Registration Using Sparse Free-Form Deformations

Wenzhe Shi¹, Xiahai Zhuang², Luis Pizarro¹, Wenjia Bai¹, Haiyan Wang¹, Kai-Pin Tung¹, Philip Edwards¹, and Daniel Rueckert¹

Biomedical Image Analysis Group, Imperial College London, UK
Shanghai Advanced Research Institute, Chinese Academy of Sciences, China

Abstract. Non-rigid image registration using free-form deformations (FFD) is a widely used technique in medical image registration. The balance between robustness and accuracy is controlled by the control point grid spacing and the amount of regularization. In this paper, we revisit the classic FFD registration approach and propose a sparse representation for FFDs using the principles of compressed sensing. The sparse free-form deformation model (SFFD) can capture fine local details such as motion discontinuities without sacrificing robustness. We demonstrate the capabilities of the proposed framework to accurately estimate smooth as well as discontinuous deformations in 2D and 3D image sequences. Compared to the classic FFD approach, a significant increase in registration accuracy can be observed in natural images (61%) as well as in cardiac MR images (53%) with discontinuous motions.

1 Introduction

The classic free-form deformation registration model (FFD) [1] is widely used for medical image registration. Several improvements of the method have been proposed including different optimization strategies [2, 3]. Despite its popularity, little effort has been devoted to improve the accuracy of the formulation compared to other registration methods such as optical flow [4–7] and the Demons approach [8, 9].

In this paper, we address one main difficulty of the classic FFD approach, namely the conflict between the robustness of the registration and the ability to model discontinuous deformations. This conflict stems from the fact that the FFD uses a smooth B-spline basis to model the contribution of each control point to the deformation. To model global and smooth deformations a coarse control point spacing is typically used. To allow for very localized deformations a fine control point spacing is required, making the method less robust. A conventional approach to address this issue uses a coarse-to-fine approach in which the initial coarse control point mesh is successively subdivided [1].

The standard smoothness constraints for different registration methods [1, 4, 8] assume that the transformation within a neighbourhood changes gradually since it is caused by a smooth motion. Combining the implicit smoothness of

N. Ayache et al. (Eds.): MICCAI 2012, Part II, LNCS 7511, pp. 659–666, 2012.

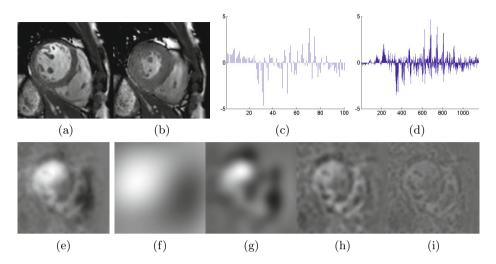


Fig. 1. (a) Target image; (b) source image; (c-d) coefficients of the B-spline basis for two levels (out of seven). Most coefficients are close to zero thus the sparsity assumption is valid; (e) deformation field induced by the FFD ϕ in x-direction; (f-i) FFD at four different levels.

the B-spline basis and explicit smoothness constraint in the regularization, the transformation model will produce smooth deformations.

As mentioned above, the control point grid spacing has a significant impact on the resulting transformation's ability to capture motion discontinuities robustly. Previous research focussed on the adaptive parametrization of the B-spline control point grid [10–12] driven by the intensity information in the images. An improved model should enable more control points to be placed in the area of the motion discontinuity.

Many other approaches to image registration have been proposed to overcome the conflict between robustness and discontinuity of the estimated motion in the field of optical flow [5–7]. Most recently, sparse coding methods have been proposed to evaluate the patch similarity between two images [13] and to constrain the transformation [14]. However there is no work ever reported to our best knowledge focused on motion discontinuity for the automatic FFD registration.

In this paper, we introduce a sparse representation for free-form deformations to estimate the transformation inspired by [13, 14]. This simple model uses the standard smoothness constraints and only imposes one assumption on the deformation, namely that the basis of the deformation is sparse in the parametric space. The assumption is generally true because the deformation between images is usually simpler than the actual images themselves. We use a multi-level FFD to represent the deformations in a parametric form. As can be seen from Figure 1, the sparsity assumption holds in general for multi-level FFDs with different control point spacings. Based on this assumption, we formulate the registration of two images using a sparse multi-level FFD representation of the control points.

We introduce a regularization term to impose smoothness at each level and a sparsity term to enforce coupled multi-level sparsity.

The novelty and contributions of this paper are the introduction of a sparsity model that avoids the *a priori* selection of an appropriate control point grid spacing. Furthermore, the approach reduces the conflict between global smoothness and local detail of the transformation by optimizing over the different FFD levels simultaneously with a sparsity constraint. These advantages allow for the robust estimation of deformation fields in the presence of discontinuous motion. We refer to this new approach as sparse free-form deformation model (SFFD). In the evaluation, we demonstrate that the proposed method can consistently capture localized motion with high accuracy.

2 Classic Free-Form Deformation Model

In the classic FFD registration [1], a non-rigid deformation $\phi = [X \ Y \ Z]^T$ is represented using a B-spline model in which the deformation is parameterized using a set of control points $\psi = [U \ V \ W]^T$ such that

$$\phi = \begin{bmatrix} B & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & B \end{bmatrix} \psi \,, \tag{1}$$

where B denotes the matrix of the B-spline basis functions. To find the optimal deformation between two images, the registration minimizes an energy functional E written as a function of ψ , which is typically a combination of two terms $E(\psi) := E_D(I_t, I_s \circ \phi) + E_R(\psi)$. The term E_D is a data constraint measuring the similarity between the target image I_t and the transformed source image $I_s \circ \phi$. The term E_R is a regularization constraint that enforces a smooth transformation. The energy function is typically minimized using gradient descent approaches [2] or discrete optimization approaches [15].

3 Sparse Free-Form Deformation Model

To be able to deal with large, global deformations and to improve the robustness, the classic FFD registration uses a multi-level approach: First, the optimal registration parameters are determined for a control point grid with large spacing. The grid is then successively subdivided to capture local deformations [1]. This requires an a-priori choice of the initial and final control point spacing. Furthermore, each level is optimized separately and once a level has been optimized it is no longer updated, leading to suboptimal registration results as can be seen in Figure 2. It was suggested in [14] that a realistic transformation can be easily embedded into a sparse representation. We postulate that an automatic selection of control points across different levels can be achieved by optimizing all FFD levels simultaneously while using a sparsity constraint.

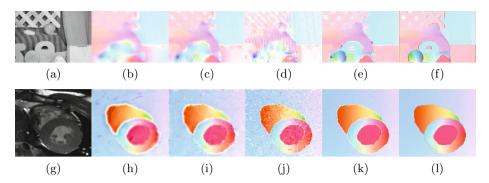


Fig. 2. Visual comparison between the classic FFD and the proposed SFFD using the colour scheme from [16]. The colour range corresponds to different direction and magnitude of the transformation. (a) and (g) show the rubber whale image from the Middlebury dataset and a frame from the cardiac B experiment, respectively;(f) and (l) display the ground truth transformations with noticeable motion discontinuities; (b-d) and (h-j) show the estimated motion with the classic FFD approach for 512/8mm, 256/4mm, and 64/1mm control point spacing; (e) and (k) exhibit the estimated motion with the proposed SFFD approach where $\lambda_S^N = 0.04$.

3.1 Sparse Free-Form Representation of Transformation

In this section, we propose to estimate the transformation ϕ with a sparse representation of the control points ψ . We use a m level FFD representation [10], $\psi = [U^1...U^m \ V^1...V^m \ W^1...W^m]^T$, as it is well suited for sparse representations. Accordingly, we utilize a multi-level B-splines basis $B = [B^1...B^m]$. The transformation ϕ is computed as in eq. (1) with the above redefinitions of ψ and B. As illustrated in Figure 1, a typical FFD is likely to be sparse in this representation.

Basis pursuit denoising [17] is a mathematical optimization problem that balances the trade-off between sparsity and reconstruction fidelity. In the context of image registration, the problem can be formulated as:

$$\arg\min_{\psi} \quad E(\psi) := ||I_t - I_s \circ \phi||_2^2 + ||\psi||_1,$$
 (2)

where the first term corresponds to the sum of squared differences (SSD) between the target and the transformed source. The second term enforces sparsity of the solution ψ . In general, an arbitrary (dis)similarity measure can be utilised in the data term $E_D(I_t, I_s \circ \phi)$, including information theoretic measures such as mutual information (MI) or its normalized counterparts (NMI) [18].

Following these principles, we propose a novel registration approach, namely the sparse free-form deformation (SFFD) model as

$$\arg \min_{\psi} \quad E(\psi) := E_D(I_t, I_s \circ \phi) + \lambda_R \sum_{i \in [1, m]} E_R(\psi^i) + \lambda_S ||\psi||_1, \qquad (3)$$

with constants $\lambda_R, \lambda_S \in \mathbb{R}^+$ weighting the regularization term and the sparsity term, respectively. Note that the regularization term imposes smoothness at each level of the multi-level FFD independently, while the sparsity term enforces coupled multi-level sparsity. This allows us to actively determine the importance of the control points across all levels in a joint manner, not independently as in the classic FFD framework. This strategy allows us to estimate deformations fields robustly and to preserve motion discontinuities, as it will be seen in the experimental validation. The SFFD model has the same computational complexity order as the classic FFD model.

We optimize eq. (3) using the interior point method of [19] that uses a log barrier function to make the sparsity term differentiable. The parameter λ_S is normalized between the data and the sparsity terms using the finite convergence to zero property. That is, for the L_1 -regularized least squares problem, convergence is achieved for a finite value λ_{max} of λ_S . The value of λ_{max} can be determined using eq. (10) in [19]. In our experiments we use:

$$\lambda_S^N = \lambda_S / \lambda_{max} \,. \tag{4}$$

For completeness, the reader will find the partial derivatives of the similarity measures with respect to ψ in [20] for the SSD and in [2] for the NMI.

4 Results

4.1 Datasets

In this work, we have evaluated the proposed SFFD against the classic FFD model on four different datasets. The datasets we have used for evaluation include the Middlebury benchmarking dataset, 2D cardiac MR images with synthetic smooth (cardiac A) and discontinuous motion (cardiac B), and 3D cardiac MR image sequences (cardiac C).

For basic benchmarking we have used six pairs of 2D greyscale natural images from the Middlebury benchmark's training dataset [16]. The Middlebury benchmark's dataset contains deformations with multiple independently moving rigid objects and background. For the dataset the ground truth deformation between each pair of images is available.

In addition, we have tested our approach using 2D and 3D cardiac MR images. For the 2D cardiac MR images a synthetically generated transformation has been applied to the images using sin function as proposed in [21]. A single sin function is used to generate a group of 10 data with smooth motion using different magnitudes and frequencies. Also, a set of 10 discontinuous motions is generated where multiple sin functions are fused into a discontinuous motion using segmentation information which can be seen in Figure 2l. The registration recovers the synthetic motion between the original image and the transformed image. For the above datasets, average error between the ground truth deformation and the deformation obtained after registration is measured.

We have also used 3D cardiac cine MR images from 13 normal volunteers to assess the proposed registration framework. The image resolution is $1.25 \times 1.25 \times 8mm$. We estimate the accuracy based on the tracking of myocardial boundaries of the left ventricle. Manual segmentation of endo- and epicardial surfaces are provided at both end-diastolic and end-systolic phases by a clinician. We then register the end-systolic (ES) phase directly to the end-diastolic (ED) phase. We evaluate surface distance between propagated myocardial surface and manual myocardial surface at end-systolic phase.

4.2 Implementation Details

During the optimization of the classic FFD, we use seven different levels of image resolution. The coarsest level has a size of around 64 voxels in each dimension. We use B-spline interpolation as evidence suggests that B-spline based interpolation is superior than linear interpolation [7, 2]. The coefficient for the smoothness penalty λ_R is set to 0.001 as in the original paper [1] for the NMI and 1 for the SSD due to the magnitude of the similarity metric.

One of most crucial parameters of the classic FFD registration is the control point grid spacing. We create a spacing at each image pyramid level by subdividing previous level's FFD. We have evaluated different initial control point spacings at the coarsest level varying from 512mm to 64mm where the final spacings at the final level varies from 8mm to 1mm. For the SFFD, we use a multi-level FFD with the coarsest level having a control point grid spacing of 64mm and finest level having a spacing of 1mm. The coarsest level should be reasonable according to data and the finest level should be around voxel size to embed a dense deformation. The number of image pyramid levels and the smoothness penalty are the same for both methods. Finally, NMI is used as similarity measures for the 3D cardiac MR images and SSD is used for the 2D cardiac MR images with simulated motion as well as for the Middlebury benchmark.

4.3 Evaluation

For the classic FFD, different spacing leads to significantly different results as showed in Table 1. Moreover, different datasets require different initial spacings to achieve best performances. It can be seen from the final column in Table 1 that the SFFD is robust against the choice of λ_S^N compare to the choice of spacing in FFD. There is little need to adjust λ_S^N across datasets to achieve near very good performance for individual data.

The median results using the multi-level FFD representation without sparsity constraint where $\lambda_S^N=0$ are 0.68mm,0.025mm,0.072mm and 1.74mm for Middlebury and cardiac A,B,C datasets respectively. Thus all datasets benefit from the sparsity constraint with a consistent increase in registration accuracy. Moreover, the SFFD exhibits a significant improvement against the best results from the classic FFD where discontinuous motion presents. The improvement is most significant against ground-truth from the Middlebury dataset and from the 2D cardiac MR image dataset with synthetic discontinuous motion. An increasing

Table 1. Median of accuracies and standard deviation of median from the classic FFD with different initial and final spacing and from the SFFD with different normalized sparsity constraint λ_S^N . The units of the accuracy metrics are in the bracket. For the results, n/a means that the image is not large enough to contain the initial control point grid spacing, bold font means the best result and * means statistically significantly different from the best result of the classic FFD using ttest with p-value 0.05. Final column is the median of the standard deviations of the results.

Initial/final spacing (FFD)	64/1 mm	128/2 mm	256/4 mm	512/8 mm	STD
Middleburry (mm)	0.69 ± 0.30	0.62 ± 0.27	$\boldsymbol{0.61 \pm 0.22}$	0.67 ± 0.33	0.108
Cardiac A (mm)	$.033 \pm .020*$	$.017 \pm .010*$	$.015 \pm .012$	$.012\pm.010$	0.014
Cardiac B (mm)	$.067 \pm .042*$	$.045\pm.016$	$.048 \pm .011$	$.059 \pm .024*$	0.034
Cardiac C (mm)	$\boldsymbol{1.68 \pm 0.28}$	$1.86 \pm 0.32*$	$1.94 \pm 0.40*$	n/a	0.619
$\lambda_S^N(SFFD)$	0.01	0.02	0.04	0.08	STD
Middleburry (mm)	$0.25 \pm 0.27*$	$0.25 \pm 0.28*$	$\textbf{0.24} \pm \textbf{0.27}^*$	$0.25 \pm 0.28*$	0.010
Cardiac A (mm)	$.010 \pm .006$	$.008 \pm .001$	$.007 \pm .001*$	$.006\pm.001^*$	0.001
Cardiac B (mm)	$.023 \pm .005*$	$.022 \pm .005*$	$.021 \pm .005*$	$.021\pm.005^*$	0.001
Cardiac C (mm)	1.57 ± 0.32	1.55 ± 0.31	$\boldsymbol{1.54 \pm 0.32}$	1.64 ± 0.32	0.091

ability to capture discontinuous motion while maintain robustness over smooth regions can be confirmed from a visual comparison in Figure 2.

Finally, in the 3D cardiac MR image sequences, we witness a limited improvement, $1.54 \pm 0.32 \ vs \ 1.68 \pm 0.28$, in the result. Due to the lack of an objective ground truth, we measured the errors only on the LV myocardial surfaces, and hence it can only partially demonstrate the registration accuracy.

5 Conclusion

In this paper, we have developed a sparse free-form deformation model for registration which addresses some most important short-comings of the original FFD registration model. Control points across different FFD levels has been optimized simultaneously using a sparse representation. Compared to the classic FFD, the SFFD requires less parameter tuning across different datasets. The user no longer needs to choose an appropriate control point spacing a-priori. Our experiments have shown a consistent improvement compared to the original FFD aproach. The most significant improvement can be observed in Middlebury dataset $(0.61 \pm 0.22~vs~0.24 \pm 0.27)$ and Cardiac B experiment $(0.045 \pm 0.016~vs~0.021 \pm 0.005)$ where the deformation field exhibits both smooth and discontinuous motion.

References

- Rueckert, 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 Transactions on Medical Imaging 18(8), 712–721 (1999)
- 2. Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S.: Fast free-form deformation using graphics processing units. Computer Methods and Programs in Biomedicine 98(3), 278–284 (2010)

- 3. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.: Elastix: a toolbox for intensity-based medical image registration. IEEE Transactions on Medical Imaging 29(1), 196–205 (2010)
- Horn, B.K.P., Schunck, B.G.: Determining optical flow. Artificial Intelligence 17(1-3), 185–203 (1981)
- 5. Mémin, E., Pérez, P.: Hierarchical estimation and segmentation of dense motion fields. International Journal of Computer Vision 46(2), 129–155 (2002)
- Roth, S., Black, M.J.: On the spatial statistics of optical flow. In: ICCV 2005, vol. 1, pp. 42–49 (2005)
- Sun, D., Roth, S., Black, M.J.: Secrets of optical flow estimation and their principles. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2432–2439. IEEE (2010)
- 8. Thirion, J.P.: Image matching as a diffusion process: an analogy with maxwell's demons. Medical Image Analysis 2(3), 243–260 (1998)
- Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Symmetric Log-Domain Diffeomorphic Registration: A Demons-Based Approach. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 754–761. Springer, Heidelberg (2008)
- Schnabel, J.A., Rueckert, D., Quist, M., Blackall, J.M., Castellano-Smith, A.D., Hartkens, T., Penney, G.P., Hall, W.A., Liu, H., Truwit, C.L., Gerritsen, F.A., Hill, D.L.G., Hawkes, D.J.: A Generic Framework for Non-rigid Registration Based on Non-uniform Multi-level Free-Form Deformations. In: Niessen, W.J., Viergever, M.A. (eds.) MICCAI 2001. LNCS, vol. 2208, pp. 573–581. Springer, Heidelberg (2001)
- Rohlfing, T., Maurer Jr., C.R.: Intensity-Based Non-rigid Registration Using Adaptive Multilevel Free-Form Deformation with an Incompressibility Constraint. In: Niessen, W.J., Viergever, M.A. (eds.) MICCAI 2001. LNCS, vol. 2208, pp. 111–119. Springer, Heidelberg (2001)
- 12. Hansen, M.S., Larsen, R., Glocker, B., Navab, N.: Adaptive parametrization of multivariate b-splines for image registration. In: CVPR, pp. 1–8 (2008)
- Roozgard, A., Barzigar, N., Cheng, S., Verma, P.: Dense image registration using sparse coding and belief propagation. In: International Conference on Signal Processing and Communication Systems, pp. 1–5 (2011)
- Shen, X., Wu, Y.: Sparsity model for robust optical flow estimation at motion discontinuities. In: CVPR, pp. 2456–2463 (2010)
- 15. Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N.: Dense image registration through mrfs and efficient linear programming. Medical Image Analysis 12(6), 731–741 (2008)
- Baker, S., Scharstein, D., Lewis, J.P., Roth, S., Black, M.J., Szeliski, R.: A database and evaluation methodology for optical flow. In: ICCV, pp. 1–8 (2007)
- Donoho, D.L., Huo, X.: Uncertainty principles and ideal atomic decomposition. IEEE Transactions on Information Theory 47(7), 2845–2862 (2001)
- Studholme, C., Hill, D., Hawkes, D., et al.: An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition 32(1), 71–86 (1999)
- Kim, S.J., Koh, K., Lustig, M., Boyd, S., Gorinevsky, D.: An interior-point method for large-scale 11-regularized least squares. IEEE Journal of Selected Topics in Signal Processing 1(4), 606–617 (2007)
- Kybic, J., Unser, M.: Fast parametric elastic image registration. IEEE Transactions on Image Processing 12(11), 1427–1442 (2003)
- Pizarro, L., Delpiano, J., Aljabar, P., Ruiz-del-Solar, J., Rueckert, D.: Towards dense motion estimation in light and electron microscopy. In: ISBI: From Nano to Macro, pp. 1939–1942 (2011)