Hierarchical vs. Simultaneous Multiresolution Strategies for Nonrigid Image Registration

Wei Sun[1], Wiro J. Niessen[1,2], and Stefan Klein[1]

[1] Biomedical Imaging Group Rotterdam,
Departments of Radiology and Medical Informatics, Erasmus MC, Rotterdam, The Netherlands
{w.sun,w.niessen,s.klein}@erasmusmc.nl
[2] Department of Image Science and Technology,
Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands

Abstract. Nonrigid image registration algorithms commonly employ multiresolution strategies, both for the image and the transformation model. Usually a hierarchical approach is chosen: the algorithm starts on a level with reduced complexity, e.g. a smoothed and downsampled version of the input images, and with a limited number of degrees of freedom for the transformation. Gradually the level of complexity is increased until the original, non-smoothed images are used, and the transformation model has the highest degrees of freedom. In this study, we define two alternative approaches in which low- and high-resolution levels are considered simultaneously. An extensive experimental comparison study is performed, evaluating all possible combinations of multiresolution schemes for image data and transformation model. Publicly available CT lung data, with annotated landmarks, are used to quantify registration accuracy. It is shown that simultaneous multiresolution strategies can lead to more accurate registration.

Keywords: Nonrigid Registration, Multiresolution, Hierarchical, Transformation, Scale Space.

1 Introduction

Nonrigid registration can be regarded as a large scale numerical optimization problem, which finds the optimal parameters for a selected transformation model to recover the deformation between images [1, 2, 3]. In practical registration tasks, local minima often exist in the optimization space. How to avoid these local traps, and reach the "correct" minimum, is a major challenge for registration algorithms. To tackle this issue, multiresolution strategies have become popular. Lester and Arridge [4] provided a comprehensive review on multiresolution strategies. They classified the multiresolution strategies into three groups: increasing data complexity, increasing warp complexity, and increasing model complexity. In most existing implementations of nonrigid registration algorithms, one or more of these multiresolution strategies are incorporated. Rueckert *et al.* [5] adopted both increasing data and warp complexities to implement a coarse-to-fine registration with free-form deformations (FFD).

B.M. Dawant et al. (Eds.): WBIR 2012, LNCS 7359, pp. 60–69, 2012.

For registration of lung data, Yin *et al.* [6] applied transformation models at two levels to one image resolution level. Gholipour *et al.* [7] presented several multiresolution approaches that can be used in brain data. Recently, Risser *et al.* [8] proposed a multiresolution strategy for large deformation diffeomorphic metric mapping. Besides the above mentioned works, hierarchical strategies have been widely used in many other registration tasks [9,10,11,12,13,14].

Most current multiresolution approaches adopt a step-by-step approach: the fine-scale registration will not be executed until all of the coarser registrations have been carried out. For example, a common strategy for FFD registration with B-splines is to combine the coarsest B-spline grid with the most blurred image at the beginning of optimization. After the optimization on this combination is complete, a denser B-spline grid and higher resolution image are used for further optimization. So, in these methods both transformation and data complexities are increased *hierarchically*.

Different from these hierarchical methods, several *simultaneous* multiresolution approaches were also presented previously. Stralen and Pluim [15] proposed a simultaneous multiresolution registration approach using a directed acyclic graph (DAG) and dynamic programming (DP). First, they constructed a DAG based on control points at different resolution levels. The DAG cost was defined as the sum of image dissimilarity at multiple scales and the difference between control point displacements in adjacent resolution levels. Then, they applied DP to find the optimum control point displacements. Somayajula *et al.* [16] also proposed a simultaneous multiresolution method for nonrigid registration. They defined corresponding scale-space feature vectors from multiresolution stacks of fixed and moving images at each voxel. Then, they used mutual information to align these feature vectors. In this way, different resolution levels were registered simultaneously, because the elements of each feature vector contained information from different resolution levels.

In this paper, we define three multiresolution concepts, named Hierarchical (H), Simultaneous (S) and Hierarchically Simultaneous (HS) respectively. These strategies can be implemented both for image data (D) and transformation model (T):

- Image data
  - DH = start with most blurred image, then less blurred image, and so on, until original image resolution;
  - DS = use the entire scale stack of different resolutions at once;
  - DHS = start with blurred image, then use blurred and less-blurred, and so on, until the entire scale stack is used.

- Transformation model
  - TH = start with coarsest B-spline grid, next level use finer B-spline grid, and so on, until the finest control point spacing;
  - TS = optimize coarse and fine B-spline deformations simultaneously;
  - THS = start with coarse B-spline, then add finer scale while still optimizing coarse scale, and so on, until all scales are being optimized simultaneously.

Combining these strategies gives 3×3 possibilities, which we implemented and compared in an experiment on publicly available CT lung data with manually annotated landmarks. In the following sections a detailed explanation of the proposed multiresolution strategies is given, followed by a description of the evaluation study.

2 Method

2.1 Multiresolution Strategies for Image Data

The *N*-dimensional moving and fixed images can be denoted by $M(\boldsymbol{x})$ and $F(\boldsymbol{y})$, where $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$ represent the image coordinates in M and F, respectively. Moving and fixed images on resolution level s can be generated by convolution of the original images with a Gaussian kernel:

$$\begin{cases} M(\boldsymbol{x}, s) = G(\sigma_s) * M(\boldsymbol{x}) \\ F(\boldsymbol{y}, s) = G(\sigma_s) * F(\boldsymbol{y}) \end{cases}, \tag{1}$$

where $G(\cdot)$ is the Gaussian kernel. σ_s is the variance of the Gaussian filter corresponding to resolution level s of the image. For a larger s, σ_s has a smaller value.

Nonrigid image registration is a process which aligns moving image M to fixed image F using a nonrigid transformation model. Mathematically, registration is formulated as an optimization problem, in which the nonrigid transformation $\boldsymbol{T}_{\hat{\mu}}$ is estimated by minimizing the difference C_{diff} between moving and fixed images:

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \left(C_{diff}\left(F, M \circ \boldsymbol{T}_{\boldsymbol{\mu}}\right)\right), \tag{2}$$

where $\boldsymbol{\mu}$ represents the parameters of the transformation $\boldsymbol{T}$. By making C_{diff} dependent on the resolution level we can introduce a multiresolution scheme for the image data. Below, we define the objective functions that correspond to DH, DS, and DHS:

$$\text{DH: } C_{diff}^{p} = C_{diff}\left(F(\boldsymbol{y}, p), M(\boldsymbol{x}, p) \circ \boldsymbol{T}_{\boldsymbol{\mu}}\right), \tag{3}$$

$$\text{DS: } C_{diff}^{p} = \sum_{s=1}^{S} C_{diff}\left(F(\boldsymbol{y}, s), M(\boldsymbol{x}, s) \circ \boldsymbol{T}_{\boldsymbol{\mu}}\right), \tag{4}$$

$$\text{DHS: } C_{diff}^{p} = \sum_{s=1}^{p} C_{diff}\left(F(\boldsymbol{y}, s), M(\boldsymbol{x}, s) \circ \boldsymbol{T}_{\boldsymbol{\mu}}\right), \tag{5}$$

where $p \in [1, S]$ is the current resolution level of the registration, and S denotes the number of resolution levels. Figure 1(a) provides an overview of these three multiresolution strategies, where $S = 3$. Note that the objective function for DS is actually independent of *p*, since all levels of the image scale stack are taken into account simultaneously.

2.2 Multiresolution Strategies for the Transformation

A classic FFD transformation model based on B-splines [5] can be defined as follows:

$$\boldsymbol{T}_{\boldsymbol{\mu}}(\boldsymbol{y}) = \boldsymbol{y} + \sum_{\boldsymbol{y}_i \in I_{\boldsymbol{y}}} \boldsymbol{c}_i \beta^r\left((\boldsymbol{y} - \boldsymbol{y}_i)/g\right), \tag{6}$$

where $\boldsymbol{y}_i$ is a control point of the B-spline grid, and $I_{\boldsymbol{y}}$ represents the set of control points within a compactly supported region of the B-spline at $\boldsymbol{y}$. $\boldsymbol{c}_i$ is the B-spline coefficient vector corresponding to control point $\boldsymbol{y}_i$, and the parameter vector $\boldsymbol{\mu}$ is formed by the elements of all $\boldsymbol{c}_i$. $\beta^r(\cdot)$ is the selected *r*th order multidimensional B-spline polynomial, and g is the spacing between grid points. By making the definition of the transformation model dependent on the resolution level p, we can define a multiresolution scheme for the transformation complexity. Below, we define the transformation models that correspond to TH, TS, and THS, in which we also introduce a dependence on s, in order to couple the image scale to the transformation complexity:

$$\text{TH: } \boldsymbol{T}_{\boldsymbol{\mu}}^{p}(\boldsymbol{y}, s) = \boldsymbol{y} + \sum_{\boldsymbol{y}_i \in I_{\boldsymbol{y}}^{p}} \boldsymbol{c}_i^{p} \beta^r\left((\boldsymbol{y} - \boldsymbol{y}_i)/g(p)\right), \tag{7}$$

$$\text{TS: } \boldsymbol{T}_{\boldsymbol{\mu}}^{p}(\boldsymbol{y}, s) = \boldsymbol{y} + \sum_{l=1}^{s} \sum_{\boldsymbol{y}_i \in I_{\boldsymbol{y}}^{l}} \boldsymbol{c}_i^{l} \beta^r\left((\boldsymbol{y} - \boldsymbol{y}_i)/g(l)\right), \tag{8}$$

$$\text{THS: } \boldsymbol{T}_{\boldsymbol{\mu}}^{p}(\boldsymbol{y}, s) = \boldsymbol{y} + \sum_{l=1}^{\min(s,p)} \sum_{\boldsymbol{y}_i \in I_{\boldsymbol{y}}^{l}} \boldsymbol{c}_i^{l} \beta^r\left((\boldsymbol{y} - \boldsymbol{y}_i)/g(l)\right), \tag{9}$$

where p is the current resolution level of the registration. $\boldsymbol{T}_{\boldsymbol{\mu}}^{p}(\boldsymbol{y}, s)$ represents the transformation at registration level p for a point $(\boldsymbol{y}, s)$ in the scale stack defined by (1). $l \in [1, S]$ denotes the B-spline grid level. $\boldsymbol{c}_i^{p}$ and $\boldsymbol{c}_i^{l}$ are the B-spline coefficient vectors at levels p and l, with corresponding grid spacing $g(p)$ and $g(l)$; the grid spacing $g(l)$ reduces with increasing *l*. Figure 1(b) provides an overview of these three multiresolution strategies for the transformation. With TH, the transformation is upsampled after each resolution (i.e., $\boldsymbol{c}_i^{p}$ are determined based on $\boldsymbol{c}_i^{p-1}$ such that $\boldsymbol{T}_{\boldsymbol{\mu}}^{p}(\boldsymbol{y}, s) = \boldsymbol{T}_{\hat{\boldsymbol{\mu}}}^{p-1}(\boldsymbol{y}, s)$ at the start of level p) and only the currently finest level is being optimized, so $\boldsymbol{\mu}$ at level p consists of the elements of $\boldsymbol{c}_i^{p}$. With TS, the transformation model is independent of p, and formed by a summation of multiple B-spline models with different grid spacings; the parameter vector $\boldsymbol{\mu}$ consists of all elements of $\boldsymbol{c}_i^{l}$, $\forall l \in [1, S]$. With THS, the model is similar to TS, but the finer B-spline models are only used in the later resolution levels; at resolution level p, the parameter vector $\boldsymbol{\mu}$ consists of all elements of $\boldsymbol{c}_i^{l}$, $\forall l \in [1, p]$.

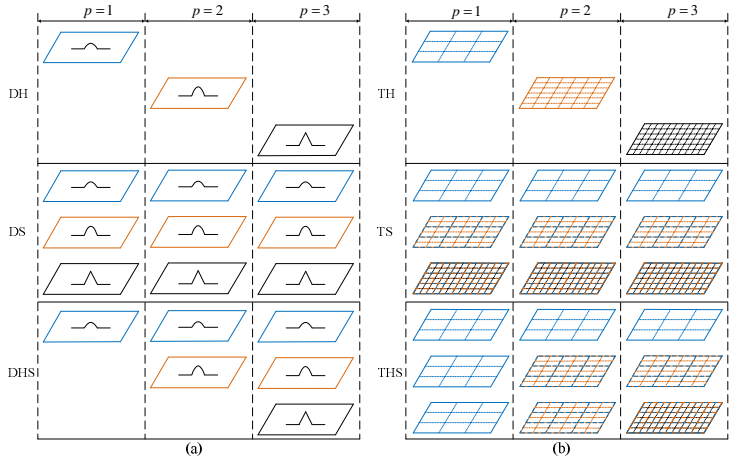

Fig. 1. Multiresolution strategies of data and transformation: (a) Most blurred, less blurred and original images are marked in blue, orange and black, respectively. (b) Coarsest, finer and finest B-spline grid are marked in blue, orange and black, respectively.

2.3 Combinations of Multiresolution Strategies

Because there are three different multiresolution strategies for both data and transformation, we can construct 3×3 combinations for multiresolution registration. Figure 2 presents all possible combinations in which registration processes have three multiresolution levels. In these combinations, the traditional multiresolution strategy is the combination of TH and DH. We impose the restriction that transformation level l (corresponding to the B-spline model with grid spacing $g(l)$) can only be applied to the finer image resolutions $s \in [l, S]$. According to this principle, the combination of TH and DHS becomes equivalent to the traditional multiresolution strategy TH-DH. In addition, TS-DH and TS-DHS are equivalent to THS-DH and THS-DHS, respectively.

2.4 Implementation Details

All experiments were performed with elastix [17], which is an open source package for registration. For C_{diff}, we used the common mean squared difference measure. Image intensities at non-grid positions were obtained by trilinear interpolation. Third order ($r = 3$) B-splines were adopted for the transformation model. The adaptive stochastic gradient descent optimizer (ASGD) [18] was selected as optimization method. In each iteration of ASGD, a small, randomly selected, subset of samples from the entire image is used. Downsampling the image is not necessary because the computation time is independent of the size of the image. To facilitate the optimization of combined B-spline levels of TS and THS, a diagonal preconditioning matrix $\boldsymbol{B}$ was defined to scale the parameters corresponding to the different transformation levels:

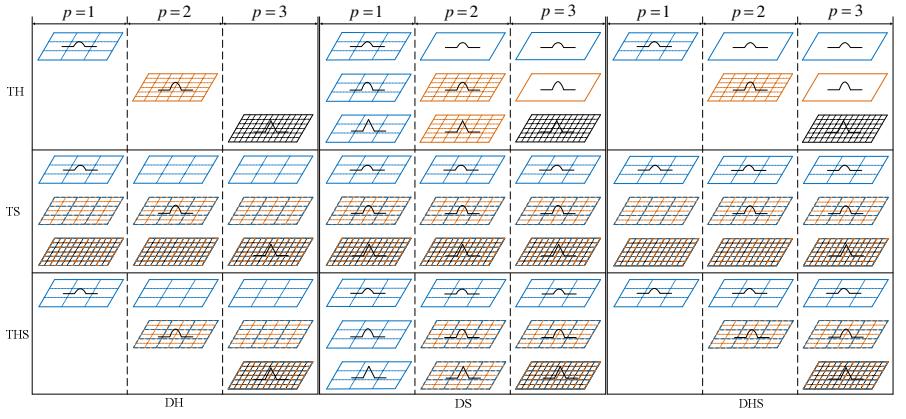

Fig. 2. Different combinations of multiresolution strategies of data and transformation for multiresolution registration

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k + \alpha_k \boldsymbol{B}\boldsymbol{d}\left(\boldsymbol{\mu}_k\right), \tag{10}$$

where k is current iteration number. $\boldsymbol{\mu}_{k+1}$ and $\boldsymbol{\mu}_k$ denote the new and current parameter vector, respectively. $\boldsymbol{d}\left(\boldsymbol{\mu}_k\right)$ is the derivative of the cost function with respect to $\boldsymbol{\mu}$. α_k is a scalar gain factor that determines the step size [18], and $\boldsymbol{B}$ is a diagonal matrix $diag($ $[b_{1,1}\, b_{1,1}\, b_{1,1}\, \ldots\, b_{l,l}\, b_{l,l}\, b_{l,l}\, \ldots\, b_{S,S}\, b_{S,S}\, b_{S,S}]$ $)$ with $b_{l,l} = \varepsilon^{-(S-l)}$. Based on initial trial-and-error experiments on one of the datasets (c1, described below), we set $\varepsilon = 4$. Here $\boldsymbol{B}$ works as a preconditioning strategy [19], which can enhance the convergence rate.

3 Experiments and Results

3.1 Experimental Data and Settings

To evaluate the performances of different multiresolution combinations, a set of lung data from DIR-lab [20] was used. Table 1 provides a description of these data.

Since manually marked landmarks have been provided in these data, mean of target registration error (mTRE) [21] can be used to evaluate the registration accuracy:

$$\mathrm{mTRE}\left(\boldsymbol{p}^{reg}, \boldsymbol{p}^{gold}\right) = \frac{1}{n}\sum_{i=1}^{n}\left\|\boldsymbol{p}_i^{reg} - \boldsymbol{p}_i^{gold}\right\|, \tag{11}$$

where $\boldsymbol{p}^{reg}$ and $\boldsymbol{p}^{gold}$ represent the registered and ground truth landmarks. $n = 300$ is the number of landmarks in all test cases.

Table 1. Description of experimental data

Case ID	Dimensions	Voxelsize (mm)	Landmarks	Initial mTRE(voxel)
c1	256 x 256 x 94	0.97 x 0.97 x 2.5	300	1.97
c2	256 x 256 x 112	1.16 x 1.16 x 2.5	300	2.12
c3	256 x 256 x 104	1.15 x 1.15 x 2.5	300	3.36
c4	256 x 256 x 99	1.13 x 1.13 x 2.5	300	4.42
c5	256 x 256 x 106	1.10 x 1.10 x 2.5	300	3.69

We selected the image data at exhale as moving image, and the image data at inhale as fixed image. In all test cases, $S = 4$ resolution levels were used. The image scale stacks were generated using $\{\sigma_1,...,\sigma_S\} = \{8, 4, 2, 1\}$. For the transformation, the coarsest grid spacing $g(1)$ was set to 64mm, isotropically. This value is a reasonable choice because it is almost one fourth of the image size. In the experiments the finest grid spacing $g(S)$ was set to 8mm, 10mm, 13mm, or 16mm. So the grid schedule for four transformation levels can be calculated as $\left\{g(1), g(S)\left(g(1)/g(S)\right)^{2/3}, \; g(S)\left(g(1)/g(S)\right)^{1/3}, g(S)\right\}$. For example, the grid schedule for $g(S) = 8\text{mm}$ is $\{64, 32, 16, 8\}$. For each iteration of optimization, the number of random samples was set to 16000 for all combinations. Note that with the DS and DHS approaches these 16000 samples are spread over multiple levels of the image scale stack, whereas with DH all 16000 samples are placed in the current active level $s = p$. The number of iterations was set to 2000 per resolution level.

3.2 Comparison of Different Multiresolution Strategies

The different multiresolution combinations are evaluated using five data pairs with four different finest grid spacings of the B-spline transformation. Figure 3 shows the registration results of all these combinations. As described in Section 2.3, TH-DHS, TS-DH and TS-DHS are actually equivalent to TH-DH, THS-DH and THS-DHS, respectively. So the results of TH-DH, THS-DH and THS-DHS are assigned to their equivalent combinations. In this way we can still make comparison among different multiresolution strategies of data and transformation in a general view. From Figure 3, it can be seen that the differences in most test cases are small, and THS generates better results than TH and TS in most of test cases. In addition, DS generates higher accuracy than DH and DHS. The traditional TH-DH approach has relatively worse performance in most cases. Especially in data pair c4, TH-DH results in unsatisfactory results. As shown in Table 1, c4 has larger average landmarks displacements than the other four data. So this significant deterioration could be caused by too large deformation of data.

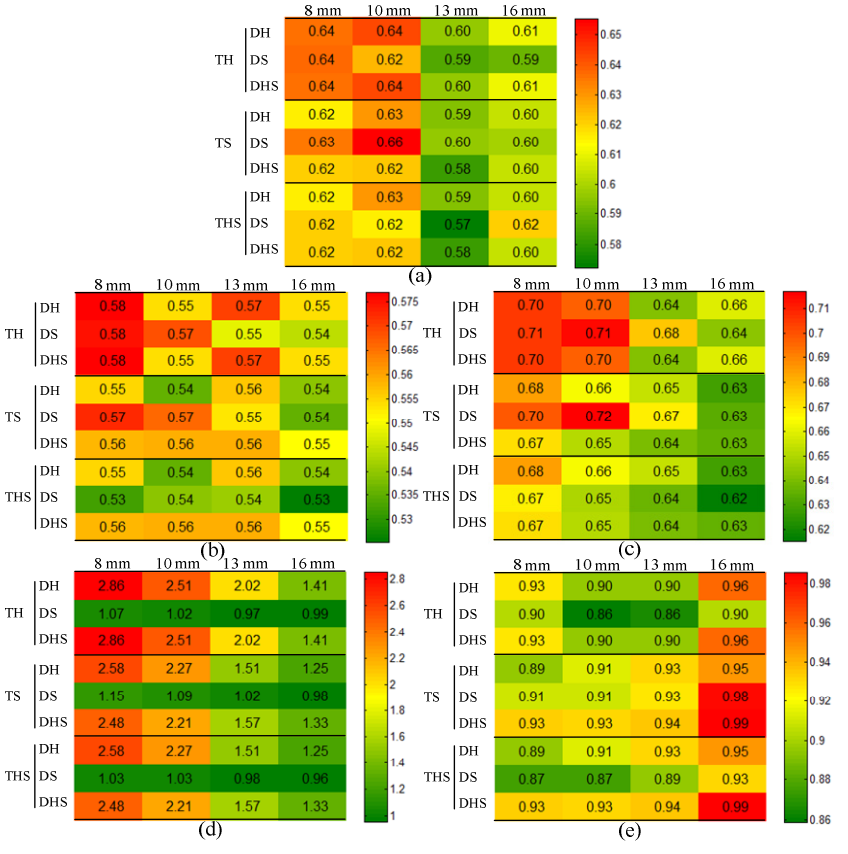

Fig. 3. Performance comparison of different multiresolution combinations. (a)-(e) are the results of lung data c1 to c5. The numbers represent the mTRE in voxels.

To make a further comparison among these combinations, a ranking of the 9 methods was made for each of the 5×4 test cases. The average rank of each method over all 5×4 test cases is presented in Figure 4. We can see that THS-DS has the best registration accuracy. The traditional TH-DH approach has the highest average rank number. It can also be noticed that the combinations with THS have lower rank than the other two multiresolution transformation strategies, when keeping the image resolution strategy the same.

4 Conclusions and Future Work

In this study different multiresolution strategies of data and transformation were compared on a publicly available lung CT dataset. Most observed differences among these combinations were small, and perhaps not statistically significant in this small number of datasets. However, some patterns could be observed. In current test cases, THS

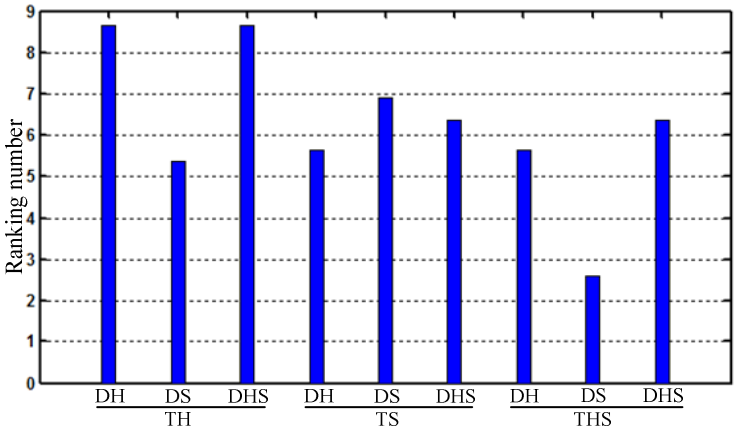

Fig. 4. Average rank of different multiresolution combinations in 20 test cases

performed better than the other multiresolution strategies for the transformation. Compared to DH and DHS, the combinations using DS had better registration results in most cases. The rank analysis indicated that the combination of THS and DS is indeed the best choice in this application. These results suggest that 1) keeping the low-dimensional B-spline transformation active while going to finer control point is advantageous, and 2) the simultaneous use of image data from multiple scales helps to improve registration quality. A limitation of the current work, is that we evaluated the different techniques within the context of one application. In future work, we plan to repeat this comparative evaluation on different data.

References

1. Zitová, B., Flusser, J.: Image registration methods: a survey. Image and Vision Computing 21, 977–1000 (2003)
2. Holden, M.: A review of geometric transformations for nonrigid body registration. IEEE Transactions on Medical Imaging 27, 111–128 (2008)
3. Fischer, B., Modersitzki, J.: Ill-posed medicine - an introduction to image registration. Inverse Problems 24 (2008)
4. Lester, H., Arridge, S.R.: A survey of hierarchical non-linear medical image registration. Pattern Recognition 32, 129–149 (1999)
5. 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, 712–721 (1999)
6. Yin, Y., Hoffman, E.A., Ding, K., Reinhardt, J.M., Lin, C.-L.: A cubic B-spline-based hybrid registration of lung CT images for a dynamic airway geometric model with large deformation. Physics in Medicine and Biology 56, 203–218 (2011)
7. Gholipour, A., Kehtarnavaz, N., Briggs, R., Devous, M., Gopinath, K.: Brain functional localization: A survey of image registration techniques. IEEE Transactions on Medical Imaging 26, 427–451 (2007)

8. Risser, L., Vialard, F., Wolz, R., Murgasova, M., Holm, D.D., Rueckert, D.: Simultaneous multi-scale registration using large deformation diffeomorphic metric mapping. IEEE Transactions on Medical Imaging 30, 1746–1759 (2011)
9. Cole-Rhodes, A.A., Johnson, K.L., LeMoigne, J., Zavorin, I.: Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. IEEE Transactions on Image Processing 12, 1495–1511 (2003)
10. Metz, C.T., Klein, S., Schaap, M., van Walsum, T., Niessen, W.J.: Nonrigid registration of dynamic medical imaging data using nD+t B-splines and a groupwise optimization approach. Medical Image Analysis 15, 238–249 (2011)
11. Mattes, D., Haynor, D.R., Vesselle, H., Lewellen, T.K., Eubank, W.: PET-CT image registration in the chest using free-form deformations. IEEE Transactions on Medical Imaging 22, 120–128 (2003)
12. Dinggang, S.: Fast image registration by hierarchical soft correspondence detection. Pattern Recognition 42, 954–961 (2009)
13. Kybic, J., Unser, M.: Fast parametric elastic image registration. IEEE Transactions on Image Processing 12, 1427–1442 (2003)
14. Musse, O., Heitz, F., Armspach, J.P.: Topology preserving deformable image matching using constrained hierarchical parametric models. IEEE Transactions on Image Processing 10, 1081–1093 (2001)
15. van Stralen, M., Pluim, J.P.W.: Optimal discrete multi-resolution deformable image registration. In: 6th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 947–950 (2009)
16. Somayajula, S., Joshi, A.A., Leahy, R.M.: Mutual information based non-rigid mouse registration using a scale-space approach. In: 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1147–1150 (2008)
17. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.W.: elastix: A toolbox for intensity-based medical image registration. IEEE Transactions on Medical Imaging 29, 196–205 (2010)
18. Klein, S., Pluim, J.P.W., Staring, M., Viergever, M.A.: Adaptive stochastic gradient descent optimisation for image registration. International Journal of Computer Vision 81, 227–239 (2009)
19. Bertsekas, D.P.: Nonlinear Programming. Athena Scientific, Massachusetts (1999)
20. Castillo, R., Castillo, E., Guerra, R., Johnson, V.E., McPhail, T., Garg, A.K., Guerrero, T.: A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. Physics in Medicine and Biology 54, 1849–1870 (2009)
21. van de Kraats, E.B., Penney, G.P., Tomazevic, D., van Walsum, T., Niessen, W.J.: Standardized evaluation methodology for 2-D-3-D registration. IEEE Transactions on Medical Imaging 24, 1177–1189 (2005)