Fast Nonrigid Image Registration using Salient Region Feature and Free Form Deformations

*Xiaolei Huang[1], Ph.D., Chenyang Xu[2], Ph.D. and *Yulin Song[3], Ph.D.*

[1]Computer Science and Engineering Department
Lehigh University
19 Memorial Drive West
Bethlehem, PA 18015
Phone: +1 (610) 758-4818
Fax: +1 (610) 758-4096
Email: huang@cse.lehigh.edu

[2]Imaging and Visualization Department
Siemens Corporate Research, Inc.
755 College Road East
Princeton, NJ 08540
Phone: +1 (609) 734-6535
Fax: +1 (609) 734-6565
Email: chenyang.xu@siemens.com

[3]Department of Medical Physics
Memorial Sloan Kettering Cancer Center
New York, N.Y. 10021

Introduction

Image registration aims to spatially align one image to another and establish point-to-point correspondences between the images. It is an active and important topic of research in medical image analysis because of its numerous applications including change detection in longitudinal studies, multi-modal image fusion, motion correction in time series, image-guided intervention, anatomical atlas generation and group analysis. Registration happens at both global and local levels. Global registration methods recover parameters of a global transformation model, such as rigid and affine, to account for global differences such as translation, rotation, and scaling. In many cases of medical image registration, local registration that utilizes an elastic or non-rigid deformation model is also required to cope with local differences between the images. Such local differences are due to, for examples, scanner-induced deformations, movement and/or respiratory motion of the patient, surgical intervention, or different anatomy (e.g. in atlas generation). This chapter will review the fundamentals of medical image registration, and describe a novel efficient non-rigid image registration method. The method is based on finding salient region-feature correspondences between images and solving for parameters of a Free Form Deformations (FFD) model in closed form.

Background

There has been intensive research on image registration [Collignon et al. 1995, Viola et al. 1995, Maes et al. 1997, Can et al. 2002, Hartkens et al. 2002, Keller et al. 2003, Deligianni et al. 2006] and the proposed methods differ in two main aspects. One aspect is *Transformation,* which refers to the selected global, local, or hierarchical (global-to-local) geometric transformation model, used for transforming a *moving* image to achieve high spatial correspondence with a *fixed* image. Global transformation models apply to an entire image; and examples are rigid, similarity, affine and perspective. Local transformation models can represent pixel-wise deformations that deform an image locally and non-rigidly; and examples include optical flow [Paragios et al. 2003], Thin Plate Splines (TPS) [Chui et al. 2000], Radial Basis Functions [Fornefett et al. 1999], and space deformation techniques such as Free Form Deformations (FFD) [Huang et al. 2006, Rueckert et al. 1999]. Hierarchical models that typically consist of a global transformation followed by local deformations are also popular since they cover the entire transformation domain. The other aspect in registration is the *Registration Criterion*, which is the approach used to recover optimal transformation parameters. One can classify existing approaches into three categories: feature-based methods, intensity-based methods, and hybrid methods that integrate the previous two. Feature-based methods [Thirion et al. 1996, Maurer et al. 1998, Can et al. 2002] establish explicit correspondences between sparse geometric features such as points, curves, and/or surface patches and then estimate the optimal transformation parameters based on these correspondences. These methods are relatively efficient and can register images in arbitrary poses, although there are concerns on the robustness of feature extraction and correspondence computation. Intensity-based registration methods [Viola et al. 1995, Collignon et al. 1995, Rueckert et al. 1999, Maes et al. 1998] operate directly on pixel intensity values from the full image content, and iteratively solve for optimal transformation parameters that would minimize a dissimilarity (or maximize a similarity) energy function defined on the two images. These methods do not require feature extraction and can be made fully automatic, thus they are widely used in medical image registration applications; they can also register images of multiple modalities (e.g. CT & MR, CT & PET, MR & PET) by using appropriate similarity measures. The most commonly-used similarity measures include Mutual Information [Viola et al. 1995] and Normalized Cross Correlation [Goshtasby et al. 1984]. Intensity-based methods often tend to be more computationally expensive than feature-based methods however, due to the use of full image content and the need for optimization on complex, non-convex energy functions. In recent years, hybrid methods are also proposed to integrate the merits of both feature-based and intensity-based methods. Most hybrid methods [Shen et al. 2002, Hartkens et al. 2002, Keller et al. 2003] focus on incorporating user-provided or automatically-extracted geometric feature constraints into intensity-based energy functions to achieve smoother and faster optimization. Some [Huang et al. 2004] extract scale-invariant local *region* features instead of *point* features so that robust similarity measures (including those capable of multi-modal registration) can be applied to matching the interior intensities of region features and finding reliable correspondences.

Challenges remain in image registration to cope with conditions such as structure appearing/disappearing, intensity differences in multi-modal images, image noise, intensity inhomogeneity, and partial matching. Particularly in nonrigid (or elastic) registration, most state-of-the-art methods are intensity-based methods that are too slow to be useful in clinical applications, because the number of transformation parameters (i.e. degrees of freedom) to be estimated is very large. To address these challenges, we propose a novel hybrid method for nonrigid image registration. The method is very efficient, in that it utilizes sparse *salient region* feature correspondences to estimate both global rigid transformation and local deformation field in closed form. The local deformations are represented by a Cubic B-spline based Free Form Deformations (FFD) model, which offers compact support so that registration can be locally constrained which especially allows to deal with local changes in medical images due to, for instance, tumor resection. In a multi-resolution manner, the method can also recover small to large non-rigid deformations. The resulting deformation field gives one-to-one mapping between the two images being registered and the field is smooth with guaranteed C^1 continuity at FFD control points and C^2 continuity everywhere else. We demonstrate this new approach on registering MR brain images.

Fast Non-rigid Image Registration using a Closed-form FFD Solution

This registration method is a hybrid framework that generally comprises the following steps.

1. *Salient Region Feature Extraction*. Detecting scale-invariant salient region features on the fixed image and on the moving image;
2. *Feature Correspondences and Rigid Alignment*. For every feature X_i on the fixed image, find its corresponding feature $X_i^{'}$ on the moving image, by matching the region-feature interior intensities using intensity similarity measures such as Mutual Information (MI) and Squared Normalized Cross Correlation. Outlier correspondences are pruned according to geometric constraints between features. The resulting region-feature correspondences are used to estimate a global rigid transformation that brings the moving image to global alignment with the fixed image;
3. *Closed-form FFD Solution for Non-rigid Registration*. In order to accurately recover the many parameters in a local free-form deformation (FFD) model, more geometric features including corner features are extracted on the globally-aligned moving and fixed images. Corner feature correspondences are established by local search and neighborhood similarity matching. Then the salient-region-feature correspondences from step 2, together with the newly detected corner-feature correspondences, are used to solve for a dense deformation field that establishes point-to-point correspondences between the moving and the fixed images. The deformation field is represented using a Cubic B-

spline based FFD model, whose parameters are estimated in closed-form from feature correspondences.

Salient Region Feature Extraction.

Scale-invariant salient region features on an image are found using an entropy-based detector [Kadir et al. 2001], which selects regions with highest local saliency in both spatial and scale spaces. For each pixel $\mathbf{x}$ on an image, let us consider a local circular region, $R(s, \mathbf{x})$, of certain scale described by radius s and centered at pixel $\mathbf{x}$. An intensity probability density function, $p(s, \mathbf{x})$, is computed from all the pixel intensities within the circular region. Then the differential entropy of intensities in this local region is defined by:

$$H(s,\mathbf{x}) = -\int p_i(s,\mathbf{x}) \log_2 p_i(s,\mathbf{x}) di$$

where i takes on values in the set of possible intensity values, and $p_i(s, \mathbf{x})$ is the probability of intensity value i in the local region. The “best” scale of the local region centered at $\mathbf{x}$ is selected as the one that maximizes the local region entropy: $S_{\mathbf{x}} = \arg\max_s H(s,\mathbf{x})$. The saliency value, $A(S_{\mathbf{x}}, \mathbf{x})$, for the region with the best scale is defined by the extreme entropy value, weighted by the best scale and a differential self-similarity measure in the scale space:

$$A(S_{\mathbf{x}},\mathbf{x}) = H(S_{\mathbf{x}},\mathbf{x}) \cdot S_{\mathbf{x}} \cdot \int \left\| \frac{\partial}{\partial s} p_i(s,\mathbf{x}) \middle| S_{\mathbf{x}} \right\| di$$

Since the saliency metric is applicable over both spatial and scale spaces, the saliency values of region features at different locations and scales are comparable. After computing the best scale and saliency value at every pixel in the image, we extract salient region features as those local regions centered at pixels with local maxima saliency values. The scale (or radius) of a salient region is defined by the best scale associated with the pixel at its center. The main advantages of such extracted region features are that they are theoretically invariant to rotation, translation and scaling, and their interior areas provide intensities that can be matched using intensity-based similarity measures. Examples on salient region features can be seen in Figure 1(II).

Feature Correspondences and Rigid Alignment.

To detect salient-region-feature correspondences between the fixed image and the moving image, a matching algorithm not only evaluates the likelihood of each hypothesized fixed-moving pairing of features, but applies geometric constraints (e.g. relative position and orientation) between features to ensure the global consistency of individual matches and thus eliminates false feature matches. Given a pair of region features, (f_i, m_j), where f_i is the ith feature on the fixed image f, and m_j is the jth feature on the moving image m, the likelihood of this pairing being a true correspondence is determined by the intensity similarity between the two features. We measure such similarity using Normalized Mutual Information (NMI), which can cope with multi-modal matching. Suppose the intensity random variable of region f_i is I_1, and the intensity random variable of region m_j is I_2. Denote the intensity probability density function (p.d.f.) of f_i as p_{f_i}, the p.d.f. of m_j as p_{m_j}, and the joint p.d.f. of f_i and m_j as p_{f_i,m_j}. Let the differential entropy of intensities in region f_i be:

$$H(f_i) = -\int p_{f_i}(I_1)\log_2 p_{f_i}(I_1)dI_1,$$

the differential entropy of intensities in region m_j be:

$$H(m_j) = -\int p_{m_j}(I_2)\log_2 p_{m_j}(I_2)dI_2,$$

and the joint entropy between intensities in region f_i and intensities in region m_j be:

$$H(f_i, m_j) = -\iint p_{f_i,m_j}(I_1, I_2)\log_2 p_{f_i,m_j}(I_1, I_2)dI_1 dI_2,$$

Then the normalized mutual information between f_i and m_j is defined by:

$$NMI(f_i, m_j) = 2 - \frac{2H(f_i, m_j)}{H(f_i) + H(m_j)}$$

We make the NMI measure invariant to translation, rotation, and scaling, by aligning the two region centers in a local coordinate system, normalizing their scales by supersampling the smaller region to match the scale of the larger region, and taking the maximum NMI value under all possible rotations. (In practice, we sample the rotation parameter space and take the maximum NMI value under all sample rotations.) Once having the NMI value computed for every hypothesized fixed-moving feature pair, we take the pairs with the greatest NMI values to be top candidate correspondences. Feature correspondences found this way are shown in Figure 1(III.a-b). We then apply a generalized Expectation-Maximization algorithm [Huang et al. 2004] to prune outliers in the detected individual feature correspondences and estimate a rigid transformation that transforms the moving image *m* to be globally aligned with the fixed image *f* (see Fig. 1(III)).

Closed-form FFD Solution for Nonrigid Registration.

In order to estimate robustly a local deformation field and establish dense point-to-point correspondences between *m* and *f*, we need as many reliable feature correspondences as possible. Thus in addition to the salient-region-feature correspondences above, we apply the Harris corner detector [Harris et al. 1988] to extract corner features on the moving and the fixed images. Since *m* and *f* are now globally aligned, correspondences between corners on *m* and on *f* are established by local search and normalized cross correlation. Now given the set of feature (including region and corner feature) correspondences between *m* and *f*, and representing the local deformations using a Cubic B-spline based Free Form Deformations (FFD) model, we can solve an inverse inference problem to recover parameters of the optimal FFD deformation field that can explain the position displacements between corresponding features.

FFD is a space warping technique that deforms an image by manipulating a regular control lattice P overlaid on its volumetric embedding space [Rueckert et al. 1999, Huang et al. 2006]. To facilitate notation, we describe FFD in the 2D case. Let us consider a regular lattice of control points $P_{m,n} = (P^x_{m,n}, P^y_{m,n}); m = 1,..., M, \ \ n = 1,..., N$, overlaid to an image domain $\Gamma=\{\mathbf{x}\}=\{(x,y)| \ l_x \le x \le h_x, \ l_y \le \ y \le h_y\}$. Denote the initial regular configuration of the control lattice as P^0, and the deforming control lattice as $P=P^0+\delta P$. The parameters of FFD are the deformations of the control points in both x and y directions;

$$\Theta = \{(\delta P^x_{m,n}, \delta P^y_{m,n})\}; (m,n) \in [1, M] \times [1, N]$$

Given the deformation of the control lattice from P^0 to P, the deformed location $L(\mathbf{x})=(x',y')$ of a pixel $\mathbf{x}=(x,y)$, is defined in terms of a tensor product of Cubic B-splines [Huang et al. 2006]:

$$L(\mathbf{x}) = \mathbf{x} + \delta L(\mathbf{x}) = \sum_{k=0}^{3}\sum_{l=0}^{3} B_k(u)B_l(v)(P^0_{i+k,j+l} + \delta P_{i+k,j+l})$$

In the definition, $P_{i+k,j+l}$; $k=0,\ldots,3$, $l=0,\ldots,3$, refer to the coordinates of the nearest sixteen (4 by 4) control points in the neighborhood of pixel $\mathbf{x}$; $i = \left\lfloor \frac{x - l_x}{h_x - l_x} \right\rfloor \cdot (M-1)$, $j = \left\lfloor \frac{y - l_y}{h_y - l_y} \right\rfloor \cdot (N-1)$ are the smallest x and y indices of control points in the neighborhood; and $B_k(u)$ is the k^{th}, and $B_l(v)$ is the lth basis function of Cubic B-splines.

Based on the linear precision property of B-splines, the initial coordinates of the pixel $\mathbf{x} = \sum_{k=0}^{3}\sum_{l=0}^{3} B_k(u)B_l(v)P^0_{i+k,j+l}$. Thus the displacement of pixel $\mathbf{x}$, given FFD control lattice deformation from P^0 to P, is:

$$\delta L(\mathbf{x}) = L(\mathbf{x}) - \mathbf{x} = \sum_{k=0}^{3}\sum_{l=0}^{3} B_k(u)B_l(v)\delta P_{i+k,j+l}$$

Suppose we have a number of n feature correspondences $(\mathbf{x}_c, \mathbf{x}_c'); c = 1,\ldots,n$, where c is the index of a feature pair. $\mathbf{x}_c = (x_c, y_c)$ is the center location of a (region or corner) feature on the fixed image, and $\mathbf{x}_c' = (x_c', y_c')$ is the center location of the feature's correspondence on the globally-aligned moving image. Assuming that under the correct local deformations, $\mathbf{x}_c'$ equals the deformed location of $\mathbf{x}_c$ which is $L(\mathbf{x}_c) = \mathbf{x}_c + \delta L(\mathbf{x}_c)$, and considering that the process of finding the correspondences introduces errors in real applications, our approach of non-rigid registration aims to find a closed-form solution for the Ordinary Least-Squares problem: $\min_{\delta P} \sum_{c=1}^{n} \left\| \mathbf{x}_c' - L(\mathbf{x}_c) \right\|$. The solution can be derived by solving a system of linear equations expressing the model:

$$U=Sp$$

where U is the displacement matrix between the correspondence pairs:

$$U=\begin{pmatrix} x_1^{'}-x_1 & y_1^{'} & - & y_1 \\ x_2^{'}-x_2 & y_2^{'} & - & y_2 \\ \vdots & & \vdots & \\ x_n^{'}-x_n & y_n^{'} & - & y_n \end{pmatrix}_{n\times 2},$$

S is the Cubic B-spline basis matrix:

$$S=\begin{pmatrix} \cdots & \cdots & [b_{i_1+0,j_1+l}] & \cdots & [b_{i_1+1,j_1+l}] & \cdots & [b_{i_1+2,j_1+l}] & \cdots & [b_{i_1+3,j_1+l}] & \cdots\cdots \\ & \ddots & & \ddots & & \ddots & & \ddots & & \ddots \\ \cdots & [b_{i_c+0,j_c+l}] & \cdots & [b_{i_c+1,j_c+l}] & \cdots & [b_{i_c+2,j_c+l}] & & [b_{i_c+3,j_c+l}] & \cdots & \cdots\cdots \\ & \ddots & & \ddots & & \ddots & & \ddots & & \ddots \\ \cdots & \cdots\cdots & [b_{i_n+0,j_n+l}] & \cdots & [b_{i_n+1,j_n+l}] & \cdots & [b_{i_n+2,j_n+l}] & \cdots & [b_{i_n+3,j_n+l}] & \cdots\cdots \end{pmatrix},$$

And p consists of the FFD parameters in a matrix:

$$p=\begin{pmatrix} \delta P_{1,1}^{x} & \delta P_{1,1}^{y} \\ \delta P_{1,2}^{x} & \delta P_{1,2}^{y} \\ \vdots & \vdots \\ [\delta P_{i_c+0,j_c+l}^{x}] & [\delta P_{i_c+0,j_c+l}^{y}] \\ \vdots & \vdots \\ [\delta P_{i_c+1,j_c+l}^{x}] & [\delta P_{i_c+1,j_c+l}^{y}] \\ \vdots & \vdots \\ [\delta P_{i_c+2,j_c+l}^{x}] & [\delta P_{i_c+2,j_c+l}^{y}] \\ \vdots & \vdots \\ [\delta P_{i_c+3,j_c+l}^{x}] & [\delta P_{i_c+3,j_c+l}^{y}] \\ \vdots & \vdots \\ \delta P_{M,N}^{x} & \delta P_{M,N}^{y} \end{pmatrix}_{(M\times N)\times 2.}$$

In both the B-Spline basis matrix S and the FFD parameter matrix p , c refers to the index of a corresponding feature pair ($c=1,\ldots n$), and (i_c, j_c) are the smallest control-point indices, along x and y directions, among all the indices for the sixteen control points in the neighborhood of feature location $\mathbf{x}_c$. Also in matrix S, $[b_{i_c+k,j_c+l}]$ is the abbreviation for:

$$[b_{i_c+k,j_c+l}] = \begin{pmatrix} B_k(u_c)B_0(v_c) & B_k(u_c)B_1(v_c) & B_k(u_c)B_2(v_c) & B_k(u_c)B_3(v_c) \end{pmatrix}.$$

In the FFD parameter matrix p , $[\delta P_{i_c+k,j_c+l}]$ is the abbreviation for:

$$[\delta P_{i_c+k,j_c+l}] = \begin{pmatrix} \delta P_{i_c+k,j_c+0} \\ \delta P_{i_c+k,j_c+1} \\ \delta P_{i_c+k,j_c+2} \\ \delta P_{i_c+k,j_c+3} \end{pmatrix},$$

Note that the column indices of $[b_{i_c+k,j_c+l}]$ in matrix S are the same as the row indices of $[\delta P_{i_c+k,j_c+l}]$ in matrix p. Thus, based on this parametric FFD model, a closed form solution for the control lattice deformation can be solved efficiently utilizing Singular value Decomposition (SVD) as: $p = S^{+}U$, where S^{+} is the pseudo-inverse matrix of S.

<u>Experiments and Validation</u>

This new image registration approach is very efficient because it solves for both rigid and non-rigid transformations in closed form, avoiding expensive iterative optimizations in traditional intensity-based methods. With elaborate feature detection and correspondence finding mechanisms, the method is also fairly accurate. The resulting deformation field has been demonstrated to be globally smooth, and minimizes the distance between target features on the fixed image and their corresponding features on the moving image after deformation. Because FFD has local compact support, any outlier correspondence would only have very local influence on the deformation field. Figure 2 shows feature correspondences and the estimated FFD deformations for an exemplary image of the brain. In this example, the moving image is a phantom image generated by artificially deforming the regular FFD control lattice overlaid on the fixed image. Accordingly, the ground-truth deformed control lattice that originally generated the moving image is known. By comparing the deformation field recovered by the proposed registration method with the ground truth deformation, the accuracy of the registration method is known. Figure 2 (III) demonstrates such comparison of the estimated space deformation and ground truth deformation for the exemplary brain image.

Future Trends

It remains challenging for intensity-based registration methods that are in use today to (1) improve efficiency of the optimization process, and (2) register images in arbitrary poses and be less dependent on the initial pose (i.e. transformation) estimate that starts the iterative parameter estimation. Hybrid image registration methods that combine the merits of intensity-based and feature-based methods have good potential in addressing these challenges as well as in tackling difficulties such as partial matching and multi-modal registration. In one direction, user-provided or automatically-extracted geometric features are integrated as part of an intensity-based energy function [Shen et al. 2002, Hartkens et al. 2002, Keller et al. 2003]. This approach stems from the observation that pixels in an image should not be treated equally during registration, so by adding feature-based energy terms, those salient features such as edges or corners are given higher weight during optimization. Having both intensity and feature based terms in one energy function, however, raises the question on how to balance their respective contributions. In another direction, much similar to the method proposed in this chapter, an overall feature-based framework is followed where features are extracted, correspondences are pursued and then transformations are estimated according to correspondences. Improvements will come in the type of features to use and on how to exploit *geometric* and *knowledge-based prior* constraints to eliminate outlier correspondences [Huang et al. 2004]. The salient-region features in this chapter, for instance, allow for multi-modal intensity-based similarity measures to be applied to matching features. Scale-invariant Feature Transform (SIFT) [Lowe 2004] and other types of scale- or affine-invariant image patches [Mikolajczyk et al. 2005] are also good candidates for the choice of features. In terms of finding correspondences, SoftAssign [Chui et al. 2000], RANSAC [Fischler et al. 1981], Expectation-Maximization [Huang et al. 2004] are some examples of robust matching; it is important to explore how to design the optimal matching scheme for a particular type of feature. And given correspondences, there are mature techniques based on Ordinary or Weighted Least-Squares estimators [Can et al. 2002, Huang et al. 2006] for global rigid transformations. In this chapter we have shown how to estimate the non-rigid deformations using a closed-form FFD solution, and other non-rigid methods describe deformation estimation using Thin Plate Spline [Chui et al. 2000] or Radial Basis Function [Fornefett et al. 1999] models.

Conclusion

We presented a novel method for efficient and robust non-rigid image registration. Instead of traditional point or edge features, the method extracts scale-invariant salient region features whose interior intensities can be matched using multi-modal similarity measures such as Mutual Information. Geometric constraints are applied in an Expectation-Maximization framework to prune outlier feature matches. Given feature correspondences, both the global transformation and non-rigid local deformations can be estimated in closed form as solutions of well-defined Ordinary Least-Squares problems. This method represents a new trend of hybrid image registration, which combines the efficiency and better invariance properties of feature-based registration with the robustness and multi-modal capability of intensity-based registration.

References

[1] Can, A., Stewart, C.V., Roysam, B., & Tanenbaum, H.L. (2002). A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):347–364, 2002.

[2] Chui, H., & Rangarajan, A. (2000). A New Algorithm for Non-Rigid Point Matching. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, pages 44–51.

[3] Collignon, A., Maes, F., Vandermeulen, D., Suetens, P., & Marchal, G. (1995). Automated multimodality image registration using information theory. *Proc. of Information Processing in Medical Imaging*, pages 263–274.

[4] Deligianni, F., Chung, A.J., & Yang, G.Z. (2006). Nonrigid 2-D/3-D Registration for Patient Specific Bronchoscopy Simulation with Statistical Shape Modeling: Phantom Validation. *IEEE Transactions on Medical Imaging*, 25(11): 1462-1471.

[5] Fischler, M.A., & Bolles, R.C. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, 24: 381--395.

[6] Fornefett, M., Rohr, K., & Stiehl, H. (1999). Elastic Registration of Medical Images using Radial Basis Functions with Compact Support. *IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, pages 1402–1409.

[7] Goshtasby, A., Gage, S.H., & Bartholic, J.F. (1984). A Two-Stage Cross-Correlation Approach to Template Matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(3): 374-378.

[8] Harris, C., & Stephens, M.J. (1988). A combined corner and edge detector. *Proc. Of Alvey Vision Conference*, pages 147–152.

[9] Hartkens, T., Hill, D.L., Castellano-Smith, A.D., Hawkes, D.J., et al. (2002). Using points and surfaces to improve voxel-based nonrigid registration. *Proc. of International Conf. on Medical Imaging Computing and Computer-Assisted Intervention*, vol. 2, pages 565–572.

[10] Huang, X., Paragios, N., & Metaxas, D. (2006). Shape Registration in Implicit Spaces using Information Theory and Free Form Deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(8): 1303-1318.

[11] Huang, X., Sun, Y., Metaxas, D., Sauer, F., & Xu, C. (2004). Hybrid Image Registration based on Configural Matching of Scale-Invariant Salient Region Features. *Proc. of IEEE Workshop on Image and Video Registration*, pages 167-175.

[12] Kadir, T., & Brady, M. (2001). Saliency, scale and image description. *International Journal of Computer Vision*, 45(2):83-105.

[13] Keller, Y., & Averbuch, A. (2003). Implicit similarity: A new approach to multi-sensor image registration. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, pages 543–548.

[14] Lowe, D.G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision,* 60(2): 91-110.

[15] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., & Suetens, P. (1997). Multi-modality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198.

[16] Maurer, C.R., Maciunas, R.J., & Fitzpatrick, J.M. (1998). Registration of head CT images to physical space using a weighted combination of points and surfaces. *IEEE Transactions on Medical Imaging*, 17(5):753–761.

[17] Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., et al. (2005). Comparison of Affine Region Detectors, *International Journal of Computer Vision*, 65(2): 43-72.

[18] Paragios, N., Rousson, M., & Ramesh, V. (2003). Nonrigid Registration Using Distance Functions. *Computer Vision and Image Understanding*, 89:142–165.

[19] Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M., & Hawkes, D. (1999). Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images. *IEEE Transactions on Medical Imaging*, 8: 712–721.

[20] Shen, D., & Davatzikos, C. (2002). HAMMER: Hierarchical attribute matching mechanism for elastic registration. *IEEE Transactions on Medical Imaging*, 21(11):1421–1439.

[21] Thirion, J.-P. (1996). New feature points based on geometric invariants for 3D image registration. *International Journal of Computer Vision*, 18(2):121–137.

[22] Viola, P., & Wells, W. (1995). Aligment by Maximization of Mutual Information. *Proc. of IEEE International Conf. on Computer Vision*, pages 16–23.

<u>Terms and Definitions</u>

Image Registration – is the process of transforming different sets of image data into one coordinate system and establishing point-by-point correspondences between the images. Registration is necessary to compare or integrate data acquired at different times or from different measurements. Medical image registration often requires nonrigid (or elastic) registration to cope with elastic deformations of the body parts imaged. Nonrigid registration of medical images can also be used to register a patient's data to an anatomical atlas.

Transformation Model – A transformation model represents in mathematics terms a geometric mapping of one space (or coordinate system) onto another or onto itself. Transformation can cause some or all of the elements in an image or a shape to change its current location. Such transformations include global effects such as translation, rotation, scaling, and local effects such as nonrigid warping and various specialized deformations.

Closed-form Solution - In Mathematics, an equation or a system of equations is said to have a closed-form solution if, and only if, at least one solution can be expressed analytically in terms of a bounded number of certain *well-known* functions. Typically, these well-known functions are defined to be elementary functions built from a finite number of exponentials, logarithms, constants, one variable, and roots of equations through composition and combinations using the four elementary operations (+ - × ÷).

Least Squares - Also known as Ordinary Least Squares (OLS) analysis, is a method to determine the value of unknown quantities in a statistical model by minimizing the sum of the residuals (i.e. the difference between the predicted and observed values) squared. Many types of optimization problems can be expressed in a least squares form, by either minimizing energy or maximizing entropy. Weighted Least Squares (WLS) is similar to OLS, but instead of weighting all points (or values) equally, they are weighted such that points with a greater weight contribute more to parameter fitting.

Modality - The type of input in image acquisition. For instance, black-and-white, color and infrared are three different modalities for the acquisition of a natural image. In medical imaging, common modalities include Computed Tomography (CT), Magnetic Resonance (MR) Imaging, Positron Emission Tomography (PET), among others.

Figure Legends

Figure 1. Registration of a pair of brain images using salient region features and feature correspondences. (I.a) Original PD-weighted MR brain image. (I.b) Original T1-weighted MR brain image. (II.a) The fixed image *f*. (II.b) Salient region features on *f*. (II.c) The moving image *m*. (II.d) Salient region features on *m*. (III.a-b) Feature correspondences selected by the Expectation-Maximization algorithm for global registration. (III.c) The transformed moving image *t* that matches with *f*. (III.d) Edges from *t* superimposed on the fixed image *f*.

Figure 2. Non-rigid image registration and validation of the estimated deformation field. (I) Region and corner feature-center points on the fixed image (left), and corresponding feature-center points on the moving image (right). Corresponding points between the two images are noted by the example numeric labels. (II) Non-rigid registration result; fixed image (II.a), deformed moving image after registration which has point-by-point correspondence with the fixed image (II.b), and the original moving image (II.c). (III) Comparison of the estimated space deformation field (drawn in green) and the ground truth deformation (drawn in red).

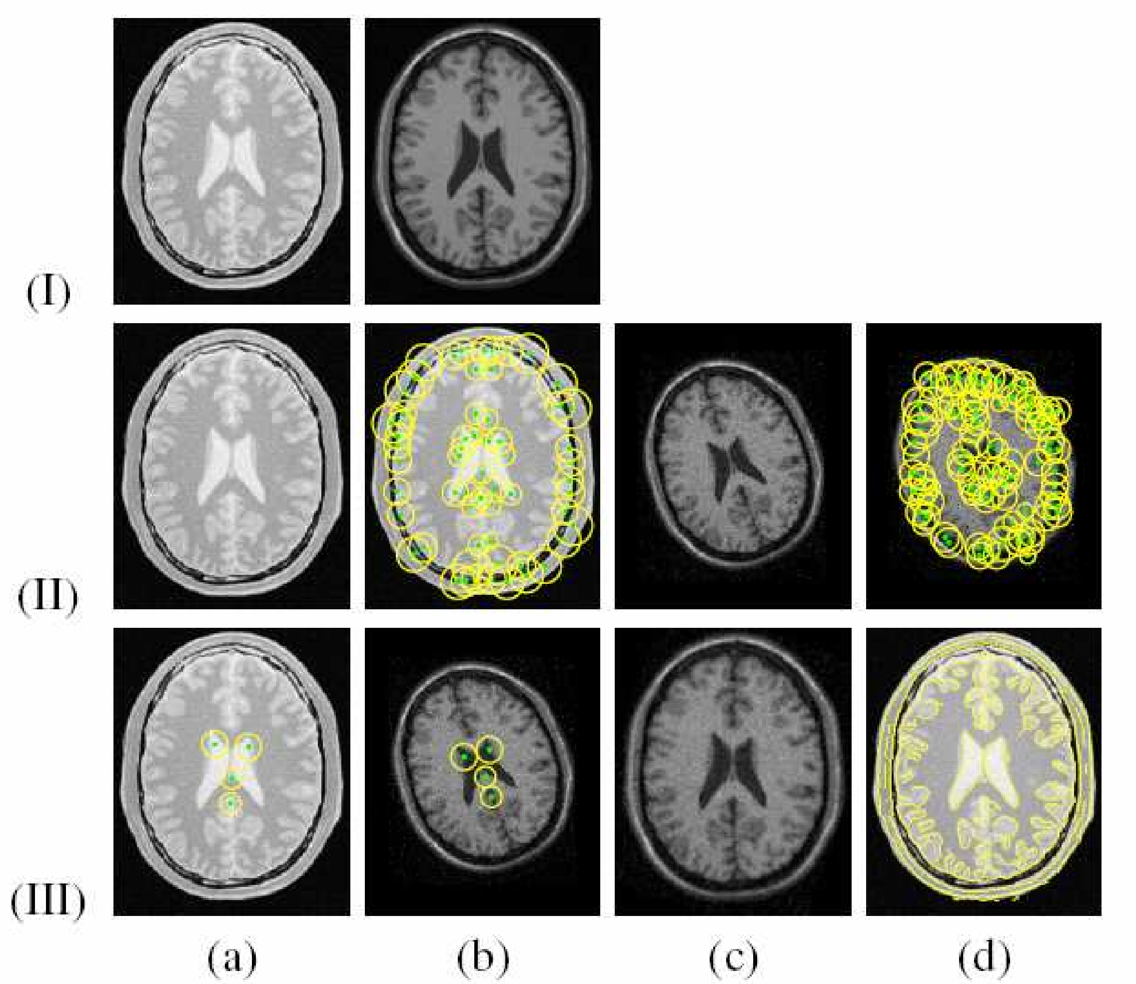

Figure 1

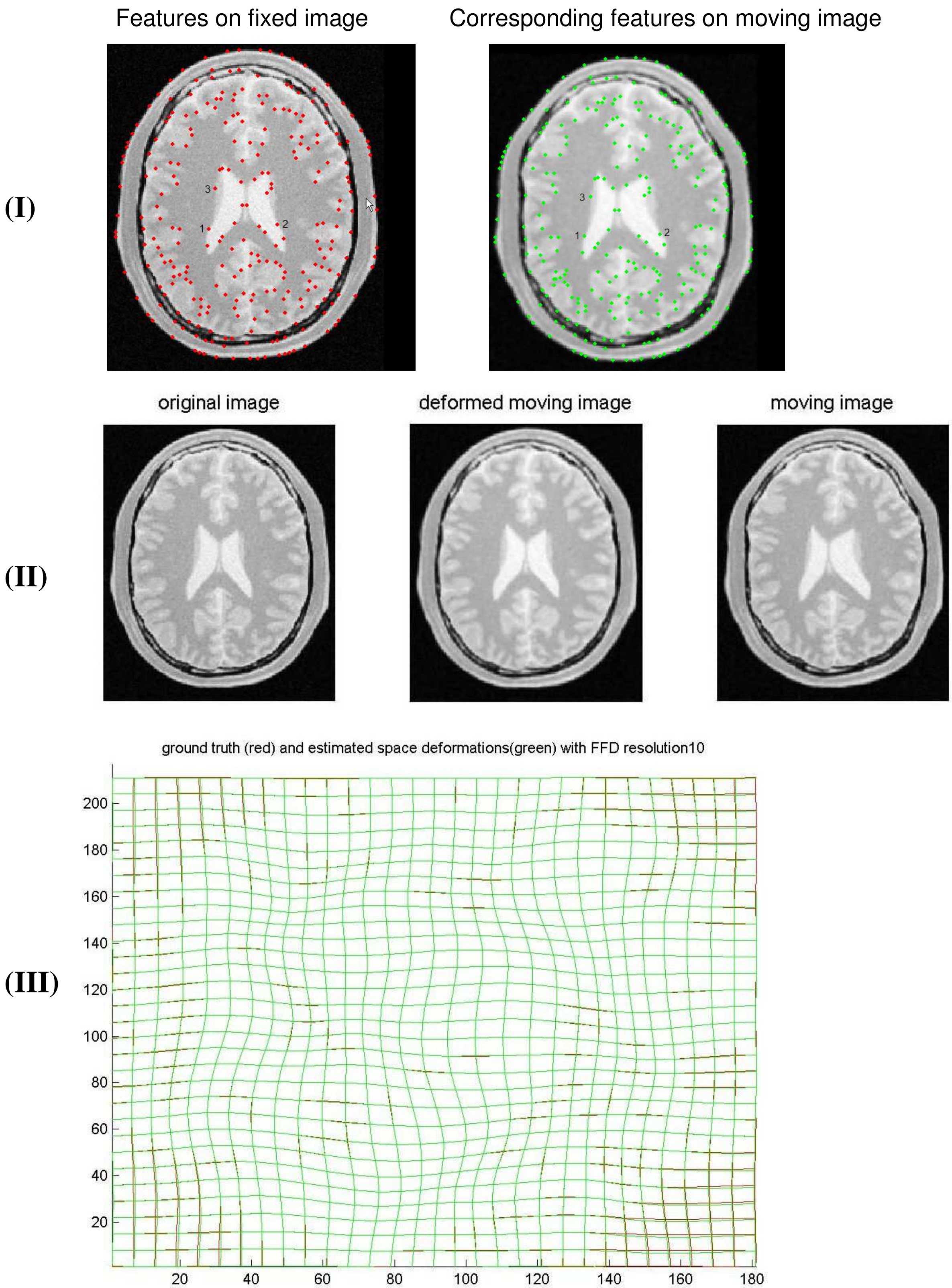
Features on fixed image
Corresponding features on moving image
(I)
3
1
2
original image
deformed moving image
moving image
(II)
ground truth (red) and estimated space deformations(green) with FFD resolution10
(III)
200
180
160
140
120
100
80
60
40
20
20
40
60
80
100
120
140
160
180

Figure 2