cerebral cortex leads to an increase of blood flow and, hence, to a decrease of deoxyhemoglobin concentration.[1] This results in a measurable alteration of the magnetic field and in a consequent increase of signal intensity in the appropriately weighted MRI images [blood oxygen-level dependent (BOLD)]. It is therefore difficult to compensate for the unwanted magnetic field inhomogeneities induced mainly by the spatially varying magnetic susceptibility of the subject [2]. In contrast to conventional MRI, where the number of excitations per slice is equal to the number of scan lines, in EPI the magnetic field gradients must encode two coordinates simultaneously in one excitation. As one of the gradients (the so-called phase-encoding gradient) is several orders of magnitude weaker than the other, the inhomogeneous magnetic field will manifest itself mainly as a geometrical distortion of the 2-D slice image along the direction of this gradient. This effect is clearly visible in Fig. 1. Since the stronger gradient is less affected, the distortion is essentially unidirectional. Letting g be the unknown warping (deformation) function, we have

$$f^o(g(x, y), y) \simeq f^u(x, y) \tag{1}$$

where f^o is the observed EPI image and f^u is the hypothetical ideal undistorted EPI image. We can consider each slice separately, as the shift in the z axis due to patient's movement is insignificant because his head is attached. Should there be such a displacement, it can by readily corrected by existing algorithms [3].

B. The Reasons to Unwarp

The amplitude of the deformation g can be as large as 3–5 mm [4] (confirmed by our own observations), which typically amounts to several pixels. In some cases, as in Fig. 1, specifically intended to illustrate EPI distortion, the deformation can be even more pronounced. Moreover, g can vary significantly from slice to slice and from acquisition to acquisition. For localization applications like stereotactic surgery, this inaccuracy is much larger than the required limit of 1 mm and therefore EPI cannot currently be used to this end. It also severely hinders the performance of the statistical processing of sets of fMRI images used to obtain activation information. Since the task-induced signal changes represent typically only 5–10% of the mean signal intensity in fMRI [1], [5], they will not stand out clearly unless the perturbations caused by the deformation g are undone.

C. Existing Distortion Correction Techniques

One approach consists in changing the acquisition procedure [2], [4], [6]. However, this is often not practical due to technical or organizational limitations, for example, lack of support or approval. Furthermore, while the alternative acquisition sequences reduce the distortion, the distortion is never removed completely and the methods usually sacrifice either sensitivity or acquisition speed.

The second group of methods uses a two-step procedure [4], [7]. First, a field map or a deformation map is obtained, e.g., from an image of a phantom. In the second step this information is used to compensate for the deformation on real images.

Manuscript received July 9, 1999; revised December 13, 1999. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was N. Ayache. *Asterisk indicates corresponding author.*

*J. Kybic, P. Thévenaz, and M. Unser are with Biomedical Imaging Group, DMT/IOA, Swiss Federal Institute of Technology Lausanne, CH-1015 Lausanne EPFL, Switzerland (e-mail: Jan.Kybic@epfl.ch).

A. Nirkko is with Inselspital, Bern, Switzerland.

Publisher Item Identifier S 0278-0062(00)02300-4.

[1]This effect prevails over the increase of oxygen consumption.

0278–0062/00$10.00 © 2000 IEEE

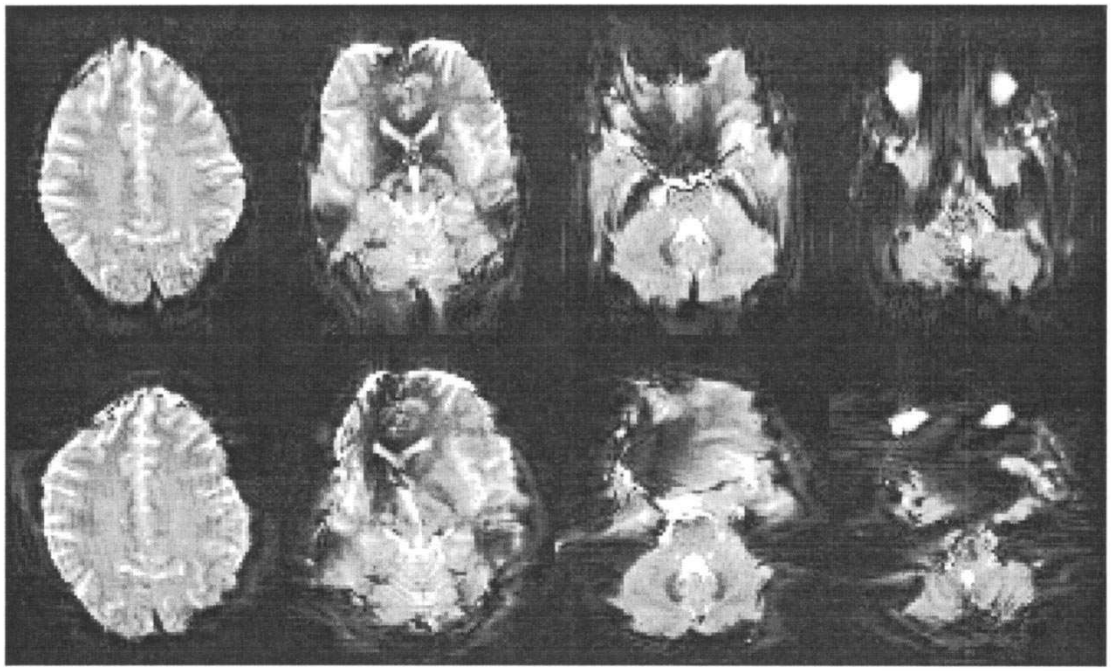


Fig. 1. Demonstration of EPI distortion. Selected EPI brain slices taken with two different phase-encoding gradient orientations. (Top) Anterior–posterior. (Bottom) Left–right. The vertical, respectively, horizontal deformation of the upper, respectively, bottom row images is clearly visible.

The major drawback of these methods is that it is impractical to build a phantom that would exactly duplicate the biological system being imaged [4], which limits the compensation only to field distortions other than those caused by the individual subject. Moreover, these other distortions are most likely already compensated for by the scanner manufacturer or operator.[2]

D. Unwarping by Registration

We propose a third approach which, to the best of our knowledge, has never been applied to this particular problem. It consists of registering the distorted EPI image with a corresponding geometrically correct anatomical MRI image. In this way, we can recover the deformation g from a single EPI slice obtained by an unaltered standard procedure. The registration can be performed manually [8], but this is tedious, time consuming, and prone to errors. An automatic procedure is advantageous because it is faster, more precise, and does not require an expert.

Our goal in this paper is to present an automatic registration algorithm we have developed specifically for this problem; i.e., identifying a nonlinear unidirectional 2-D warping. More precisely, given the observed EPI slice f^o and the corresponding undistorted anatomical MRI reference slice f^r, the task is to find a warping g so that the warped test image $f^o(g(x, y), y)$ matches as well as possible (in a sense defined later) the reference image f^r.

II. Proposed Algorithm

We categorize registration algorithms according to the warp space used. In general, a deformation function g_V from a warp space V is described by a finite set of parameters c_k by means of a warping model.

At one end of the scale we have nonparametric local methods. These methods are formulated either as variational, defining a scalar criterion to minimize, or (more generally) using PDE's. The continuously defined correspondence function that minimizes a given criterion (respectively, that solves a given PDE) is sought for in a very large and unrestrictive function space, e.g., the Sobolev space W_2^2. The essence of these methods is entirely in the criterion (respectively, PDE). The PDE come from the optical flow approach (gradient methods) [9], viscous fluid model [10]–[12], elastic deformations with physical analogs [13], [14], or without [15]. Some elastic deformations can also be modeled as potential fields [16].

At the other end, we have parametric global methods that describe the correspondence function using a global model with a relatively small number of parameters [17]. The model mostly consists of expressing the warping function in a linear [3], global polynomial [18], [19], or harmonic basis [20]. For these methods, the deformation model corresponding to a specific warp space is as important as the criterion being minimized.

In this article, we consider mainly intensity-based registration methods, which directly take into account the voxel values (c.f., [21]). Other methods are based on matching surfaces [21]–[23], curves [24], or interpolating landmarks using radial basis functions, especially thin-plate splines [21], [25], [27].

A. Semilocal Model

The model proposed in this article is situated between the above-mentioned local and global methods, combining the advantages of both.

We parametrize the warp space by a scale parameter h and denote it V_h. The scale parameter corresponds loosely to the

[2]This procedure is called shimming and is generally repeated before each series of acquisitions.

density of knots or landmarks. By changing h, we can approach either of the two limit cases or choose a compromise offering the best tradeoff. Big h yields a global model which has just a few parameters. Such a model is rather constrained, which is approximatively equivalent to strong explicit regularization. On the other hand, a small h gives a local model with many parameters, which generally leads to more complicated optimization. In exchange, a small h generally permits one to approximate any given g in V_h well, because the space V_h is big. (Arbitrarily small precision can be achieved as $h \to 0$.) This roughly corresponds to weak explicit regularization. Thus, the scale parameter can partly assume the role of an explicit regularization factor, unlike in local methods where the regularization is a part of the criterion.

In the next section, we give our motivation for the algorithm and for our particular choice of the warp model—uniform *cubic splines* represented by a linear combination of cubic *B-splines* [28], [29].

B. Univariate Case

For the sake of explanation, let us begin with the one-dimensional case (1-D), i.e., with the task of recovering an univariate warping function g: $\mathbb{R} \to \mathbb{R}$. An example of such a function could be the affine map $g(x) = \alpha x + \beta$. Such linear dependence is rather frequently encountered in practice; for example, it arises when the acquisition techniques use different coordinate systems or when there is a movement between acquisitions.

The univariate equivalent of our matching problem becomes $f^o(g(x)) \simeq f^r(x)$. The landmark method would consist of (either manually or automatically) identifying a set of landmark pairs $\{(x_l^r, x_l^o)\}_{l=1}^L$ so that a feature found at location x_l^r in the reference image can be found (or associated with) a location x_l^o in the observed test image. For interpolation, this gives a set of L constraints for g: $g(x_l^r) = x_l^o$. To get a well-posed problem we shall require g to minimize some criterion J. We do not want J to penalize linear dependencies, i.e., we want J to be zero for linear g. We want the resulting $g = \arg\min J$ to be linear with respect to the landmark coordinates $\{x_l^o\}$ and invariant with respect to a linear (affine) transformation of $\{x_l^r\}$, which means that if we take a linear combination of two sets of landmark coordinates then the resulting g should be the same linear combination of the solutions corresponding to the two sets of landmarks. In other words, we want a solution that is invariant to the choice of a particular coordinate system, or to the choice of units. The simplest criterion satisfying these requirements is $J = \int |g''(x)|^2\, dx$, which is compatible with the elasticity theory, as it corresponds to strain or bending energy [25].

In addition to interpolation other approximation schemes can be applied, the most popular being least squares fitting. It consists of minimizing an extended criterion $J' = J + \gamma \sum_l (g(x_l^r) - x_l^o)^2$. This has the advantage of accommodating uncertainty (noise) in landmark positions.

For both interpolation and least squares fitting, as well as for any other criterion of the general form $J' = J + \sum_l \mathcal{L}(g(x_l^r), x_l^o)$ (with arbitrary function $\mathcal{L}$), the function g that minimizes the criterion can be shown to be a cubic spline [30], [31], i.e., a piecewise cubic function that is twice continuously differentiable, with knots (the boundaries between polynomial pieces) at points $\{x_l^r\}$. In Appendix A, we show that the solution can be expressed as a linear combination of radial basis functions $|x - x_l|^3$. However, these functions are not convenient to work with because of their instability and global support. Fortunately, it is possible to localize the $|x - x_l|^3$ functions using divided differences, which yields a base made of cubic B-splines with local support that generates the same space [25].

The automatic landmark method is rather difficult to apply because there is no automatic landmark detection algorithm available that would be sufficiently robust and precise, especially for our class of medical images which typically exhibit only a few distinct features. Moreover, it is difficult to automatically find common features in both anatomical MRI and EPI modalities. For this reason, instead of trying to work with landmarks we introduce a data criterion E (defined in Section II-D) taking into account the entire image and measuring the discrepancy between the warped version of the observed image $f^o(g(x))$ and the reference image $f^r(x)$. Then, we seek a deformation g such that the two images are as similar as possible, i.e., when E is minimized.

For the reasons mentioned above, we choose to search the warping g in the cubic spline space as in the landmark case. However, there are now no explicit landmarks available to put the knots on. We also do not know, how much useful information each part of the image can provide. Therefore, we will distribute the knots uniformly over the image. It follows that the function g will be a uniform cubic spline, which can be uniquely represented as a linear combination of uniformly spaced cubic B-splines [28], [29]:

$$g(x) = \sum_{k \in K} c_k \beta_3(x/h - k) \tag{2}$$

where K is the set of the indexes of the spline functions, the support of which intersects with the image and h is the knot spacing (the B-splines will be centered at points kh for $k \in K \subseteq \mathbb{Z}$). Working with uniform splines is also significantly faster with respect to nonuniform splines (see Appendix B for a definition of a cubic B-spline). Note that in order to get a complete control over g, it is useful to put some spline knots outside the image. For cubic splines, we need to put one such exterior knot at each side. Consequently, for an image size N and knot spacing h, we have $N/h + 2$ knots.

We have thus transformed the registration task into a nonlinear finite-dimensional optimization problem: find a set of coefficients $\boldsymbol{c}$ minimizing some criterion E.

C. Splines—A Perfect Fit

Let us now show several important properties of the spline model.

1) Good Approximation Properties: The error of a cubic spline approximation decreases asymptotically as h^4 (measured by any L_p or l_p norm, $p \in \{1, 2, \cdots, \infty\}$). Quantitative

analyzes indicate that splines perform well in comparison with other wavelet-like basis functions [33].

2) Speed: Cubic splines have a short compact support of length four. They are symmetric and piecewise cubic. To evaluate $\beta_3(x)$ at one particular point only five arithmetic operations (additions or multiplications) and three comparisons are needed. In multiple dimensions, where we will use tensor products of cubic splines as basis functions (see Section II-E), the computational complexity stays low thanks to separability. The number of operations needed to evaluate g also does not depend on the total number of basis functions (and thus the number of parameters c_k).

3) Plausibility: The spline model corresponds to a wide range of physical situations where the restoring force can be approximated as being linearly dependent on the displacement. In such situations, the generated deformation is physically plausible. It is also a good approximation for cases when a better model is not known, such as the deformation of EPI images, even though it is not of mechanical origins.

4) Simplicity: The model is linear in the parameters c_k and polynomial with respect to the position x. It is thus possible to truncate the Taylor expansion such that it is exact in some neighborhood of x with a typical diameter of $h/2$.

5) Scalability: Thanks to the q-scale relation $\beta_3(x/q) = \sum_k \zeta(k)\beta_3(x-k)$ where $q \in \mathbb{N}$, we have the embedding $V_{hq} \subseteq V_h$, i.e., the transition from a coarse space V_{hq} to a finer space V_h is exact [34].

D. Data Criterion

A reasonable and most often [3] used way to measure the discrepancy between two images is the sum of squared differences (SSD) criterion

$$E = \sum_{i \in I} (f^o(g(i)) - f^r(i))^2 \tag{3}$$

where the sum is over all pixels in the image. Note that minimizing this criterion is equivalent to calculating the maximum likelihood estimate of the unknown parameters assuming that the difference is an independently identically distributed (i.i.d.) Gaussian noise and that the true test image is indeed a geometrically distorted version of the reference one. Moreover, the SSD criterion is also algorithmically advantageous because it is easy to evaluate (including its derivatives) and because it depends smoothly on the parameters. Interpolation (cf. Section II-G) is needed to evaluate this criterion, as it calls for values of f^o on generally noninteger coordinates.

E. Bivariate Case

The transition to the bivariate case is straightforward. The criterion becomes

$$E = \sum_{\boldsymbol{i} \in I} (f^o(\boldsymbol{g}(\boldsymbol{i})) - f^r(\boldsymbol{i}))^2 \tag{4}$$

where we have taken the convenience notation

$$\boldsymbol{g}(\boldsymbol{i}) = \boldsymbol{g}(i, j) = \begin{bmatrix} g(i, j) \\ j \end{bmatrix} \tag{5}$$

and where the warping function is now described by a 2-D array of coefficients

$$\begin{aligned} g(\boldsymbol{x}) &= g(x, y) \\ &= \sum_{\boldsymbol{k} \in K} c_{\boldsymbol{k}} \boldsymbol{\beta}_3(\boldsymbol{x}/\boldsymbol{h} - \boldsymbol{k}) \\ &= \sum_{(k_x, k_y) \in K} c_{k_x, k_y} \beta_3(x/h_x - k_x)\beta_3(y/h_y - k_y). \end{aligned} \tag{6}$$

Tensor products of splines were also used in [35].

F. Optimization Algorithm

The criterion is minimized with respect to the coefficients c_k using a regularized version of the Newton method [3], [36] inspired by the Marquardt–Levenberg algorithm. This algorithm smoothly varies between the gradient-descent and the Newton approach, which gives it robustness and quadratic convergence near the optimum.

The algorithm uses first two derivatives of the criterion E with respect to c_k, i.e., $\nabla_{\boldsymbol{k}} E$ and $\nabla^2_{\boldsymbol{k}} E$. Thanks to the spline representation (2), the derivatives can be calculated exactly and at a small cost. As the number of components each pixel contributes to remains constant, the cost of evaluating E, $\nabla_{\boldsymbol{k}} E$ and $\nabla^2_{\boldsymbol{k}} E$ does not depend on the number of coefficients c_k (or, equivalently, the spacings h_x, h_y) used to describe the deformation (see Appendix C for explicit formulas).

At each step we update the vector of all coefficients $\boldsymbol{c}$ to $\boldsymbol{c}+\Delta\boldsymbol{c}$ by taking

$$-(\nabla^2_c E(\boldsymbol{c}) + \lambda \boldsymbol{I})\Delta\boldsymbol{c} = \nabla_c E(\boldsymbol{c}) \tag{7}$$

where the regularization factor λ is divided by a constant λ_f if the previous step resulted in a decrease in E, otherwise it is multiplied by the same amount.[3] We iterate as long as the relative and absolute change of E stay above *a priori* given thresholds.

When the number of coefficients exceeds a certain limit, the Hessian matrix $\nabla^2_c E$ gets too big for the linear equation set (7) to be efficiently solvable and this algorithm ceases to be practicable. Note that the theoretical computational complexity of solving (7) increases as h^{-6}. The asymptotical memory requirements grow as h^{-4}. While iterative linear equation solvers generally speed up the solution of (7), it is at the cost of a loss in precision and the overall gain in our case is insignificant.

When the size of the Hessian matrix exceeds our computational capacities, we replace the Marquardt–Levenberg optimizer by a conjugate gradient method [36], which converges quadratically too without explicitly calculating and storing the Hessian matrix. Even though the conjugate gradient method needs more iterations for smaller problems, it outperforms Marquard–Levenberg for bigger ones.

G. Image Interpolation Model

To calculate the derivatives, as well as to evaluate the criterion (4), an image interpolation model is needed to get a continuous form ${}^c f(\boldsymbol{x})$ from a discrete image $f^o(\boldsymbol{i})$. Because of their good

[3] We use $\lambda_f = 10$.

approximation properties, simple analytic form, and effective algorithms available, we use cubic B-splines here as well

$$ {}^{c}f(x, j) = \sum_{i} b_{ij}\beta_3(x - i) \qquad \text{where} $$
$$ {}^{c}f(i, j) = f^{o}(i, j), \forall (i, j) \in I. \tag{8} $$

The coefficients b_{ij} can be obtained prior to registration by an efficient filtering algorithm [29] which incurs negligible overhead. For the filtering, we are using mirror boundary conditions on the image. In this way we have the same number of coefficients b_{ij} as there are pixels in the original image.

H. Multiresolution

The robustness and efficiency of the algorithm can be significantly improved by a multiresolution approach. The task at hand is first solved at a coarse scale. Then, the results are propagated to the next finer level and used as a starting guess for solving the task at that level. This procedure is iterated until the finest level is reached.

In our algorithm, multiresolution is used twice. First, we build an image pyramid, a set of gradually reduced versions of the original image [37]. This pyramid is compatible with our image representation (8) and is optimal in the l_2 sense [i.e., in the sense of the criterion (4)], which ensures that the approximation made by substituting the lower resolution image is the best possible. Based on an image of size $N_x \times N_y$, we create a sequence of images with sizes $\mathcal{M}_I = \{(N_x^1, N_y^1), (N_x^2, N_y^2), \cdots, (N_x^{m_I}, N_y^{m_I})\}$, where $N_x^{m_I} = N_x$, $N_y^{m_I} = N_y$ and where $N_x^j = \lfloor N_x^{j+1}/2 \rfloor$, $N_y^j = \lfloor N_y^{j+1}/2 \rfloor$. The optimum starting size (N_x^1, N_y^1) depends on the image. We chose $16 < \min(N_x^1, N_y^1) \leq 32$, which works well in most cases.

Second, we use multiresolution for the warping function as well. We start with a deformation g described with very few parameters c_k and with a large distance h between knots. After the optimization of c_k is complete, we halve the distance between knots. This approximately corresponds to doubling the number of knots in each direction, i.e., quadrupling the number of coefficients c_k. Because of the two-scale spline relation, we can exactly represent the warping function from the old coarse space in the new finer space. More precisely, the sequence of knot spacings is going to be $\mathcal{M}_W = \{(h_x^1, h_y^1), (h_x^2, h_y^2), \cdots, (h_x^{m_W}, h_y^{m_W})\}$ where $h_x^1 = N_x$, $h_y^1 = N_y$; the final element $(h_x^{m_W}, h_y^{m_W})$ corresponds to the user chosen target grid size. The sequence obeys $h_x^{j+1} = h_x^j/2$, $h_y^{j+1} = h_y^j/2$. The process starts with g being identity.

The global strategy that combines the two multiresolutions is depicted in Fig. 2. Symbolically, to construct the double multiresolution $\mathcal{M}$ we first extend the shorter of $\mathcal{M}_I$, $\mathcal{M}_W$ to the length $m = \max(m_I, m_W)$ by repeating the last element. Then, $\mathcal{M}$ consists of pairs of elements from $\mathcal{M}_I$, $\mathcal{M}_W$ in alternate progression: $\mathcal{M} = \{(\mathcal{M}_I^1, \mathcal{M}_W^1), (\mathcal{M}_I^1, \mathcal{M}_W^2), (\mathcal{M}_I^2, \mathcal{M}_W^2), (\mathcal{M}_I^2, \mathcal{M}_W^3), \cdots, (\mathcal{M}_I^m, \mathcal{M}_W^m)\}$.

The consequence of using multiresolution is that the algorithm works best for images and deformations that follow the multiresolution model, i.e., when a low-resolution version is a good approximation of the finer resolution version.

I. Invertibility

We implicitly assumed that the deformation $\boldsymbol{g}$ is recoverable and recoverability requires invertibility. In case the EPI artifacts cause the signal from adjacent pixels to blend due to the fold-over effect, the invertibility condition might not be satisfied and the deformation is not recoverable. However, this does not happen in practice. To ensure the stability of the registration it is necessary also to enforce the invertibility on the trial solutions during the optimization process, i.e., to perform constrained optimization. If we assume that the deformation conserves orientation, the sufficient condition ensuring local invertibility is the positivity of the Jacobian $\det(\nabla_{\boldsymbol{x}}\boldsymbol{g}) > 0$ everywhere. In our case, the invertibility condition reduces to $\forall (x, y) \in S$; $\partial g/\partial x > 0$ where S is the image domain.

Although an iterative algorithm can be found verifying this condition exactly, it is not practical to apply it for performance reasons. Instead of checking that $\partial g/\partial x > 0$ on a continuous domain, we sample this condition on pixel coordinates. At this scale, the two conditions are essentially equivalent.

There is little hope of finding an algorithm capable of solving a constrained optimization problem of our complexity (highly nonlinear criterion, hundreds of parameters, and tens of thousands of constraints) in a reasonable time. We have therefore chosen to convert the constrained search into an unconstrained one using an exponential penalty cost function. A set of constraints

$$ 0 < u_{\boldsymbol{p}} = \left.\frac{\partial g}{\partial x}\right|_{(x, y)=\boldsymbol{p}} \tag{9} $$

is replaced by a penalty function E_p

$$ E_p = \gamma \sum_{\boldsymbol{p}} e^{-\alpha u_{\boldsymbol{p}}}. \tag{10} $$

Consequently, we will minimize the combined criterion

$$ E_c = E + E_p. \tag{11} $$

The choice of the constants α and γ is a tradeoff between speed and precision. As α increases the penalty function gets steeper, which improves precision in the vicinity of the constraints. At the same time, the criterion becomes highly nonlinear, which slows down the optimization. We set α and γ so that $E_p \approx E$ on the boundary of the permissible space and $E_p \approx 10^{-5}E$ for the initial configuration when g is an identity.

J. Preprocessing

To apply the SSD criterion, we need to make the test and reference images more similar, so that their difference after warping is as close as possible to white noise. We choose therefore to apply a preprocessing step that consists of high-pass filtering and histogram equalization. The effect can be seen in Fig. 3.

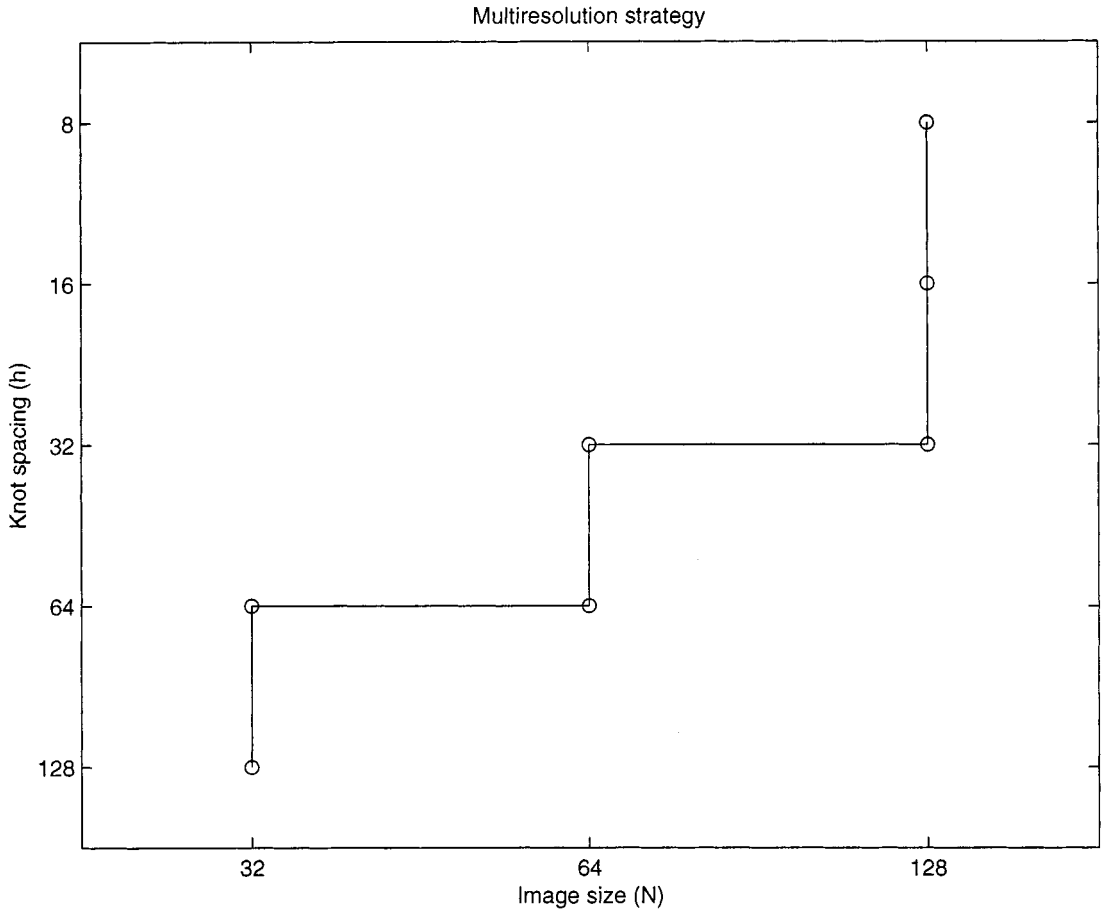


Fig. 2. The multiresolution strategy starts with a small version of the image (horizontal axis) and a small warping control grid (vertical axis). After several steps, and by augmenting resolution alternatively in the two domains, we reach the original image resolution and the desired size of the control grid.

The preprocessing makes it unnecessary to add a special parameter accounting for differences in intensity profiles of the two images. It also helps to compensate for the intensity distortion due to in-plane dephasing, which is also caused by the magnetic field inhomogeneity. The dephasing cannot be compensated completely, as it is impossible to differentiate between effects of dephasing and warping on a single pixel.

III. Experiments

A. General Comments

We tested the performance of our algorithm on several hundreds of images. In addition, we compare to the automatic method the results of the registration of 30 image pairs by three different people, including one experienced practitioner. For the manual registration, we use the standard thin-plate spline method [8], [25], [38].

Unless stated otherwise, the tests are performed on 128 × 128-pixel spin-echo anatomical MRI images of the brain, on a SUN Ultra workstation, and the published numbers are arithmetical means of the results of experiments made on all of the 30 horizontal slices of a brain volume. A typical pair consisting of corresponding anatomical and EPI images is shown in Fig. 3(a) and (b).

B. Error measurement

As a main measure to compare different solutions to the registration problem, we use the warping index defined in [3] as $\varpi = (1/\|R\|) \sum_{\boldsymbol{i} \in R} |g(\boldsymbol{i}) - g^*(\boldsymbol{i})|$ where g^* is the reference solution (ideally, the true deformation) and $R \subseteq I$ is the region of interest (in our case, the interior of the brain). This corresponds to the mean per-pixel precision of the result.

We shall also use the SSD, as defined by (3). It measures the similarity of the reference and warped test images as perceived by the algorithm and corresponds to the quantity minimized.

C. Sources of Error

There are several reasons why we cannot expect a perfect registration (indicated by $\varpi = 0$).

1) *Different Images*: Despite the assumptions made when deriving our criterion, the images we are asked to register are not geometrically deformed versions of each other. To this we must add the effects of discretization, quantization, and noise. Any of them may result in spurious minima of the criterion E, misleading the algorithm. This problem can be alleviated by preprocessing, or by choosing a different criterion, but it can never be completely removed. The situation is more favorable when a small number of

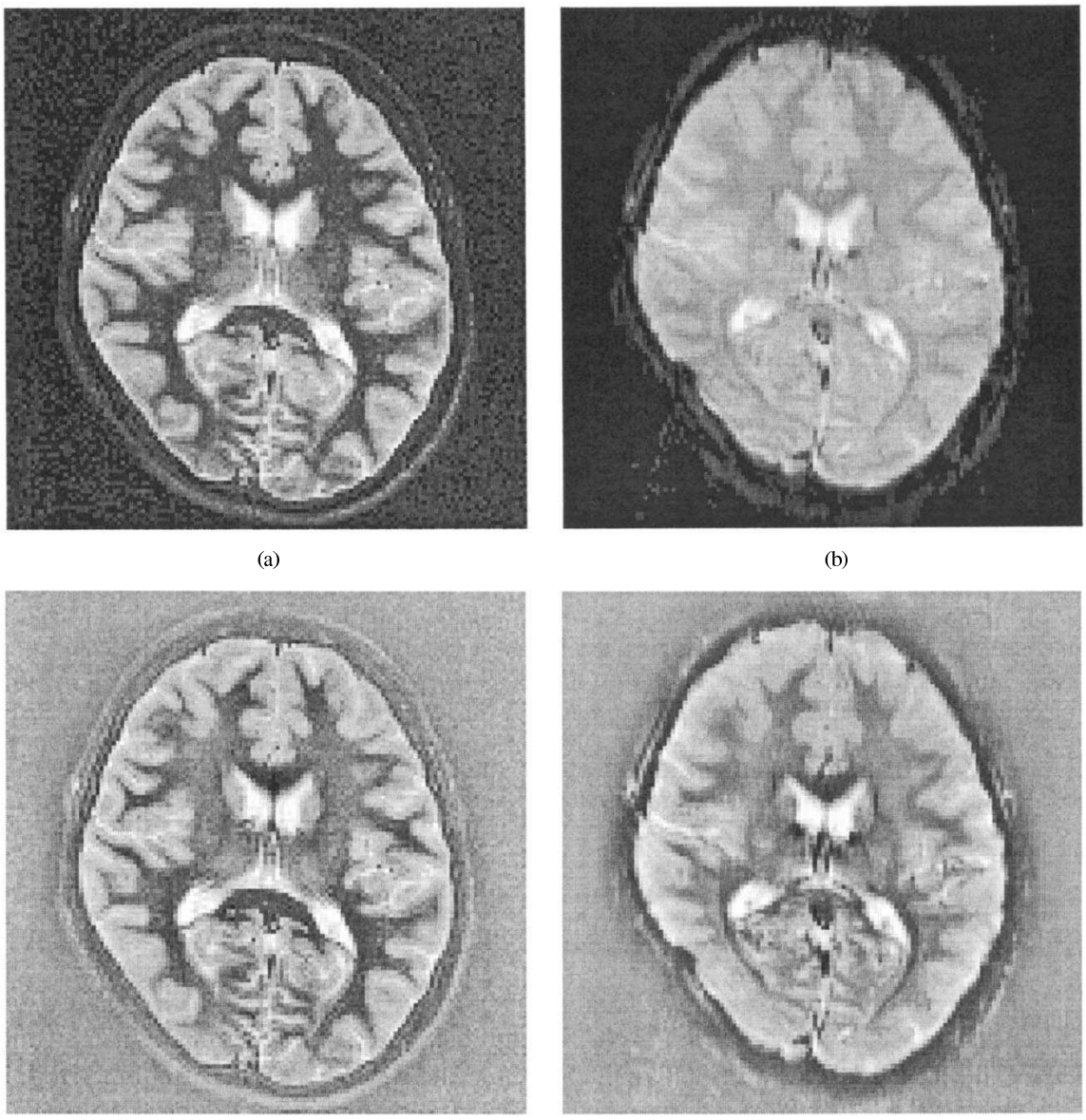


Fig. 3. (a) An anatomical MRI image. (b) A corresponding EPI image. Both images represent a horizontal slice through the middle of a human brain. (c) and (d) The same images after preprocessing.

parameters is sought, for the random effects tend to cancel out.

2) *Local Minima*: The algorithm with its local vision cannot distinguish between local and global minima of the criterion and can therefore get trapped in a local minimum. Multiresolution improves significantly this aspect, but it does not solve it entirely. Again, the situation is much more favorable for a small number of parameters.
3) *Warp Space*: The true warping does not necessarily belong to the warp space V_h. We define $\varpi_{\min}$ as the minimum achievable registration error, which is equal to the distance between the true warping and the closest point in the warp space V. We call this closest point (which is the best possible approximation of g in V) a projection $\mathcal{P}_V g$. In the L_2 or l_2 sense it is equal to $\sum_k \mu_k \varphi_k$ where the coefficients μ_k verify $\langle g, \varphi_l \rangle = \sum_k \mu_k \langle \varphi_k, \varphi_l \rangle$ and where we have denoted the basis functions by φ_k in order to simplify the notation. For other norms, iterative procedure must be used. In a controlled experimental environment we can thus calculate $\varpi_{\min}$ exactly and compare it with the results of our algorithm.
4) *Lack of Details*: The precision is limited by the local resolution of the images. When two corresponding sharp edges occur in both images, subpixel registration precision is often attainable. On the other hand, in homogeneous textureless regions there is little hope of recovering any information whatsoever.
5) *Criterion Surface Complexity*: Depending on the images being registered, the dependence of the criterion on the parameters can be nonlinear, nonconvex, and generally very complicated. In such a situation the convergence is extremely slow. If the time is limited, we must stop the algorithm before any significant improvement has been realized.

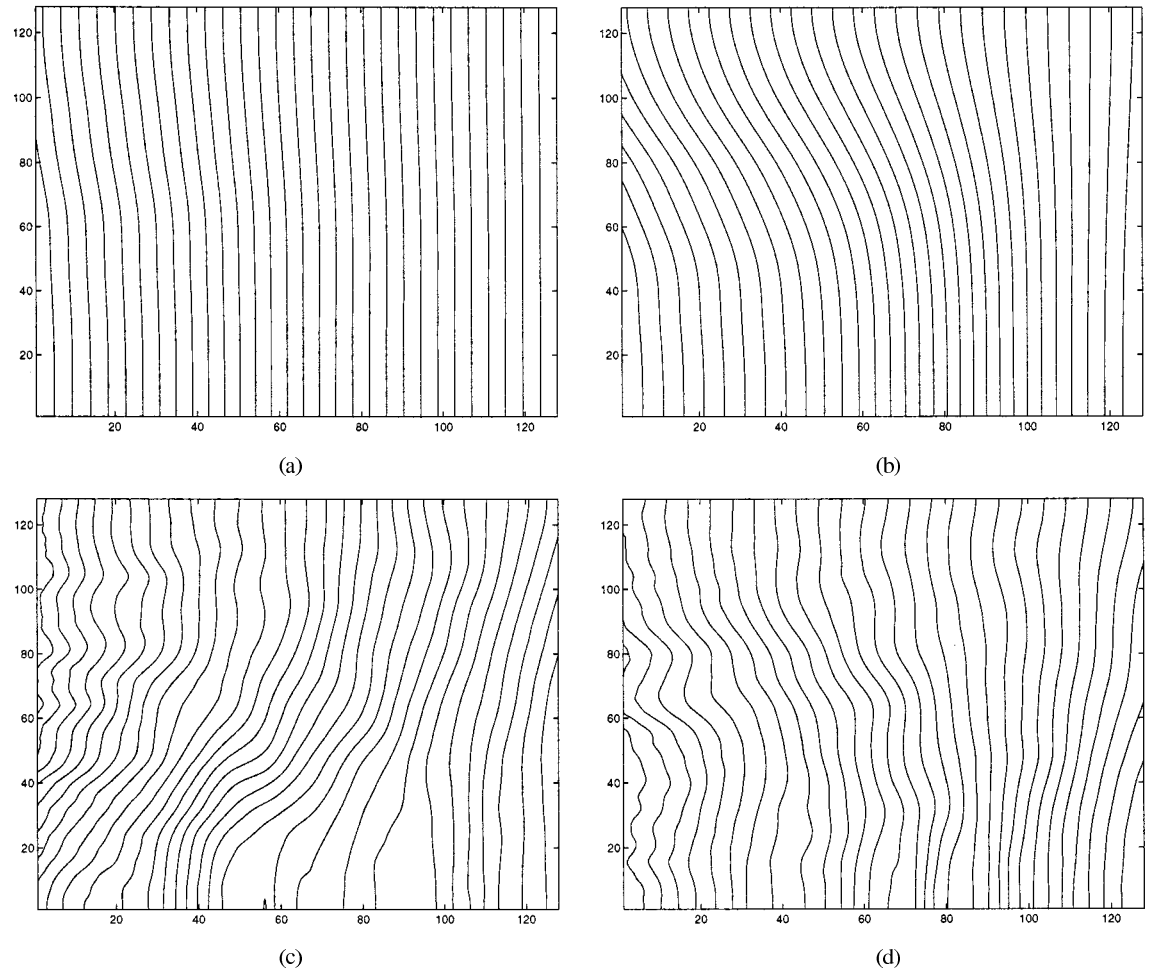


Fig. 4. Examples of deformation presented as contour plots. (a) and (b) were generated with $W_H(10,\ 1.6)$ and (c) and (d) with $W_H(10,\ 0.6)$.

6) *Numerical Precision*: Insufficient numerical precision can hinder the performance of the optimizer. We have encountered this problem only rarely, when iterative methods were used to solve the normal equation set (7).

We have tried to design the experiments to separate the effect of the above-mentioned factors whenever possible. However, this will rarely be the case in a real-life situation.

D. Deformation Generator

To test our algorithm, we have implemented a wavelet-based deformation generator. We want to generate a random Sobolev-type deformation: a deformation lying in a prescribed Sobolev space W_2^r. The parameter r refers to the regularity. It is equal to the number of derivatives in the L_2 sense and it is also strongly related to pointwise differentiability. The higher the r, the more regular are the functions from W_2^r. Wavelets are known to be good bases for functions lying in Sobolev spaces and the decay of wavelet coefficients across scales is directly related to a Sobolev-type regularity. Let $\theta_{j,k}$ denote the coefficients of a wavelet expansion[4]

$$g(x) - g_I(x) = \sum_{j,k} \theta_{j,k}\, \psi(2^j x - k) \tag{12}$$

[4]For brevity, we shall deal here with the unidimensional case only.

where g_I is an identity transform and ψ is an orthogonal wavelet. Then the displacement $g - g_I$ belongs to a Sobolev space W_2^r if and only if

$$\sum_{j,k} 2^{2jr} |\theta_{j,k}|^2 < \infty \tag{13}$$

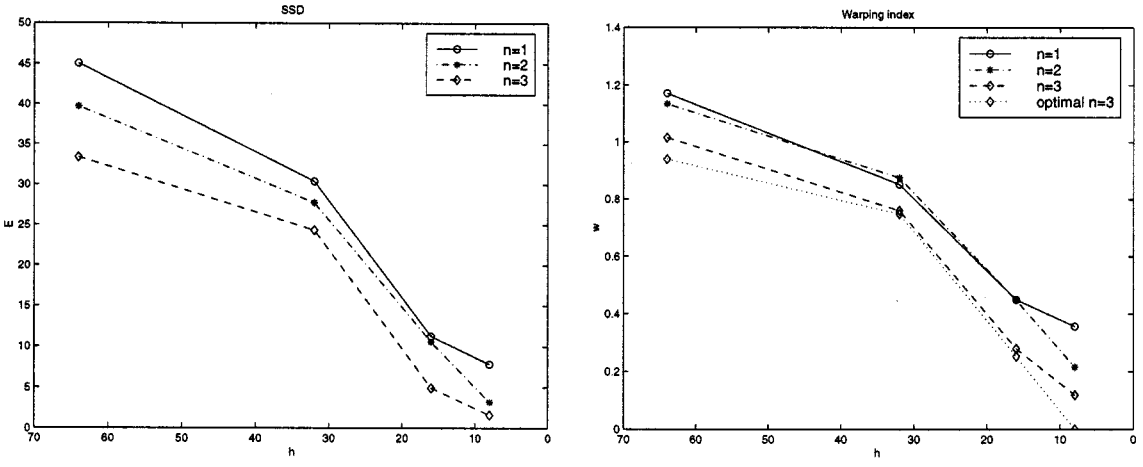
provided that the regularity of ψ is greater than r [39]. It follows that for (13) to hold, the necessary condition is

$$\exists\, C \in \mathbb{R};\ \forall\, j,\ k;\ |\theta_{j,k}|^2 < C\,\xi^j, \qquad \text{with } \xi = 2^{-2r}. \tag{14}$$

Practically, we generate our deformations using zero-mean normally distributed coefficients with variance

$$\boldsymbol{E}\{\theta_{j,k}^2\} = \sigma_0^2\, \xi^j \tag{15}$$

where σ_0^2 governs the total energy of the deformation. Such a deformation will be denoted $W_H(\sigma_0,\ r)$ where the relation between r and ξ is given by (14). Note that the generated displacements are white noise for $r = 0$ and become progressively smoother as r increases. The regularity of the deformation converges to that of the generating wavelet ψ for $r \to \infty$. For moderate to large r we get a hierarchical warping: a deformation comprising displacements at several scale levels with gradually decreasing amplitudes, from important large-level deformation toward progressively smaller finer level details. The algorithm



(a) (b)

Fig. 5. The quality of the registration as a function of the warp spline degree and the knot spacing. The initial values ranges (prior to the registration) were $E = 150$ and $\varpi = 3.5$. Each point shown is an average of 30 experiments.

should work well for such deformations, which are compatible with the multiresolution strategy.

Finally, the deformation can be projected onto V_h if needed, in which case we denote it $W_H(\sigma_0, r, h)$. Typically, we use Battle–Lemarié wavelets of order four, $\sigma_0 = 5$, and $r = 0.5 \sim 1.6$, depending of what aspect of the algorithm we want to highlight. For some experiments we have also added an affine component. Examples of generated deformations are shown in Fig. 4.

E. Ideal Case

We begin our series of experiments with an ideal case. The test (f^o) and reference (f^r) images are identical except for a known transformation, no noise is present. We use a random Sobolev deformation $W_H(5, 1.6, 8)$ as defined in the previous section. We show how the criterion E and the warping index ϖ decrease with the knot spacing h. Moreover, we want to demonstrate the advantage of using cubic splines to represent the warping, as opposed to linear and quadratic ones.[5] (Linear splines are sometimes used for motion estimation [40], [41].)

Fig. 5 displays E and ϖ as a function of the degree n of the splines used to represent the deformation g and the knot spacing h. It clearly shows the benefit of using a high degree n in high-precision applications. The minimum achievable error $\varpi_{\min}$ (cf., Section II-C2) is shown by a dotted line for the cubic case and marked optimal.

F. Image Interpolation Order

As many other registration algorithm use only linear interpolation on the image, we want to show in this experiment that higher order image interpolation is advantageous from the point of view of both precision and speed.

Table I demonstrates the dependency of the registration accuracy and speed on the degree n_i of the splines used for image interpolation. As a deformation, we have chosen $W_H(5, 0.5, 64)$. The stopping criterion was identical in all cases.

[5]For splines of even degrees, there is an additional shift term of $1/2$ in (2), e.g., $\beta_2(x/h - k + 1/2)$, which minimizes the number of knots needed. Other changes are straightforward.

TABLE I
THE QUALITY OF THE REGISTRATION AS A FUNCTION OF THE IMAGE INTERPOLATION SPLINE DEGREE. THE WARPING FUNCTION WAS INTERPOLATED BY CUBIC SPLINES

h		32			16		
n_i	before	1	2	3	1	2	3
E	113.4	2.53	1.84	1.50	1.12	1.00	0.58
ϖ	2.93	0.130	0.111	0.098	0.074	0.067	0.057
time [s]		1010	766	69	1983	1500	103

In addition to providing a less accurate approximation, linear splines are penalized by not providing a meaningful second derivative estimate of the image, which forces us to use the slower conjugated gradient algorithm in place of the Marquardt–Levenberg optimizer.

Table I shows that although all results are accurate to a fraction of a pixel, the convergence speed varied a great deal. This proves that the benefit from better approximation properties of higher order splines with respect to linear, or even quadratic ones, indeed outweighs the increase in computational complexity per iteration. This is consistent with other findings in the literature [3]. By being able to better estimate the image derivatives $\nabla_{\boldsymbol{x}^c} f$, one is able to better estimate the criterion derivatives $\nabla_{\mathfrak{c}} E$ and $\nabla^2_{\mathfrak{c}} E$. This permits the optimization algorithm to acquire a more precise local model of $E(\boldsymbol{c})$, which in turn leads to more efficient optimization steps and faster convergence. The ability to precisely represent an image from a few samples is crucial at coarse levels of the image pyramid. The iterations there are relatively inexpensive compared to finer levels because we process much less data. It is therefore beneficial to get as close as possible to the optimum. In this manner we provide a good starting estimate for the next finer level, where each iteration costs at least four times more, and thus save the overall effort.

G. Noisy case

Fig. 6 demonstrates the dependence of the registration accuracy on the signal-to-noise ratio. For this series of experiments the test images were obtained from a known transformation of a reference image with a various level of white Gaussian

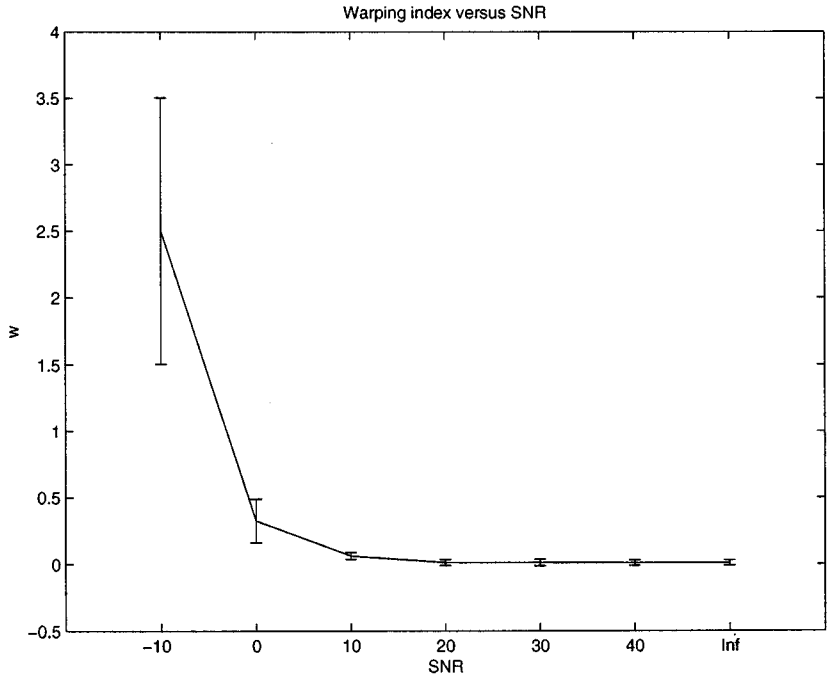
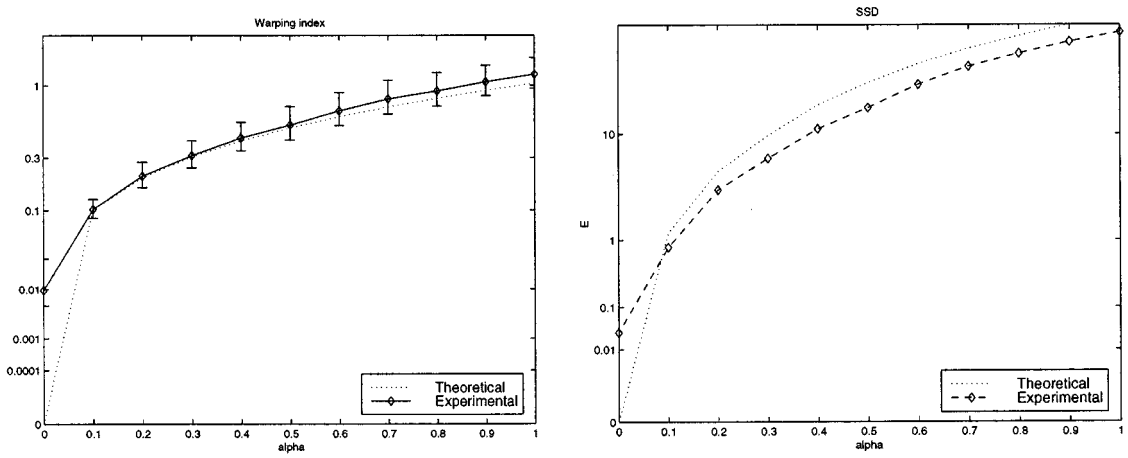


Fig. 6. The quality of the registration as a function of the SNR in dB. $h = 32$, $\varpi_{\text{before}} = 3.34$. The error bars mark one standard deviation.



(a) (b)

Fig. 7. (a) Warping index and (b) SSD for an out-of-space deformation. The mean initial values (prior to the registration) ranged between $E = 146$, $\varpi = 2.8$ for $\alpha = 0$ and $E = 181$, $\varpi = 3.47$ for $\alpha = 1$. Error bars show one standard deviation.

noise added. Here, an in-space deformation $W_H(5, 1.16, 32)$ was used. We observe that the degradation of the algorithm's performance is graceful for SNR > 10 dB.

H. Out-of-Space Deformation

This experiment illustrates the behavior of the algorithm as the deformation progressively moves out of the warp space. Specifically, we have used $W_H(5, 0.5, 32)+\alpha W'_H(5, 0.5)$ with $0 \leq \alpha \leq 1$ where the first term corresponds to the in-space part of the deformation and W'_H is the displacement associated with W_H, i.e., $W_H - g_I$. Fig. 7 compares the attained and theoretically attainable warping index and the corresponding SSD.

It can be seen that, the experimentally recovered deformation is almost as close to the true one as theoretically possible, with the exception of $\alpha = 0$, i.e., a deformation lying exactly in the warp space. On the other hand, we obtain mostly a smaller SSD than what corresponds to the projection. This demonstrates that the relationship between the criterion we optimize (E) and the true error we make (ϖ) is far from straightforward. The minimum of the difference (E) does not correspond exactly to the true solution as measured by ϖ. We will see this sort of behavior in Section III-J, as well.

I. Multiresolution Strategy

In Table II, we compare three multiresolution strategies: the strategy actually used; a strategy without multiresolution in the image size; and a strategy without multiresolution in the warping grid size. The results show that both multiresolutions

TABLE II
DIFFERENT MULTIRESOLUTION STRATEGIES. IMAGE SIZE IS DENOTED BY N, GRID SPACING BY h. FROM TOP TO BOTTOM: JOINT, IMAGE SIZE ONLY AND WARP GRID SIZE ONLY MULTIRESOLUTIONS

Warping index before registration $\varpi = 9.3$.

h	128	64	64	32	32	16	
N	32	32	64	64	128	128	
ϖ	0.43	0.40	0.15	0.15	0.0038	0.0029	total
time	2.2	2.0	29.0	1.2	20.0	23.2	78

h	16	16	16	
N	32	64	128	
ϖ	1.40	0.86	0.69	total
time	14.0	27.1	129.8	171

h	128	64	32	16	
N	128	128	128	128	
ϖ	0.145	0.0043	0.0038	0.0031	total
time	164.9	101.6	24.6	34.4	326

improve accuracy as well as speed of the algorithm because they reduce the amount of data to be treated and provide a smoothed version of the problem.

Multiresolution in the warp grid size significantly improves accuracy because it avoids that the algorithm is trapped in a local minimum. The multiresolution in the image size reduces the amount of data to be treated, and consequently the execution time, mainly for coarse control grids.

The joint strategy combines the advantages of both multiresolutions and yields the best results.

J. Artificial EPI images

By filtering and performing histogram modification of the anatomical MRI images, we obtain images that are visually equivalent to the corresponding EPI images (compare Figs. 3 and 8) and we use them as test images. We warp the anatomical images with a known deformation $W_H(5, 0.9)$ and use them as reference images.

The reference/test image pairs are registered automatically as well as manually using the landmark method by three people, including one experienced practitioner. The results are presented in Table III. For the manual case, the best results obtained (in the sense of ϖ among all the attempts of all the participants) are shown. The column marked ideal shows the minimum attainable warping index $\varpi_{\min}$, given the fact that the target deformation does not belong to the space searched by the algorithm. The results demonstrate that the automatic method is vastly superior to the manual one.

K. Real EPI Images

Fig. 9 shows a typical pair of corresponding anatomical and EPI images with superimposed contours of the anatomical image before and after manual and automatic registration. It illustrates that the automatic procedure leads to subjectively comparable or better results than the manual one.

IV. CONCLUSION

We have suggested a new approach for undoing nonlinear unidirectional deformations in EPI images. We proceed by registering them with corresponding geometrically correct anatomical MRI images. We have developed a fully automatic image registration algorithm specialized for this task. Our technique increases the geometrical precision of EPI images and thus improves the quality of information obtainable from these images. This will allow the use of EPI images in many clinical and diagnostic applications where they could not have been used previously, as well as an increase of their usefulness in existing applications.

As an additional benefit, our method can be extended to compensate for other causes of geometrical distortion of EPI images besides imperfect magnetic field, such as heart beat and respiration. (In this case, we would look for bidirectional warping.)

The novelty of our registration algorithm stems from a high-order spline model for the warping. It has good approximation properties and lends itself well to a multiresolution approach, while permitting an efficient implementation. We have also taken advantage of a spline model for the image being warped, leading to a second dimension of the multiresolution strategy, and yielding additional computational savings. Finally, we have replaced the customary regularization criterion by a scale parameter of the search space.

We have also presented many experiments to demonstrate the performance of our algorithm.[6] We plan to carry out more clinical experiences to prove conclusively the potential of our method in a real-world setting.

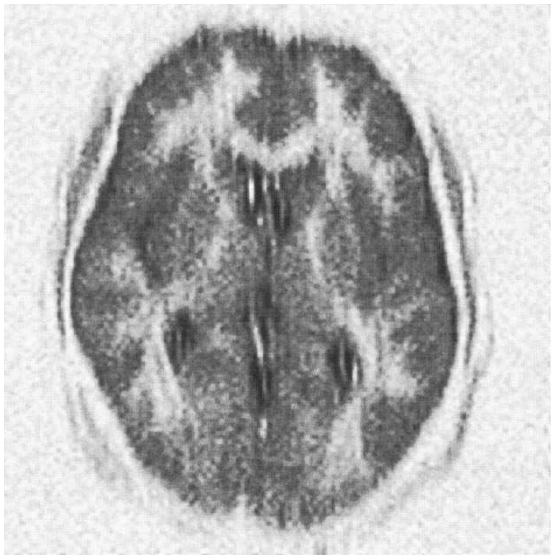


Fig. 8. Artificial EPI image.

TABLE III
MANUAL VERSUS AUTOMATIC REGISTRATION. ϖ IS THE WARPING INDEX, E IS THE MEAN-SQUARE DIFFERENCE

	before	manual	automatic	ideal	true
ϖ	2.60	1.01	0.44	0.19	0
E	246	178	161	163	162
time		~ 5 min	~ 30 s		

[6]An online demonstration of our algorithm is available on our WEB page http://bigwww.epfl.ch/.

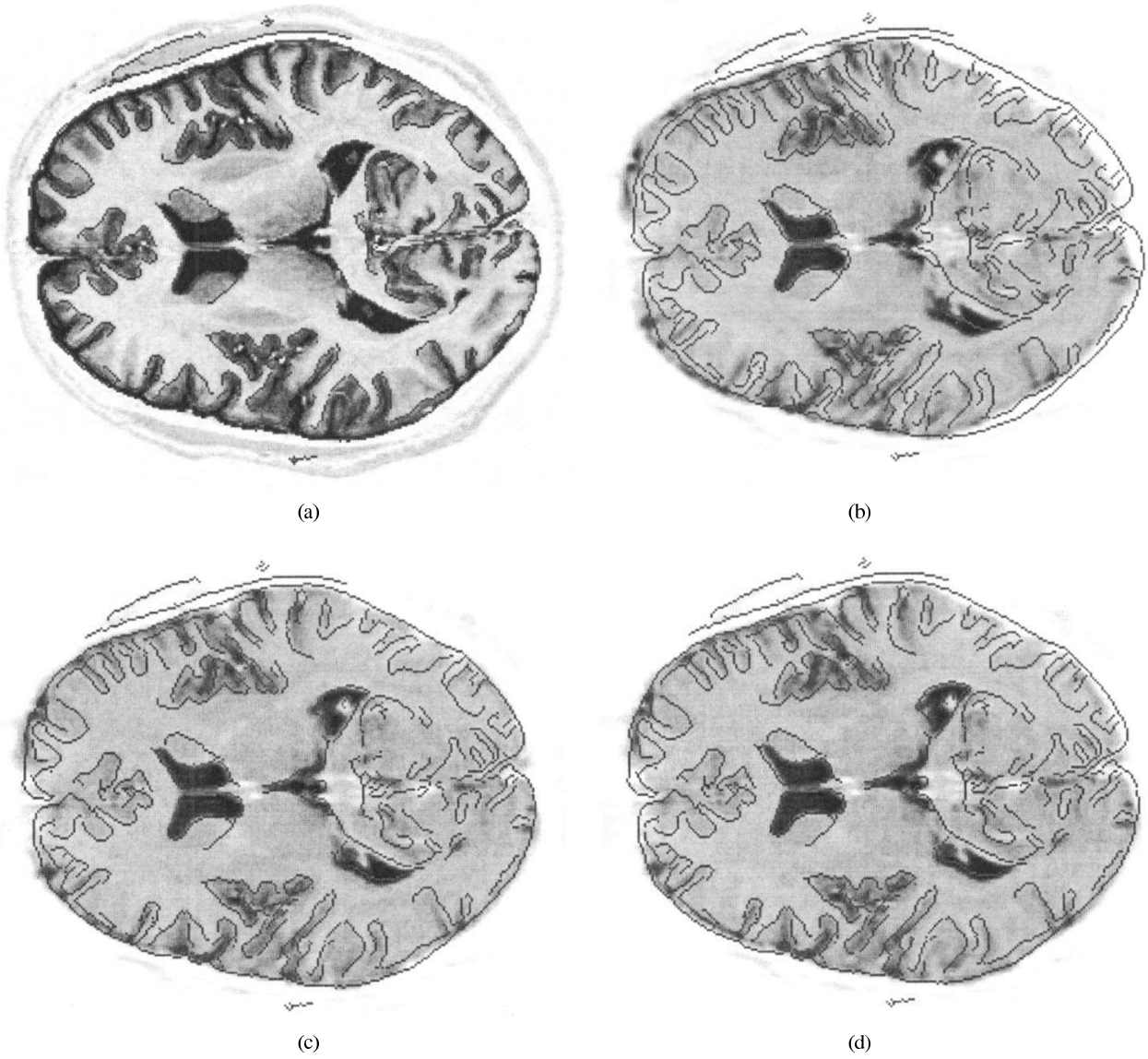


Fig. 9. (a) Anatomical and (b) EPI images before registration, with superimposed contours from the anatomical images. (c) EPI image after automatic and (d) manual registration.

APPENDIX A
OPTIMALITY OF THE CUBIC SPLINE MODEL

The landmark fitting from Section II-B is mathematically equivalent to the smoothing spline problem frequently encountered in statistics [44]. We present here an informal derivation of the variational property of cubic splines. We use an original Fourier-based technique which is instructive and concise.

Consider the following approximation problem. Given a set $\{(x_i, y_i) \in \mathbb{R}^2\}$ and a function $\mathcal{L}$: $\mathbb{R}^2 \rightarrow \mathbb{R}$, find a function g: $\mathbb{R} \rightarrow \mathbb{R}$ minimizing the functional criterion

$$J'(g) = J(g) + \sum_i \mathcal{L}(g(x_i), y_i) \qquad \text{with}$$

$$J(g) = \int |g''(x)|^2 \, dx. \tag{16}$$

It is not difficult to show that the solution of this problem belongs to the same class as a solution of the interpolation problem, that would consist of finding g minimizing $J(g)$ under the constraints $g(x_i) = z_i$, equivalent to landmark interpolation from Section II-B. We will therefore concentrate on the interpolation case here.

Using the Lagrange multiplier method, we construct an augmented criterion

$$J_\lambda = \int |g''(x)|^2 \, dx - 24 \sum_i \lambda_i (g(x_i) - z_i). \tag{17}$$

We then express J_λ in terms of $\hat{g}(\omega)$, the Fourier transform of g

$$J_\lambda = \frac{1}{2\pi} \int \omega^4 |\hat{g}(\omega)|^2 \, d\omega - 24 \cdot \sum_i \lambda_i \left(\frac{1}{2\pi} \int \overline{\hat{g}(\omega)} \, e^{-j\omega x_i} \, d\omega - z_i \right) \tag{18}$$

where we have used $\hat{g}(\omega) = \overline{\hat{g}(-\omega)}$. To find an optimal $\hat{g}(\omega)$, we differentiate J_λ with respect to $\overline{\hat{g}(\omega)}$ and we impose the first order change to be zero

$$\begin{aligned}\Delta J_\lambda &= 2\,\mathrm{Re}\left(\frac{1}{2\pi}\int \overline{\Delta\hat{g}}\left(\hat{g}\omega^4 - 12\sum_i \lambda_i\, e^{-j\omega x_i}\right) d\omega\right)\\ &= 0, \qquad \text{for all } \overline{\Delta\hat{g}(\omega)}. \end{aligned} \tag{19}$$

This gives

$$\hat{g} = \sum_i \lambda_i \frac{12\, e^{-j\omega x_i}}{\omega^4}. \tag{20}$$

By interpreting $e^{-j\omega x_i}$ as a phase shift, we get

$$g(x) = \sum_i \lambda_i\, \varrho(x - x_i) \tag{21}$$

where ϱ is the inverse Fourier transform of $12/\omega^4$, which gives a general form of all solutions to our interpolation or approximation problems as

$$g(x) = \sum_i \lambda_i |x - x_i|^3. \tag{22}$$

The basis function $|x|^3$ is twice continuously differentiable and piecewise cubic. Hence, g is a cubic spline. In order to be complete this solution needs to be augmented by a linear term generating the null space of J. For a rigorous treatment and generalization to multiple dimensions, we refer to [38].

APPENDIX B
B-SPLINES

A B-spline β_r of degree r is recursively defined as

$$\beta_r = \beta_{r-1} * \beta_0, \qquad \text{for } r > 0$$

$$\beta_0(x) = \begin{cases} 1, & \text{if } x \in \left(-\frac{1}{2}, \frac{1}{2}\right) \\ 0, & \text{otherwise.} \end{cases}$$

B-splines, as defined above, have a compact support $(-r/2 - 1/2,\ r/2 + 1/2)$, are symmetric, and are $(r-1)$-times continuously differentiable. Specifically

$$\beta_3(x) = \begin{cases} 2/3 - (1 - |x|/2)x^2, & \text{if } 0 < |x| \le 1 \\ (2 - |x|)^3/6, & \text{if } 1 < |x| < 2 \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

APPENDIX C
EXPLICIT DERIVATIVES

Given E by (4), (6), and (8), let us express the components of $\nabla_{\mathfrak{L}} E$ and $\nabla^2_{\mathfrak{L}} E$.

$$\frac{\partial E}{\partial c_{\boldsymbol{k}}} = 2\sum_{\boldsymbol{i}\in I} e_{\boldsymbol{i}} \frac{\partial e_{\boldsymbol{i}}}{\partial c_{\boldsymbol{k}}} \qquad \text{where } e_i = {}^c f(\boldsymbol{g}(\boldsymbol{i})) - f^r(\boldsymbol{i}) \tag{24}$$

$$\begin{aligned}\frac{\partial e_{\boldsymbol{i}}}{\partial c_{\boldsymbol{k}}} &= \left.\frac{\partial^c f}{\partial x}\right|_{\boldsymbol{x}=\boldsymbol{g}(\boldsymbol{i})} \cdot \left.\frac{\partial g}{\partial c_{\boldsymbol{k}}}\right|_{\boldsymbol{x}=\boldsymbol{i}} \\ &= \boldsymbol{\beta}_3(\boldsymbol{i}/h - \boldsymbol{k}) \sum_q b_{qj}\beta_3'(g(\boldsymbol{i}) - q)\end{aligned} \tag{25}$$

where $\boldsymbol{i} = (i, j)$ and the summation across q only needs to be performed within the support of β_3' of length four. The second derivatives are

$$\frac{\partial^2 E}{\partial c_{\boldsymbol{k}} \partial c_{\boldsymbol{l}}} = 2\sum_{\boldsymbol{i}\in I}\left(\frac{\partial e_{\boldsymbol{i}}}{\partial c_{\boldsymbol{k}}}\frac{\partial e_{\boldsymbol{i}}}{\partial c_{\boldsymbol{l}}} + e_{\boldsymbol{i}}\frac{\partial^2 e_{\boldsymbol{i}}}{\partial c_{\boldsymbol{k}}\partial c_{\boldsymbol{l}}}\right) \tag{26}$$

$$\frac{\partial^2 e_{\boldsymbol{i}}}{\partial c_{\boldsymbol{k}} \partial c_{\boldsymbol{l}}} = \left.\frac{\partial^2 {}^c f}{\partial^2 x}\right|_{\boldsymbol{x}=\boldsymbol{g}(\boldsymbol{i})} \frac{\partial g}{\partial c_{\boldsymbol{k}}}\frac{\partial g}{\partial c_{\boldsymbol{l}}} + \frac{\partial {}^c f}{\partial x}\underbrace{\frac{\partial^2 g}{\partial c_{\boldsymbol{k}}\partial c_{\boldsymbol{l}}}}_{0} \tag{27}$$

$$\frac{\partial^2 e_{\boldsymbol{i}}}{\partial c_{\boldsymbol{k}} \partial c_{\boldsymbol{l}}} = \boldsymbol{\beta}_3(\boldsymbol{i}/h - \boldsymbol{k})\boldsymbol{\beta}_3(\boldsymbol{i}/h - \boldsymbol{l}) \sum_q b_{qj}\beta_3''(g(\boldsymbol{i}) - q) \tag{28}$$

which is easy to calculate because many terms can be reused from the calculation of the first derivatives.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for bringing to their attention a recent conference report [42], which describes a similar approach to theirs and that was published after this manuscript had been submitted, at about the same time as their conference report [43].

REFERENCES

[1] Z.-H. Cho, J. P. Jones, and M. Singh, *Foundations of Medical Imaging*. New York, NY: Wiley, 1993.
[2] X. Wan, G. T. Gullberg, D. L. Parker, and G. L. Zeng, "Reduction of geometric and intensity distortions in echo-planar imaging using a multireference scan," *Magnetic Resonance Med.*, vol. 37, no. 6, pp. 932–942, 1997.
[3] P. Thévenaz, U. E. Ruttimann, and M. Unser, "A pyramid approach to subpixel registration based on intensity," *IEEE Trans. Image Processing*, vol. 7, pp. 1–15, Jan. 1998.
[4] H. Chang and J. M. Fitzpatrick, "A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities," *IEEE Trans. Med. Imag.*, vol. 11, pp. 319–329, Sept., 1992.
[5] B. B. Biswal and J. S. Hyde, "Contour-based registration technique to differentiate between task-activated and head motion-induced signal variations in fMRI," *Magnetic Resonance Med.*, vol. 38, no. 3, pp. 470–476, 1997.
[6] A. L. Alexander, J. S. Tsuruda, and D. L. Parker, "Elimination of Eddy current artifacts in diffusion-weighted echo-planar images: The use of bipolar gradients," *Magnetic Resonance Med.*, vol. 38, no. 6, pp. 1016–1021, 1997.
[7] P. Jezzard and R. S. Balaban, "Correction for geometric distortion in echo planar images from B_0 field variations," *Magnetic Resonance Med.*, pp. 65–73, 1995.
[8] 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 in Biomedical Computing*, K. H. Höhne and R. Kikinis, Eds. Berlin, Germany: Springer-Verlag, 1996, pp. 297–306.
[9] B. Horn and B. Schunck, "Determining optical flow," *Artificial Intell.*, vol. 17, pp. 185–203, 1981.
[10] G. Christensen, "Deformable shape models for anatomy," Ph.D. dissertation, Washington Univ., Saint Louis, MO, 1994.
[11] G. Christensen, S. Joshi, and M. Miller, "Volumetric transformation of brain anatomy," *IEEE Trans. Med. Imag.*, vol. 16, pp. 864–877, Dec. 1997.
[12] M. Bro-Nielsen and C. Gramkow, "Fast fluid registration of medical images," in *Visualization in Biomedical Computing*, K. H. Höhne and R. Kikinis, Eds. Berlin, Germany: Springer-Verlag, 1996, pp. 267–276.
[13] R. Bajcsy and S. Kovačič, "Multiresolution elastic matching," *Comp. Vision Graph. Image Processing*, vol. 46, pp. 1–21, 1989.
[14] D. V. Iosifescu, M. E. Shenton, S. K. Warfield, R. Kikinis, J. Dengler, F. A. Jolesz, and R. W. McCarley, "An automated registration algorithm for measuring MRI subcortical brain structures," *Neuroimage*, pp. 14–25, 1997.

[15] Y. Tai, K. Lin, C. Hoh, S. Huang, and E. Hoffman, "Utilization of 3D elastic transformations in the registration of chest X-ray CT and whole body PET," *IEEE Trans. Nucl. Sci.*, vol. 44, pp. 1606–1612, Aug. 1997.
[16] T. Schormann, S. Henn, and K. Zilles, "A new approach to fast elastic alignment with applications to human brains," *Visualization in Biomedical Computing*, 1996.
[17] J. Bergen, P. Anandan, K. Hanna, and R. Hingorani, "Hierarchical model-based motion estimation," in *Proc. Second European Conf. Computer Vision ECCV'92*, 1992, pp. 237–252.
[18] N. Sicotte, R. Woods, and J. Mazziotta, "Automated image registration using a 105 parameter nonlinear model," *Neuroimage*, June 1996.
[19] R. Woods, S. Grafton, N. Sicotte, and J. Mazziotta, "Automated image registration—II: Intersubject validation of linear and nonlinear models," *J. Comp. Assisted Tomogr.*, vol. 22, no. 1, pp. 153–165, 1998.
[20] S. Kiebel, J. Ashburner, J. Poline, and K. Friston, "MRI and PET coregistration—A cross validation of statistical parametric mapping and automated image registration," *Neuroimage*, no. 5, 1997.
[21] A. W. Toga, Ed., *Brain Warping*. New York: Academic , 1999.
[22] P. Thompson and A. W. Toga, "A surface-based technique for warping 3-dimensional images of the brain," *IEEE Trans. Med. Imag.*, vol. 15, pp. 1–16, Aug. 1996.
[23] C. Davatzikos, "Spatial normalization of 3D brain images using deformable models," *J. Comp. Assisted Tomogr.*, vol. 4, no. 20, pp. 656–665, 1996.
[24] J. Declerck, G. Subsol, J. P. Thirion, and N. Ayache, "Automatic retreival of anatomical structures in 3d medical images," in *Proc. CVRMed'95, Lecture Notes in Computer Science*, N. Ayache, Ed., Nice, France, Apr. 1995, pp. 153–162.
[25] F. Bookstein, *Morphometric Tools for Landmark Data: Geometry and Biology*. New York: Cambridge Univ. Press, 1997.
[26] B. Kim, J. Boes, K. Frey, and C. Meyer, "Mutual information for automated multimodal image warping," in *Visualization in Biomedical Computing*, K. H. Höhne and R. Kikinis, Eds. Berlin, Germany: Springer-Verlag, 1996.
[27] H. Lester and S. Arridge, "Summarizing fluid registration by thin-plate spline warps with many landmarks," *Med. Image Understanding Anal.*, July 1997.
[28] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing—Part I: Theory," *IEEE Trans. Signal Processing*, vol. 41, pp. 821–832, Feb. 1993.
[29] ——, "B-spline signal processing—Part II: Efficient design and applications," *IEEE Trans. Signal Processing*, vol. 41, pp. 834–848, Feb. 1993.
[30] J. Ahlberg, E. Nilson, and J. Walsh, *The Theory of Splines and Their Applications*. New York: Academic, 1967.
[31] I. Schoenberg, "Spline functions and the problem of graduation," *Proc. Nat. Acad. Sci.*, vol. 52, pp. 947–950, 1964.
[32] M. Unser and T. Blu, "Fractional splines and wavelets," SIAM Rev., 1999, to be published.
[33] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques—Part I: Interpolators and projectors," *IEEE Trans. Signal Processing*, vol. 10, pp. 2783–2795, Oct. 1999.
[34] 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, pp. 277–285, Mar. 1991.
[35] G. Subsol, "Construction automatique d'atlas anatomiques morphométriques à partir d'images médicales tridimensionnelles," Ph.D. dissertation, Ecole Centrale, Paris, France, 1995.
[36] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. New York: Cambridge Univ. Press, 1992.
[37] F. Müller, P. Brigger, K. Illgner, and M. Unser, "Multiresolution approximation using shifted splines," *IEEE Trans. Signal Processing*, vol. 46, pp. 2555–2558, Sept. 1998.
[38] J. Duchon, "Splines minimizing rotation-invariant semi-norms in Sobolev spaces," in *Constructive Theory of Functions of Several Variables*, W. Schempp and K. Zeller, Eds. Berlin, Germany: Springer-Verlag, 1977, pp. 85–100.
[39] S. Mallat, *A Wavelet Tour of Signal Processing*. New York: Academic, 1998.
[40] P. Moulin, R. Krishnamurthy, and J. Woods, "Multiscale modeling and estimation of motion fields for video coding," *IEEE Trans. Image Processing*, vol. 6, pp. 1606–1620, Dec. 1997.
[41] R. Szeliski and J. Coughlan, "Spline-based image registration," *Int. J. Comp. Vision*, vol. 22, pp. 199–218, 1997.
[42] C. Studholme, R. T. Constable, and J. S. Duncan, "Incorporating an image distortion model in nonrigid alignment of EPI with conventional MRI," in *Proc. Int. Conf. Information Processing Medical Imaging, Lecture Notes Computer Science*, Visegrád, Hungary, June 1999, pp. 454–459.
[43] J. Kybic, P. Thévenaz, and M. Unser, "Multiresolution spline warping for EPI registration," *Proc. SPIE*, vol. 3813, pp. 571–579, July 1999.
[44] G. Wahba, *Spline Models for Observational Data*. Philadelphia, PA: SIAM, 1990.