Diffeomorphic Nonlinear Transformations: A Local Parametric Approach for Image Registration

R. Narayanan[1], J.A. Fessler[1,2,3], H. Park[3], and C.R. Meyer[1,3]

[1] Department of Biomedical Engineering, University of Michigan,
Ann Arbor, Michigan 48109-2099
[2] Department of Electrical Engineering and Computer Science,
University of Michigan, Ann Arbor, Michigan 48109-2122
[3] Department of Radiology, University of Michigan,
Ann Arbor, Michigan 48109-0553
{rnz, fessler, hyunjinp, cmeyer}@umich.edu

Abstract. Many types of transformations are used to model deformations in medical image registration. While some focus on modeling local changes, some on continuity and invertibility, there is no closed-form nonlinear parametric approach that addresses all these properties. This paper presents a class of nonlinear transformations that are local, continuous and invertible under certain conditions. They are straightforward to implement, fast to compute and can be used particularly in cases where locally affine deformations need to be recovered. We use our new transformation model to demonstrate some results on synthetic images using a multi-scale approach to multi-modality mutual information based image registration. The original images were deformed using B-splines at three levels of scale. The results show that the proposed method can recover these deformations almost completely with very few iterations of a gradient based optimizer.

1 Introduction

Methods for image registration have three main components: the geometric transformation used to model deformations, the objective function, and the optimization algorithm. While using Mutual Information (MI) as the objective function has been successfully explored and validated [1], finding a simple transformation possessing the useful qualities of smoothness, compact support, and the existence of an inverse has been an ongoing effort. Rigid or affine transformations cannot be used to recover local warps. Deformation fields that are solutions to Ordinary Differential Equations (ODEs) [2, 3] have been proposed because of their ability to recover large deformations while still being invertible. These methods have large number of degrees of freedom except Arsigny's Polyrigid transforms [4] and geodesic spline representations of diffeomorphisms [5]. In contrast the parametric transformation proposed here is a more parsimonious approach in that it can be applied only in regions that need correction.

Different types of radial basis functions with global e.g., Thin Plate Splines [6], and local support [7] are used as well. This is because they have fewer degrees of freedom and can be used to recover local warps. But these types of deformations are not invertible in general. In addition, functions with global support change distant regions of the

image that may not require correction while attempting to change local regions that do. B-Splines [8] have been used successfully because of their C^2 continuity and local support, but injectivity conditions are non-trivial [9]. In Arsigny's Polyrigid transforms [4] using ODEs, the deformation vector is obtained by integrating the velocity vector that is a distance weighted sum of individual vectors corresponding to 'action' points whose solution is the trajectory equation. This method always ensures that the transform is continuous and invertible. However these weights are normalized, so the transform is global. Furthermore, methods that use ODEs do not have a closed form and the deformation is computed by integrating the velocity vector in a finite number of time steps to obtain the transformation. This paper was motivated by the ideas discussed in [4]. We introduce a nonlinear transformation that possesses the properties discussed by modifying the affine transformation, so that at the center of the region that needs correction we have an affine transform described by all the parameters of the transform, and gradual convergence to identity as we move away from the center. This convergence can be controlled using our transform model. Also our transform has a closed form and is easy and fast to compute because it is characterized by few parameters and always ensures that an inverse exists under certain trivial conditions.

We show some preliminary results using a multi-scale approach to image registration by applying corrections starting from the coarsest level of scale to the finest. We applied synthetic B-Spline based deformations to images and then corrections were applied at three levels of scale using only one seed point at each. A seed point is the center of the region that we are trying to correct. They are picked based on finding high gradients of local MI with respect to local affine transformation parameters. Results show that using these transforms could be a good alternative to current methods used in image registration.

We used normalized mutual information (NMI), first proposed by Studholme [10] as the objective function and a simultaneous perturbation based gradient optimizer [11] to maximize NMI.

2   The Locally Affine Transformation Model

A global affine transformation without shear in $\mathbb{R}^n$ for any vector $\mathbf{x}$ about the center $\mathbf{x}_0$ is

$$T(\mathbf{x}) = e^{sA}e^{sS}(\mathbf{x} - \mathbf{x}_0) + s\mathbf{t} + \mathbf{x}_0, \tag{1}$$

where $\mathbf{x} = \begin{bmatrix} x_1 \ x_2 \ \dots \ x_n \end{bmatrix}^T$, $\mathbf{x}_0 = \begin{bmatrix} x_{01} \ x_{02} \ \dots \ x_{0n} \end{bmatrix}^T$, $\mathbf{t} = \begin{bmatrix} t_1 \ t_2 \ \dots \ t_n \end{bmatrix}^T$ (translation), A is the skew symmetric matrix corresponding to the rotation matrix, S is the symmetric matrix corresponding to the scale matrix and $s \in [0,1]$.

At $s = 1$ we get the complete affine transformation about the center $\mathbf{x}_0$. Many such centers will be chosen from the image as requiring correction. The parameter s in the above equation can be parameterized in space in the form of a continuous function, say $\lambda(r)$ where $r = \|\mathbf{x} - \mathbf{x}_0\|$ so that at $r = 0$, $s = 1$ and as r increases $s \to 0$. The elegance of writing it this way is that as we move towards the center of the region that we are attempting to correct, we have an affine transformation, but the transformation converges to an identity map as we move away. The region of influence can be controlled by a

parameter of the continuous function. So any function with the above properties can be used. Although the Gaussian does not have a compact support it was used because it could be treated as being almost local for small σ. We used the Gaussian because of its C^∞ smoothness and loose bounds for an inverse to exist. The proposed transformation ($T : \mathbb{R}^n \to \mathbb{R}^n$) is

$$T(\mathbf{x}) = e^{\lambda(r)A} e^{\lambda(r)S}(\mathbf{x} - \mathbf{x}_0) + \lambda(r')\mathbf{t} + \mathbf{x}_0 \tag{2}$$

where

$$r' = \|e^{\lambda(r)A} e^{\lambda(r)S}(\mathbf{x} - \mathbf{x}_0)\| = \|e^{\lambda(r)S}(\mathbf{x} - \mathbf{x}_0)\|$$

and

$$s \equiv \lambda(r) = e^{-\frac{r^2}{2\sigma^2}}. \tag{3}$$

One can also write T in Eq. (2) as

$$T(\mathbf{x}) = (T_T \circ T_{RS})(\mathbf{x}) + \mathbf{x}_0 \tag{4}$$

where

$$T_{RS}(\mathbf{x}) = e^{\lambda(r)A} e^{\lambda(r)S}(\mathbf{x} - \mathbf{x}_0) \tag{5}$$

and

$$T_T(\mathbf{x}) = \mathbf{x} + \lambda(\|\mathbf{x}\|)\mathbf{t}. \tag{6}$$

Above, σ^2 is the variance of the Gaussian modulation function. It sets the scale at which we are working in the registration step and is also fine tuned (optimized along with the affine parameters) to match the scale at which the deformations were induced. The transformation is nonlinear and can be applied to each region individually. A region is picked if it has a large gradient of local MI. The center of this region is $\mathbf{x}_0$ about which the transformation is applied.

The transform also satisfies many desirable properties discussed in the following subsections. These properties depend on the choice of λ which can be chosen to be local and smooth. The properties discussed here are for $\lambda(r)$ chosen to be gaussian in Eq. (3). One may select the function λ based on what properties one seeks to satisfy.

2.1   Continuity and Locality

Continuity is determined by the choice of the function λ. The Gaussian ensures C^∞ continuity. Locality also depends on λ. The Gaussian function has "nearly" local support. Functions with strictly local support may also be used to arrive at different conditions for an inverse to exist.

2.2   Existence of Inverse

Current methods using spline-based deformation models have either difficult conditions to incorporate in the optimizer to prevent folding, i.e not invertible or use regularization methods that discourage folding by adding an additional smoothness term in the objective function [8]. In our method we derive loose bounds for the transformation parameters which are straightforward to implement and always ensure invertibility.

The Jacobian matrix for a transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ must be positive definite everywhere to ensure invertibility. We have found the conditions for which the determinant of the Jacobian of the transformation is positive to always guarantee an inverse (see Appendix). We picked λ to be Gaussian because of its loose bounds, infinite continuity and an easily controllable region of influence. Other functions like inverse multiquadratics or differentiable local support functions of the type proposed by Wendland [12] may also be used and lead to similar conditions.

As shown in Eq. (2) the transformation T has an inverse as long as

$$\|\mathbf{t}\| < \sigma e^{\frac{1}{2}}$$

and

$$0 < a < e^{e^{0.5}} \approx 5.2003,$$

where $\mathbf{t}$ is the translation vector and $a = max(a_x, a_y)$, the larger of the two anisotropic scales in the x and y direction. These bounds in practice were found to be very loose and we never experienced any folding in our simulations.

3   Initialization and Registration

3.1   Initialization

We implement a multi-scale approach to image registration starting from the coarsest level of scale and proceeding to the finest.

At each level of scale we pick only regions that are mis-registered and apply the algorithm. Rohde et al. [7] picked regions with large gradient of cost function with respect to radial basis function coefficients while Park et al. [13] used a mismatch measure to quantify mis-registration.

Here we pick regions based on its sensitivity to local affine deformations. Since we apply corrections based on a locally affine transformation model, the gradients computed give us a meaningful estimate on the extent of mis-registration. The way these gradients are computed is as follows. A rectangular window is picked with dimensions in correspondence with the scale and three control points are placed in a triangular fashion spanning the area of the window. The window is placed in the reference and the floating image and the control points in the floating image are perturbed and the gradient of NMI with respect to the affine coefficients is found. This window is moved over the complete reference and floating image in an overlapping fashion. If the gradients of the cost in a region is not small, then it is likely that this region is mis-registered. Regions with large magnitude of gradient norm above a selected threshold are picked and the centers of these regions denoted as seed points are used in the global registration step. If p_i are the parameters that define our affine transformation, the gradient of local NMI is computed as

$$\hat{\mathbf{g}} = \left[\frac{\partial NMI}{\partial p_1} \frac{\partial NMI}{\partial p_2} \ldots \frac{\partial NMI}{\partial p_6} \right]^T$$

We apply transformations about these points and correct for them locally using the transformation model at different levels of scale. Since these points are also fed as parameters to the optimizer they will also be allowed to move to model the deformation better.

3.2   Multi-scale Nonrigid Registration

The final deformation is computed iteratively across different levels of scale. Since the spatial support of the deformation can be constrained to be local, seed points are picked in the initialization step at these different scales, and they serve as the centers for our locally affine transformation model. Global registration is then initiated at the coarsest level of scale (large σ) and optimization is performed over all the seed points with large to smallest σ. The final transform is computed as a composition of individual transformations.

After optimizing over each region, the geometric maps are stored and this is repeated over other regions of the image. Since each of these transforms correct for only one region at a time, they have very few parameters and high local sensitivity yielding their ability to model local changes accurately. Also, only regions that are mis-registered are picked and corrected instead of placing a grid of control points and picking which ones are active (needing optimization) and inactive. This gives us a finer control over the region we are trying to correct. E.g. if we have N seed points, the final transformation is

$$T(\mathbf{x}) = (T_N \circ T_{N-1} \ldots T_2 \circ T_1)(\mathbf{x}),\qquad(7)$$

where each seed point 'i' is associated with a transformation T_i.
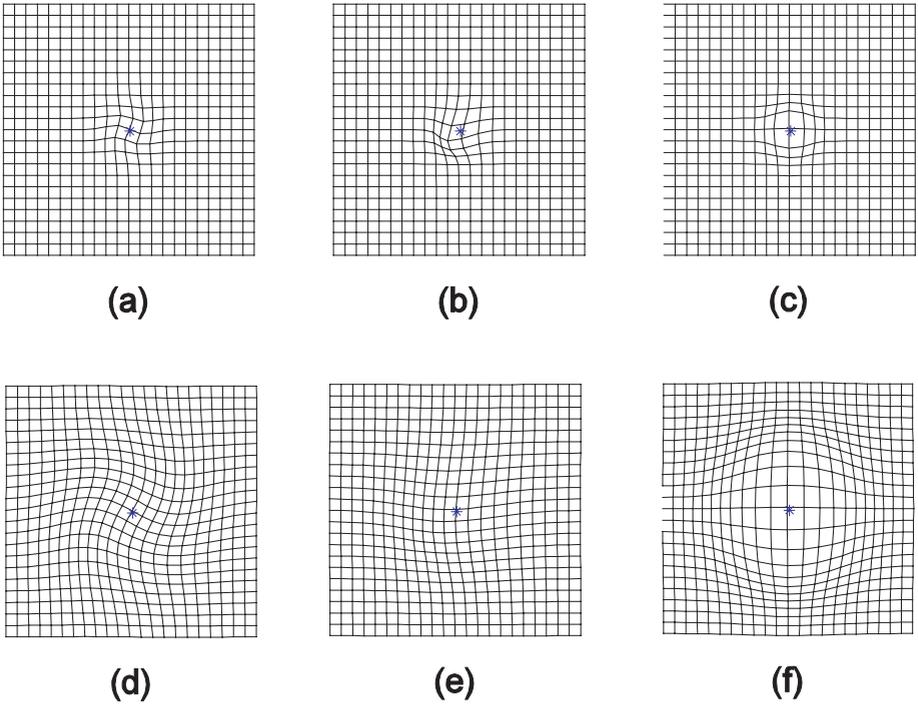

Fig. 1. Deformations applied to a uniform grid at two different levels of scale (σ). The figure shows the same amount of rotation (a and d), translation (b and e) and scale (c and f) applied individually for a small and large σ respectively
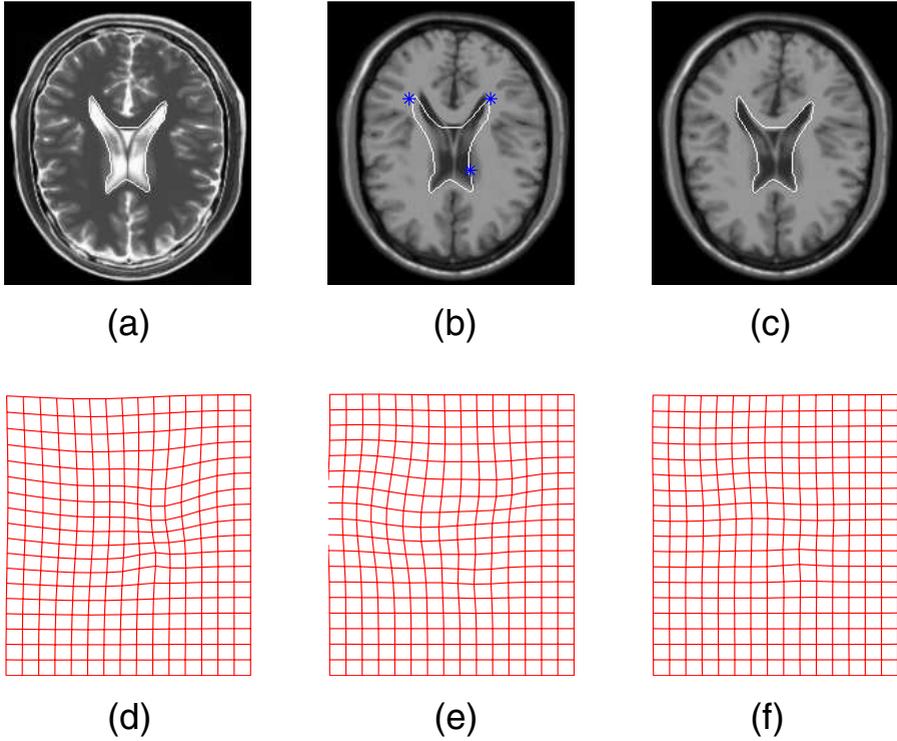
Fig. 2. Registration of T1 and T2 weighted slices using three seed points (a) Original T2 weighted reference image. (b) artificially deformed T1 weighted floating image. (c) T1 weighted floating image after registration. (d) Applied Deformation (e) Estimated inverse after registration (f) Estimated inverse applied to the induced deformation

The reference and the floating image are assumed to be already affine registered with each other before we begin the algorithm. The individual transformation parameters are computed for each seed point. There may be several seed points identified at a level of scale. Global normalized mutual information was used as the objective function and a simultaneous perturbation based gradient optimizer proposed by Spall [11] was used to arrive at the final solution. All eight parameters corresponding to the transformation were optimized: i.e. two translation parameters (t_x and t_y), rotation angle (θ), two anisotropic scale parameters in the scale matrix (a_x and a_y), two center coordinates (C_x and C_y) and a variance parameter (σ from the Gaussian function).

Algorithm

1: Initialize reference (A) and floating images (B) and set T_0 to an identity map
2: **for** $i = 1$ *to Levels of Scale* **do**
3:    $M = \#$ of seed points picked based on high local gradients
4:    **for** $k = 1$ *to M* **do**
5:        $\hat{T}_{i,k} = argmax_{T_{i,k}} NMI(A(\bullet), B((T_{i,k} \circ \hat{T}_{i,k-1} \ldots \hat{T}_{i,2} \circ \hat{T}_{i,1} \circ \hat{T}_{i-1} \circ \hat{T}_{i-2} \ldots \hat{T}_0)(\bullet)))$

6:     **end for**
7:     $\hat{T}_i = \hat{T}_{i,k}$
8: **end for**
9: $\hat{T} = \hat{T}_i$

4   Results

4.1   Examples of Locally Affine Deformations

Fig. (1) shows examples of rotate, translate and scale applied individually about one seed point for two different σ. This is to show that we can model all kinds of local and global changes using a combination of these parameters.

4.2   Registration Experiments

Fig. (2) shows a head registration example using an axial slice from T1 and T2 weighted images from Brainweb [14]. They were artificially deformed using B-Splines at three
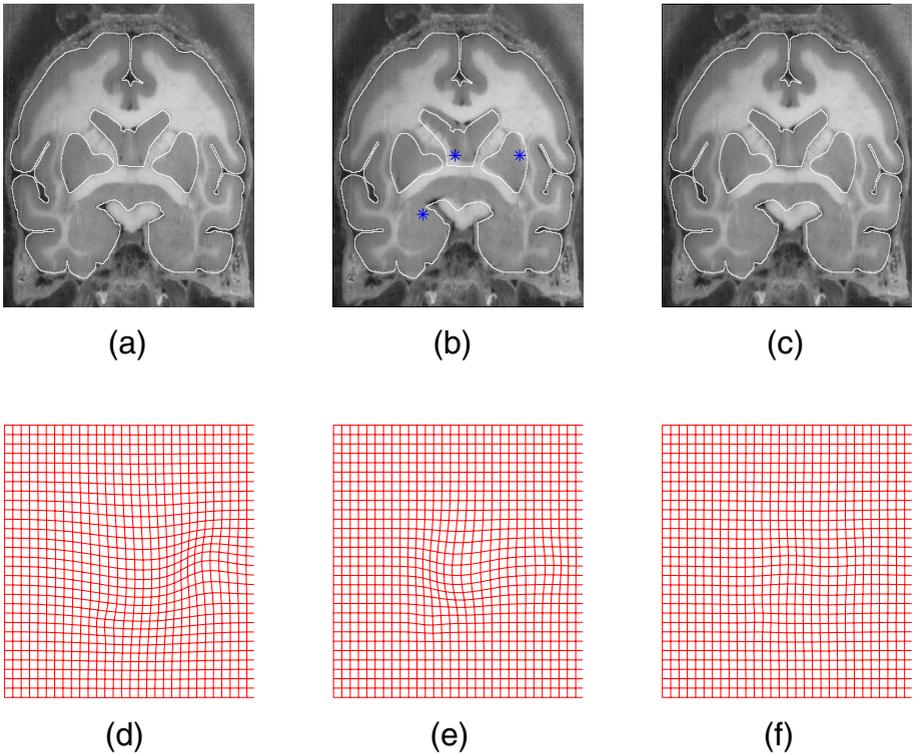

Fig. 3. Registration of a coronal slice of a vervet monkey using three seed points. (a) Original reference image. (b) artificially deformed floating image. (c) Reconstructed floating image after registration. (d) Applied Deformation (e) Estimated inverse after registration (f) Estimated inverse applied to the induced deformation
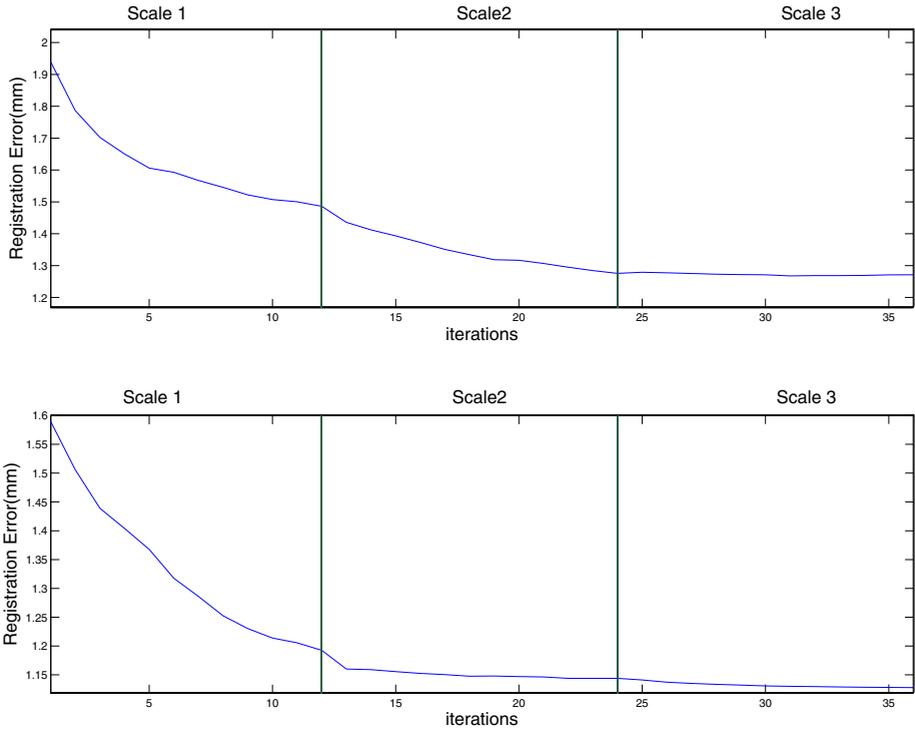
Fig. 4. Average norm of registration error vs. iterations at three different levels of scale (σ). (a) Human brain - T1, T2 weighted registration. (b) Monkey brain registration

different levels of scale. This was done by moving one knot in the B-Spline grid by a known amount at each scale, refining the grid and repeating the procedure at the next level. Registration was performed using three seed points each working at a different scale. i.e different σ. The outline for the ventricle was marked manually so that the registration performance could be visually assessed. In Fig. (2), (a) is the original T2-weighted reference image and (b) is the deformed floating image with three seed points marked. The seed points from left to right are in decreasing levels of scale (σ), i.e. three optimizations were performed, one at each scale. The registered T1 image in (c) shows that the ventricles follow the contours more tightly after registration. (d) shows the applied deformation, (e) is the estimated inverse obtained via registration and (f) shows the deformation computed(e) applied to the induced deformation(d) which should resemble a uniform grid as best as possible.

Fig. (3) shows a coronal slice from a vervet monkey atlas developed at UCLA's Laboratory of Neuroimaging [15]. The slice was deformed using B-Splines similar to the procedure described in the previous paragraph. Seed points were placed at exactly three locations each with a different variance ($\sigma_{center} > \sigma_{right} > \sigma_{left}$) for the Gaussian function that controls the support of the transformation. The contours in (c) shows that the boundaries of the caudate and putamen hug the manually segmented boundaries more tightly, a marked improvement from (b).

Fig. (4) shows the registration error versus iterations at three different levels of scale for the first and second example respectively. Only twelve iterations were performed at each scale and each experiment took less than three minutes to run on a 3.2 Ghz PC with 2 Gb memory running MATLAB 7.

5   Discussion

We have demonstrated and tested a new local nonlinear transformation using a multi-scale approach to multi-modality image registration. This transformation has good local properties and affine behavior near the region of interest. The parameters controlling the support of this transform (σ) can be initialized and changed (by the optimizer) during the course of the registration to match the level of scale of the induced deformation. Furthermore, the transformation has a closed form and there is no need to integrate the velocity vector over time as in the case of methods using ODEs, making it very fast. Since each region is optimized for one at a time, only eight parameters are used in the optimization which makes it very fast. Although we can always guarantee that folding does not occur, finding a direct inverse is not straightforward. A numerical inverse could be found finally using the optimized parameters. Since this is computed only after estimating the transformation (T in Eq. (7)), it could be easily done using any numerical method at the end if required.

5.1   Translation

Consider the case where the vector $\mathbf{x}$ is subjected to pure translation with no rotation or scaling (i.e. $T(\mathbf{x}) = \mathbf{x} + \lambda(\|\mathbf{x} - \mathbf{x}_0\|)\mathbf{t}$). Using the condition that $det(J) > 0$ we get,

$$\bar{\mathbf{x}}^T \mathbf{t} < \frac{1}{d}.$$

Applying the Cauchy-Schwarz inequality to the left hand side and substituting $\|\bar{\mathbf{x}}\| = 1$ we get the sufficient condition

$$\|\mathbf{t}\| < \sigma e^{\frac{1}{2}}, \tag{8}$$

where $\sigma e^{\frac{1}{2}}$ is the smallest value that $\frac{1}{d}$ can assume.

5.2   Rotation

For $\mathbf{x} \in \mathbb{R}^n$, the rotation matrix can be constructed as the composition of elementary rotations in planar subspaces. Each of these matrices is a Jacobi rotation matrix. The rotation matrix is invertible as long as each of these matrices has an inverse. The Jacobian for a Jacobi matrix corresponding to the planar subspace containing axes 'i' and 'j' is given by

$$J = Q_{ij} \begin{bmatrix} p_1 & p_2 & 0 \\ p_3 & p_4 & \\ 0 & & I_{n-2} \end{bmatrix} Q_{ij}$$

where

Q_{ij} is the permutation matrix

$p_1 = \cos(\lambda\theta) + r\theta d\bar{x}_i^2 \sin(\lambda\theta) + r\theta d\bar{x}_i\bar{x}_j \cos(\lambda\theta)$,

$p_2 = -\sin(\lambda\theta) + r\theta\bar{x}_i\bar{x}_j d \sin(\lambda\theta) + r\bar{x}_j^2\theta d \cos(\lambda\theta)$,

$p_3 = \sin(\lambda\theta) - r\bar{x}_i^2\theta d \cos(\lambda\theta) + r\bar{x}_i\bar{x}_j\theta d \sin(\lambda\theta)$ and

$p_4 = \cos(\lambda\theta) - r\bar{x}_i\bar{x}_j\theta d \cos(\lambda\theta) + r\bar{x}_j^2\theta d \sin(\lambda\theta)$.

The determinant of this matrix is always 1. The volume is always preserved under rotation. So the transformation $(T(\mathbf{x}) = e^{\lambda(r)A}(\mathbf{x} - \mathbf{x}_0) + \mathbf{x}_0)$ always has an inverse

5.3   Scale

Finally consider the case when the transformation consists of only scaling. (i.e. $T(\mathbf{x}) = e^{\lambda(r)S}(\mathbf{x} - \mathbf{x}_0) + \mathbf{x}_0$). One can show that the determinant of the Jacobian is $1 - rd\sum_{j=1}^{n} \bar{x}_j^2 \log(a_j)$. Applying the conditions for an inverse to exist(i.e $det(J) > 0$) we get

$$rd \sum_{j=1}^{n} \bar{x}_j^2 \log(a_j) < 1$$

where a_j are the anisotropic scales in each dimension. Let $a = max(a_1, a_2, \ldots, a_n)$. Replacing a_j above with a we get a more stringent inequality

$$rd \log(a)\|\bar{\mathbf{x}}\|^2 < 1.$$

Substituing $\|\bar{\mathbf{x}}\|^2 = 1$ and rearranging above we get

$$a < e^{\frac{1}{rd}}.$$

Since the minimum value that $\frac{1}{rd}$ can assume is easily shown to be $e^{0.5}$, the sufficient condition is

$$0 < a < e^{e^{0.5}} \approx 5.2003. \tag{9}$$

5.4   Conditions for Inverse

We have derived the bounds so that an inverse always exists for rotation, translation and scale each individually applied. Let T_R be the isomorphism for pure rotation (i.e. no scale or translation) so that $T_R(\mathbf{x}) = e^{\lambda(r)A}(\mathbf{x} - \mathbf{x}_0) + \mathbf{x}_0$ and let $T_{R'}(\mathbf{x}) = T_R(\mathbf{x}) - \mathbf{x}_0$. We need to show that T in Eq. (4) has an inverse. Let us first show that T_{RS} in Eq. (5) is invertible. The transformation T_{RS} is

$$T_{RS}(\mathbf{x}) = e^{\lambda(r)A}e^{\lambda(r)S}(\mathbf{x} - \mathbf{x}_0)$$
$$= P\, T_{R'}(\mathbf{x})$$

where

$$P = e^{\lambda(r)A} \begin{bmatrix} a_x^{\lambda(r)} & 0 \\ 0 & a_y^{\lambda(r)} \end{bmatrix} e^{-\lambda(r)A}.$$

Being similar to a diagonal matrix P is invertible. Also $T_{R'}$ is always invertible since T_R is. So T_{RS} always has an inverse as long as $a < e^{e^{0.5}} \approx 5.2003$.

We have already proved that T_T in Eq. (4) has an inverse as long as $\|\mathbf{t}\| < \sigma e^{\frac{1}{2}}$.(See Eq. (8)). So the transformation $T(\mathbf{x}) = (T_T T_{RS})(\mathbf{x}) + \mathbf{x}_0$ has an inverse as long as

$$\|\mathbf{t}\| < \sigma e^{\frac{1}{2}} \tag{10}$$

and

$$0 < a < e^{e^{0.5}} \approx 5.2003. \tag{11}$$

References

1. Viola, P., Wells, W.M.: Alignment by maximization of mutual information. Int. J. Comput. Vision **24** (1997) 137–154
2. Joshi, S.C., Miller, M.I.: Landmark matching via large deformation diffeomorphisms. IEEE Trans. Medical Imaging **9** (2000)
3. Hotel, C.C., Hermosillo, G., Faugeras, O.: Flows of diffeomorphisms for multimodal image registration. In: ISBE. (2002) 753–756
4. Arsigny, V., Pennec, X., Ayache, N.: Polyrigid and polyaffine transformations: A new class of diffeomorphisms for locally rigid or affine registration. In: MICCAI (2). (2003) 829–837
5. Marsland, S., Twining, C.J.: Constructing diffeomorphic representations for the groupwise analysis of nonrigid registrations of medical images. IEEE Trans. Medical Imaging **23** (2004)
6. Meyer, C.R., Boes, J.L., Kim, B., Bland, P.H., Zasadny, K.R., Kison, P.V., Koral, K., Frey, K.A., Wahl, R.L.: Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations. Med. Image. Anal. **1** (1997) 195–206
7. Rohde, G.K., Aldroubi, A., Dawant, B.M.: The adaptive bases algorithm for intensity based nonrigid image registration. IEEE Trans. Med. Imaging **22** (2003) 1470–1479
8. Ruckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: Application to breast MR images. IEEE Trans. Med. Imaging **18** (1999) 712–721
9. Choi, Y., Lee, S.: Injectivity conditions of 2D and 3D uniform cubic B-Spline functions. Graphical Models (**62**)
10. Studholme, C., Hill, D., Hawkes, D.: An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition **32** (1999) 71–86
11. Spall, H.C.: Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. IEEE Trans. Autom. Control **37** (1992) 332–341
12. Wendland, H.: Piecewise polynomial positive definite and compactly supported radial basis functions of minimal degree. AICM **4** (1995) 389–396
13. Park, H., Bland, P.H., Brock, K.K., Meyer, C.R.: Adaptive registration using local information measures. Medical Image Analysis **8** (2004) 465–473
14. Cocosco, C.A., Kollokian, V., Kwan, R.K.S., Evans, A.C.: Brainweb: Online interface to a 3D MRI simulated brain database. NeuroImage **5** (1997)
15. Rubins, D.J., Ambach, K., Toga, A.W., Melega, W.P., Cherry, S.R.: Development of a digital brain atlas of the vervet monkey. BrainPET **4** (1999)

Appendix

Here we derive conditions under which an inverse exists for a Gaussian weighting function. The conditions have been derived for a vector $\mathbf{x} \in \mathbb{R}^n$(i.e. $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$) for

rotation and scale about $\mathbf{x}_0 = \begin{bmatrix} x_{01} & x_{02} & \dots & x_{0n} \end{bmatrix}^T$. The bounds for 2D and 3D that we are interested in will turn out to be the same as the N dimensional case.

We will derive the conditions for translation, rotation and scale each treated individually and will show that these are sufficient conditions for an inverse to always exist.

For λ gaussian,

$$\frac{\partial \lambda(r)}{\partial x_i} = -\frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \frac{(x_i - x_{0i})}{r} = -d\bar{x}_i,$$

where $d = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}$ and $\bar{x}_i = \frac{x_i - x_{0i}}{r}$. Let $\bar{\mathbf{x}}$ be the direction cosine vector so that $\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 & \bar{x}_2 & \dots & \bar{x}_n \end{bmatrix}^T$ and $\|\bar{\mathbf{x}}\| = 1$.