Feature-Based Registration of Medical Images: Estimation and Validation of the Pose Accuracy

Xavier Pennec[1,2], Charles R.G. Guttmann[3], and Jean-Philippe Thirion[4,1]

[1] INRIA Sophia Antipolis, EPIDAURE project,France
Xavier.Pennec@sophia.inria.fr,
http://www.inria.fr/epidaure/personnel/pennec/pennec.html
[2] MIT, A.I.Lab, Cambridge, USA
[3] Brigham and Women's Hospital, Harvard Medical School, Boston, USA
[4] Focus Imaging SA, France

Abstract. We provide in this article a generic framework for pose estimation from geometric features. We propose more particularly two algorithms: a gradient descent on the Riemannian least squares distance and on the Mahalanobis distance. For each method, we provide a way to compute the uncertainty of the resulting transformation. The analysis and comparison of the algorithms show their advantages and drawbacks and point out the very good prediction on the transformation accuracy. An application in medical image analysis validates the uncertainty estimation on real data and demonstrates that, using adapted and rigorous tools, we can detect very small modifications in medical images. We believe that these algorithms could be easily embedded in many applications and provide a thorough basis for computing many image statistics.

1 Introduction

Registration is a fundamental task in medical imaging to compare images taken at different times for diagnosis or therapy. In the case of images of the same patient, one often assume that the motion between the images is rigid, and registration consists in estimating the six parameters of the 3D rotation and translation. When registration is based on features extracted from the images, the problem can be separated into two steps: (1) finding the correspondences between features (matches) and (2) computing the geometric transformation that maps one set of features to the other. In this article, we do not discuss matching methods per se, but rather the estimation of the geometric transformation from matched features. The quantification of the registration quality is also an important problem as most measurements are done after registration. For instance, registration errors can have a strong influence on the quantification of the lesion evolution [3]. Knowing the uncertainty of the transformation might even be vital in image guided surgery when it comes to operate close to important anatomical structures.

Most existing methods for computing 3D rigid motion deal with sets of matched points and minimize the sum of square distances after registration.

This is called the *orthogonal Procrustes problem* in statistics, the *absolute orientation problem* in photogrammetry and the *pose estimation problem* in computer vision. Several closed form solutions have been developed, using unit quaternions [4], singular value decomposition (SVD) [9, 11], Polar decomposition [5] or dual quaternions [12]. However, models of the real world often require more complex features like lines [2], planes, oriented points or frames. Traditional methods rely on the vector space structure of points and generalizing them directly to other types of features leads to paradoxes. For instance, depending on the representation used, the standard expectation could take an arbitrary value [7]. Moreover, if uncertainty handling is a central topic in several works, like [1], there are much fewer studies dealing with the accuracy of the estimated transformation.

We first review some notions of Riemannian geometry in order to introduce proper tools on geometric features. Then, we develop a pose estimation criterion based on the Riemannian least-squares and another based on the intrinsic Mahalanobis distance, and provide a way to compute an estimation of the result accuracy (generalizing the approache of [8] to any kind of features). In the last section, we investigate a practical case in Medical Image Analysis: the registration of MRI images of the head, where the availability of each image in two different echoes allows us to test for the uncertainty prediction.

2 Geometric Features

Geodesics Geometric features like lines, planes, oriented points, frames, etc. generally belong to a manifold and not to a vector space. In the geometric framework, one specifies the structure of a manifold $\mathcal{M}$ by a *Riemannian metric*. This is a continuous collection of dot products on the tangent space at each point x of the manifold. Thus, if we consider a curve on the manifold, we can compute at each point its instantaneous speed. The length of the curve is obtained as usual by integrating it along the curve. The distance between two points of a connected Riemannian manifold is the minimum length among the curves joining these points. The curves realizing this minimum for any two points of the manifold are called geodesics.

Exponential charts Let us develop the manifold in the tangent space at point x along the geodesics (think of rolling a sphere along its tangent plane). The geodesics going through that point are transformed into straight lines and the distance along these geodesics are conserved. This generates a chart called *the exponential chart*. It covers all the manifold except a set of null measure called the cut locus. Let $\overrightarrow{\mathrm{xy}}$ be the representation of y in this chart. Then its distance to x is $\mathrm{dist}(\mathrm{x}, \mathrm{y}) = \|\overrightarrow{\mathrm{xy}}\|$. This means that the exponential chart is a linear representation of the manifold with respect to the development point.

Invariant distance Since we are working with a transformation group that models the possible image viewpoints, it is natural to choose an invariant Riemannian metric on the manifold. This way, all the measurements based on distance are independent of the image reference frame. Denoting by $\mathrm{f} \star \mathrm{x}$ the action of transformation f on feature x, the distance is invariant if $\mathrm{dist}(\mathrm{x}, \mathrm{y}) = \mathrm{dist}(\mathrm{f} \star \mathrm{x}, \mathrm{f} \star \mathrm{y})$. Existence conditions for such a metric are detailes in [7].

Principal chart Let o be a point of the manifold that we call *the origin* and f_x be a "placement function" (a transformation such that $\mathrm{f}_\mathrm{x} \star o = \mathrm{x}$). We call *principal chart* the exponential chart at the origin and we denote by $\vec{\mathrm{x}}$ the representation of x in this chart. In this chart, the distance becomes: $\mathrm{dist}(\mathrm{x}, \mathrm{y}) = \mathrm{dist}\left(\mathrm{f}_\mathrm{x}^{(-1)} \star \mathrm{y}, o\right) = \|\mathrm{f}_{\vec{\mathrm{x}}}^{(-1)} \star \vec{\mathrm{y}}\|$. In fact, we can express all operations of interest for us from the following "*atomic operations*" and their Jacobians in the principal chart: **the action** $[\vec{\mathrm{f}} \star \vec{\mathrm{x}}]$ of a transformation and **the placement function** $[\vec{\mathrm{f}_{\vec{\mathrm{x}}}}]$.

The transformation group Since the group acts on itself, we just have to replace the action by **the composition** $[\vec{\mathrm{f}} \circ \vec{\mathrm{g}}]$ and add **the inversion** $[\vec{\mathrm{f}}^{(-1)}]$ to the atomic operations. The placement function disappears (it is the identity). An important property of the invariant metric is that it relates the exponential chart at any point f with the principal chart. Using the non-orthogonal coordinate system induced by the principal chart, we have: $\overrightarrow{\mathrm{fg}} = J_L(\vec{\mathrm{f}})(\vec{\mathrm{f}}^{(-1)} \circ \vec{\mathrm{g}})$, where $J(\vec{\mathrm{f}}) = \left.\frac{\partial(\vec{\mathrm{f}} \circ \vec{\mathrm{e}})}{\partial \vec{\mathrm{e}}}\right|_{\vec{e} = Id}$ is the Jacobian of the left translation of the identity in the principal chart. From a practical point of view, this means that we can "translate" local calculations on points to the principal chart of our transformation group by replacing $g - f$ with $\overrightarrow{\mathrm{fg}}$ and $f + \overrightarrow{\delta f}$ with $\exp_{\vec{\mathrm{f}}}(\overrightarrow{\delta \mathrm{f}}) = \vec{\mathrm{f}} \circ (J_L(\vec{\mathrm{f}})^{(-1)} \vec{\delta \mathrm{f}})$.

Example of features We have implemented this framework for 3D rigid transformations acting on frames, semi-oriented frames and points. Frames are composed of a point and an orthonormal trihedron and are equivalent to rigid transformations. The principal chart is made of the rotation vector representing the trihedron or the rotation and the translation of the point position vector. Semi-oriented frames model the differential properties of a point on a surface. In particular, they model the "extremal points" we will extract on medical images in Sec. 4. They are composed of a point and a trihedron (t_1, t_2, n) where $(t_1, t_2) \equiv (-t_1, -t_2)$ are the principal *directions* and n the normal of the surface.

3 Feature-Based Pose Estimation

3.1 Riemannian Least-Squares

Let $\{\mathrm{x_i}\}$ and $\{\mathrm{y_i}\}$ be two sets of matched features. The Least squares criterion is easily written using the invariant Riemannian distance:

$$C(\mathrm{f}) = \frac{1}{2} \sum_i \mathrm{dist}(\mathrm{y_i}, \mathrm{f} \star \mathrm{x_i})^2$$

Now, thanks to the good properties of the principal chart, it turns out that this criterion can be expressed as a classical sum of squares of vector norms. Let $\vec{\mathrm{z}}_i = \vec{\mathrm{f}}_{\vec{\mathrm{y}}_i}^{(-1)} \circ (\vec{\mathrm{f}} \star \vec{\mathrm{x}}_i)$ be the *error vector* in the principal chart. The criterion becomes $2C(\mathrm{f}) = \sum_i \mathrm{dist}\left(\mathrm{f}_{\mathrm{y_i}}^{(-1)} \circ (\mathrm{f} \star \mathrm{x_i})\,,\, \mathrm{o}\right)^2 = \sum_i \|\vec{\mathrm{z}}_i\|^2$.

From the atomic operations, and using the composition rule for differentials, we can compute the error vector $\vec{z}_i$ and its Jacobians $\frac{\partial \vec{z}_i}{\partial \vec{x}_i}$, $\frac{\partial \vec{z}_i}{\partial \vec{y}_i}$, and $\frac{\partial \vec{z}_i}{\partial \vec{f}}$. The first derivative of the criterion $C(\mathrm{f})$ is: $\Phi = \sum_i \frac{\partial \vec{z}_i}{\partial \vec{f}}^{\mathrm{T}} \vec{z}_i$. Neglecting the term in $\ddot{z}z$ with respect to the terms in $\dot{z}^2$ in the second derivatives, we obtain: $H = \frac{\partial \Phi}{\partial \vec{f}} \simeq \sum_i \frac{\partial \vec{z}_i}{\partial \vec{f}}^{\mathrm{T}} \frac{\partial \vec{z}_i}{\partial \vec{f}}$ and $\frac{\partial \Phi}{\partial \vec{z}_i} \simeq \frac{\partial \vec{z}_i}{\partial \vec{f}}^{\mathrm{T}}$.

A gradient descent algorithm Assume that f is a vector. The 2nd order Taylor expansion of the criterion is $C(f + \delta f) \simeq C(\vec{f}) + \Phi^{\mathrm{T}} \delta f + \frac{1}{2} \delta f^{\mathrm{T}} H \delta f$. The minimum of this approximation is obtained for $\delta f = -H^{(-1)} \Phi$. Now, since $\vec{f}$ is the expression of a transformation in the principal chart, we just have to replace $f + \delta f$ by $\exp_{\mathrm{f}}(\overrightarrow{\delta f}) = \vec{f} \circ (J_L(\vec{f})^{(-1)} \overrightarrow{\delta f})$, and iterate the process:

$$\vec{f}_{t+1} = \vec{f}_t \circ \left(-J_L(\vec{f}_t)^{(-1)} H_t^{(-1)} \Phi_t \right) \tag{1}$$

As an initial estimate, we can choose the identity if nothing else is given. The process is stopped when the norm $\|\overrightarrow{\delta f}_t\|$ of the adjustment transformation becomes too small (we use $\varepsilon = 10^{-10}$) or when the number of iterations becomes too high (practically, it converges in about 10 iterations).

Estimation of the uncertainty at the minimum Let $\hat{\chi}$ be the vector of observed data and $\hat{f}$ the corresponding state vector. The minimum $f(\chi)$ of the criterion for a data vector χ is characterized by $\Phi(f(\chi), \chi) = 0$. A Taylor expansion around the actual values $(\hat{\chi}, \hat{f})$ gives a modification of the state $\delta f = -\hat{H}^{(-1)} \hat{J}_\Phi \delta \chi$ for a modification of the data vector $\delta \chi$, where $\hat{J}_\Phi = \frac{\partial^2 \Phi}{\partial \chi \partial f}$. Thus, the covariance of $\hat{f}$ is $\Sigma_{\hat{f}\hat{f}} = \mathbf{E}(\delta f \delta f^{\mathrm{T}}) = \hat{H}^{(-1)} \hat{J}_\Phi \Sigma_{\hat{\chi}\hat{\chi}} \hat{J}_\Phi^{\mathrm{T}} \hat{H}^{(-1)}$. Assuming that all our measurements are independent, we can simplify $\hat{J}_\Phi \Sigma_{\hat{\chi}\hat{\chi}} \hat{J}_\Phi^{\mathrm{T}}$ to obtain:

$$\Sigma_{\hat{\mathbf{f}}\hat{\mathbf{f}}} = \hat{H}^{(-1)} \left(\sum_i \frac{\partial \hat{\Phi}}{\partial \vec{z}_i} \Sigma_{\mathbf{z_i z_i}} \frac{\partial \hat{\Phi}}{\partial \vec{z}_i}^{\mathrm{T}} \right) \hat{H}^{(-1)} \tag{2}$$

In our case, the data and the state are not vectors, but features and transformations in a Riemannian manifold. In fact, we can do the same derivation by replacing δf with $\overrightarrow{\delta f} = J(\hat{\vec{f}})(\hat{\vec{f}}^{(-1)} \circ \overrightarrow{\delta f})$ and $\delta \chi$ with a somehow similar expression. It turns out that the definition of the covariance is changed accordingly and that finally nothing is changed in equation (2).

3.2 Mahalanobis Distance Minimization

To allow different and non isotropic covariance matrices for different measures, we can minimize the sum of squared Mahalanobis distances after registration. It turns out that this Mahalanobis distance can be expressed with exactly the same error vector as before:

$$C(\mathrm{f}) = \frac{1}{2} \sum_i \mu^2(\mathbf{y_i}, \mathrm{f} \star \mathbf{x_i}) = \frac{1}{2} \sum_i \vec{z}_i^{\mathrm{T}} \Sigma_{\mathbf{z_i z_i}}^{(-1)} \vec{z}_i$$

Thus, the algorithm is the same as for least squares, but the derivatives of the criterion are different: $\Phi = \left(\frac{\partial C}{\partial \vec{f}}\right)^{\mathrm{T}} = \sum_i \frac{\partial \vec{z}_i}{\partial \vec{f}}^{\mathrm{T}} \Sigma^{(-1)}_{\mathbf{z_i z_i}} \vec{z}_i$, $H \simeq \sum_i \frac{\partial \vec{z}_i}{\partial \vec{f}}^{\mathrm{T}} \Sigma^{(-1)}_{\mathbf{z_i z_i}} \frac{\partial \vec{z}_i}{\partial \vec{f}}$ and $\frac{\partial \Phi}{\partial \vec{z}_i} \simeq \frac{\partial \vec{z}_i}{\partial \vec{f}}^{\mathrm{T}} \Sigma^{(-1)}_{\mathbf{z_i z_i}}$. Now, with these new derivatives, the Taylor expansion for the criterion is the same, and the evolution for the gradient descent is still given by equation (1). We can use the same starting value and stopping criterion as before. Practically, we have observed a convergence in about 15 iterations when starting from identity and in 5 to 10 iterations when starting from the least-squares solution. For the uncertainty of the solution, we replace the values of H and $\frac{\partial \Phi}{\partial \vec{z}_i}$ into equation (2) and obtain: $\Sigma_{\hat{\mathbf{f}}\hat{\mathbf{f}}} = \hat{H}^{(-1)}$

3.3 Algorithm Comparison

We have performed test on synthetic data simulating the MRI data of Sect. (4) to evaluate these two algorithms (denoted by LSQ and MAHA). Since all our features can be simplified into points, we took as reference the unit quaternion technique (QUAT) [4]. Concerning accuracy, we found that QUAT and LSQ perform very similarly, but MAHA is 1.2 to 1.5 times more accurate. LSQ and MAHA computation times are much higher than QUAT (by a factor 10 to 40) but the times are still reasonable (we have applied these algorithms to more than 500 registrations of MR images in next section).

To verify the uncertainty prediction, we used the validation index developped in [8]: under the Gaussian hypothesis, the Mahalanobis distance between the estimated and the exact transformation (the validation index) should be χ^2_6 distributed. By repeating the registration experiment, we can verify that the empirical mean value $I = \bar{\mu}^2 = \frac{1}{N}\sum \mu_i^2$ and variance σ_I^2 correspond to the expected values (here 6 and 12 for a χ^2_6), and that the empirical distribution corresponds to the exact distribution using the Kolmogorov-Smirnov (K-S) test. As expected, the most accuracy estimation is given by MAHA ($I = 6.05$, $\sigma_I^2 = 11.90$), which proves that the uncertainty estimation is very accurate. For LSQ, the uncertainty on the transformation is still well predicted (but it is larger that the one of MAHA) and QUAT needs a minimum number of 15 matches tp pass the K-S test since we have to estimate the noise on features from mesurements.

As a conclusion, MAHA gives the most accurate transformation and a good uncertainty in all cases, even with very few matches, but it should be initialized with QUAT to keep the computation time low.

4 Registration of Real MRI Images

The experiment is performed using multiple 2D contiguous Magnetic Resonance images (MRI) which constitute a 3D representation of the head. The images are part of an extensive study of the evolution of the Multiple Sclerosis (MS) disease performed at the Brigham and Woman's Hospital (Harvard Medical School, Boston). Each patient underwent a complete head MR examination several times during one year (up to 24 different 3D acquisitions). Each acquisition provides a first echo image and a second echo image (256 x 256 x 54 voxels of size .9375 x .9375 x 3mm) representing the same T2 weighted signal imaged at different echo times. Thus, they are expected to be approximately in the same coordinate system. This protocol was designed to optimize the contrast in the two channels for an easier tissue segmentation. Considering two acquisitions A and B, the registration of echo-1 images (A_1 to B_1) and echo-2 images (A_2 to B_2) give two relatively independent estimates of the genuine transformation from A to B. The comparison of these two transformations gives a Real Validation Index which can be tested for the accuracy of the uncertainty estimation.

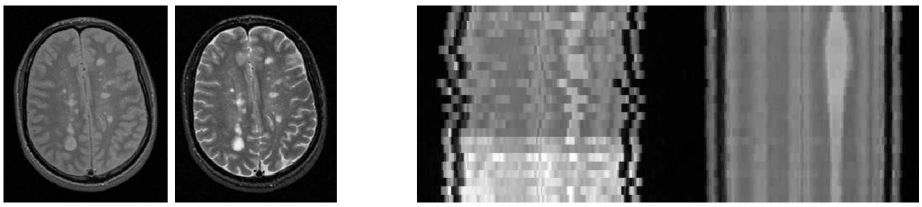
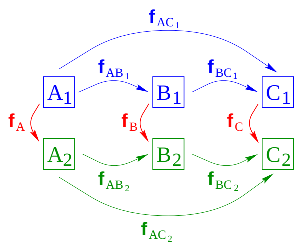


Fig. 1. Left: Example of MS images. The same slice of one acquisition in echo-1 (left) and echo-2 (right). **Right:** evolution of an image row going through a lesion across 24 time points over a year. Left: without registration; Right: after registration and intensity correction.

Our registration algorithm relies on the extraction of *Extremal Points* (see [10]) that we model as semi-oriented frames (see Sec. 2). Matches between extremal points of two images are determined using an iterative closest point algorithm adapted to such features. Typically, we match 1000 extremal points out of the about 3000 extracted with a residual mean square error (RMS) of about $1mm$. We initialize the registration with QUAT and then iterate a loop consisting of a noise estimation on features [8] followed by a MAHA registration.



In a first experiment, we compared directly the registrations between corresponding echo-1 and echo-2 images. This diagram represents three acquisitions A, B and C with the three echo-1 images (A_1, B_1, C_1) and the three echo-2 images (A_2, B_2, C_2). The echo-1 and echo-2 registrations are significantly different ($\mu^2(f_{AB_1}, f_{AB_2})$, $\mu^2(f_{AC_1}, f_{AC_2})$, $\mu^2(f_{BC_1}, f_{BC_2}) > 50$) but the intra-echo-1 and intra-echo2 registrations are compatible ($\mu^2(f_{BC_1} \circ f_{AB_1}, f_{AC_1}) \simeq 6$ and $\mu^2(f_{BC_2} \circ f_{AB_2}, f_{AC_2}) \simeq 6$). This led us to assume a global bias for each acquisition between echo-1 and echo-2 images, represented here by the transformations f_A, f_B, and f_C.

To estimate the biases, we observed first that the transformation from image A_1 to image B_2 can be written $f_{A_1 B_2} = f_B \circ f_{AB_1} = f_{AB_2} \circ f_A$. If measurements where perfect, the bias f_A could be expressed for any other image Z: $f_A = f_{AZ_2}^{(-1)} \circ f_Z \circ f_{AZ_1}$. Since measurements are noisy, we obtain an estimator of the bias f_A by taking the Fréchet mean value [7]. The biases we obtain are different for each acquisition and could be considered as translations of standard deviations $\sigma_x = 0.09$, $\sigma_y = 0.11$ and $\sigma_z = 0.13$ mm.

Although the biases appear very small, they are sufficient to explain the previous errors in the registration accuracy prediction: the mean value and standard deviation of this new index across all registrations are now very close to their theoretical value (see table 1). For the uncertainty of the transformations, we found a typical boundary precision around $\sigma_{corn} = 0.11$ mm and a typical object precision far below the voxel size: $\sigma_{obj} = 0.05$ mm for echo-1 registrations. Values are even a little smaller for echo-2 registrations: $\sigma_{corn} = 0.10$ and $\sigma_{obj} = 0.045$ mm.

	$\bar{I}$	σ_I	K-S test	num. im.	num. reg.
Theoretical values	6	$\sqrt{12} = 3.46$	0.01 – 1	$n \leq 24$	$n * (n-1)/2$
patient 1	6.29	4.58	0.14	15	105
patient 2	5.42	3.49	0.12	18	153
patient 3	6.50	3.68	0.25	14	91
patient 4	6.21	3.67	0.78	21	210

Table 1. Real Validation Index with bias correction for different patients. The mean validation index is within 10% of its theoretical value and the K-S test exhibits impressively high values.

Most of the extremal points we match are situated on the surface of the brain and the ventricles. These surfaces appear differently in echo-1 and echo-2 images due to the difference in contrast. Other artifacts such as chemical shift or susceptibility effects may also account for the observed bias as they influence the detection of extremal points. Indeed, the two echoes are acquired with different receiver RF bandwidth to improve the signal/noise ratio [6]. Therefore, the chemical shift and the susceptibility effect are different in the two echoes. In a future work, we plan to correlate the biases with diverse quantities in the images in order to understand their origin. Ultimately, we would like to predict the biases in each acquisition before registration. This would allow the indubitable validation of the registration accuracy prediction.

5 Conclusion

We show in this article how to generalize the classical least squares and least Mahalanobis distance pose estimation algorithms to generic geometric features, and how to estimate the uncertainty of the result. The uncertainty prediction on the transformation is validated on synthetic data within a bound of 5% for all algorithms, decreasing to less than 1% of inaccuracy for the algorithm MAHA. We believe that these algorithms could be easily embedded in many applications and provide a thorough basis for computing many image statistics. Further improvements could include the generalization of this framework to non rigid transformations, the estimation of a multiple registration of n sets of matched features and the coupling with statistical matching algorithms to provide a complete registration system.

From the application point of view, we show an example in medical imaging where we reached a sub-voxel registration accuracy (0.05 mm) that allowd us to detect and estimate a systematic bias in each acquisition between features extracted from the echo-1 and the echo-2 images. After the correction of the biases, multiple experiments on several patients show that our prediction of the registration accuracy is validated within a bound of 10%. These experiments demonstrate that, using adapted and rigorous tools, we can detect very small modifications in medical images. In a future work, we plan to determine the origin of the bias in order to predict it. Other applications include the automatic detection and correction of the misalignment of the slices in contiguous slice images and the statistical study of deformations in MR images.

Acknowledgments

The authors would like to thank Pr. Eric Grimson and Pr. Ron Kikinis for making this collaboration possible. Part of this work was supported by a post-doctoral fellowship from INRIA at the A.I.Lab, in Pr. Grimson team. The serial MRI exams were obtained from studies of the natural course of multiple sclerosis support by the National Institute of Health (contract no. N01-NS-0-2397).

References

1. H.F. Durrant-Whyte. Uncertain geometry in robotics. *IEEE Journal of Robotics and Automation*, 4(1):23–31, February 1988.
2. W.E.L. Grimson. *Object Recognition by Computer - The role of Geometric Constraints*. MIT Press, 1990.
3. C.R.G. Guttmann, R. Kikinis, M.C. Anderson, M. Jakab, S.K. Warfield, R.J. Killiany, H.L. Weiner, and Jolesz F.A. Quantitative follow-up of patients with multiple sclerosis using mri: reproducibility. *JMRI*, 1998. To appear.
4. B.K.P. Horn. Closed form solutions of absolute orientation using unit quaternions. *Journal of Optical Society of America*, A-4(4):629–642, April 1987.
5. B.K.P. Horn, Hilden H.M., and S. Negahdaripour. Closed form solutions of absolute orientation using orthonormal matrices. *Journal of Optical Society of America*, A-5(7):1127–1135, 1988.
6. R. Kikinis, C.R.G. Guttmann, D. Metcalf, W.M. III Wells, G.J. Ettinger, H.L. Weiner, and F.A. Jolesz. Quantitative follow-up of patients with multiple sclerosis using mri: technical aspects. *JMRI*, 1998. to appear.
7. X. Pennec and N. Ayache. Uniform distribution, distance and expectation problems for geometric features processing. *Journal of Mathematical Imaging and Vision*, 9(1):49–67, July 1998.
8. X. Pennec and J.P. Thirion. A framework for uncertainty and validation of 3D registration methods based on points and frames. *Int. Journal of Computer Vision*, 25(3):203–229, 1997.
9. P.H. Schönemann. A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, 31:1–10, 1966.
10. J-P Thirion. The extremal mesh and the understanding of 3D surfaces. *IJCV*, 19(2):115–128, 1996.
11. S. Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):376–380, April 1991.
12. M.W. Walker and L. Shao. Estimating 3-D location paramters using dual number quaternions. *CVGIP: Image Understanding*, 54(3):358–367, Nov. 1991.