ANZIAM J. 55(2014), 289–297
doi:10.1017/S1446181114000091

NEW DEVELOPMENT OF NONRIGID REGISTRATION

HSI-YUE HSIAO[1], CHIH-YAO HSIEH[1], XI CHEN[1], YONGYI GONG[2], XIAONAN LUO[3] and GUOJUN LIAO[✉1,3]

(Received 12 August, 2012; revised 7 January, 2014; first published online 5 June 2014)

Abstract

We propose a new nonrigid registration algorithm which is based on the optimal control approach. In our previously proposed methods, the Jacobian determinant and the curl vector were used as control functions. In this algorithm, we use a new set of control functions. A main advantage of using the new controls is that the positivity and normalization of the Jacobian determinant are satisfied automatically. Numerical results on large deformation brain images are provided to show the accuracy and efficiency of the algorithm.

2010 *Mathematics subject classification*: 65K10.

Keywords and phrases: nonrigid registration, optimal control, brain images.

1. Introduction

Image registration is a key technology for image analysis. It is the process of establishing a pixel (voxel in three dimensions) correspondence which maximizes a similarity measure of two similar images. Linear image registration uses affine transformation models to align two images globally by translation, rotation, and scaling. Nonrigid registration is needed to account for local, nonlinear variability of the images. In this paper we propose a new variational algorithm for nonrigid image registration. Variational techniques for nonrigid registration are classified into two categories:

(1) Unconstrained optimization of an energy functional, which consists of a data term and regularizing terms. The large deformation diffeomorphic metric mapping (LDDM) algorithm [3], and advanced neuroimaging tools (ANTs) [2], as well as the elastic and fluid methods, all belong to this category. The last

[1]Department of Mathematics, University of Texas at Arlington, Arlington, TX 76019, USA;
e-mail: hsiyue@gmail.com, josephsieh@gmail.com, chen.xi@mavs.uta.edu.
[2]School of Information Technology, Guangdong Foreign Language and International Business University, Guangzhou, China;
[3]National Engineering Research Center of Digital Life, School of Information Science & Technology, Sun Yat-sen University, Guangzhou, China.
© Australian Mathematical Society 2014, Serial-fee code 1446-1811/2014 $16.00

https://doi.org/10.1017/S1446181114000091 Published online by Cambridge University Press

two use physically motivated regularizing terms from linear elasticity and fluid mechanics. Other regularizing terms can also be used, including those penalizing the magnitude of displacement, the gradient of the displacement, and the Laplacian of the displacement, the divergence of the displacement, the curl of the displacement, etc. Auroux and Fehrenbach [1] tested seven commonly used regularizing terms on two hurricane-like images. This study confirms the concern that the regularizing terms prevent the data term from being optimally minimized.

(2) Constrained optimization of the data term only. Spline-based methods belong to this category [9, 17]. Regularity is provided by the spline model. They are not free-form deformation methods as claimed, because only the spline control points are moved somewhat freely, while all other points are interpolated by the spline formulas. In practice, the control points are at least eight pixels apart. Recently, we proposed an optimal control approach [4, 5, 7, 8, 12], which uses a differential system involving the Jacobian determinant and the curl vector as constraints. This approach generates smooth, invertible transformation for nonrigid registration, because any such transformation is generated from its Jacobian determinant and its curl vector.

A previous implementation of our approach [12] is based on the Lagrangian multiplier method, which adds extra terms to account for the constraints. The total energy functional is the sum of the data term and these (Lagrangian) terms. Since this implementation still uses an energy structure that is similar to that of methods in the first category, the accuracy is quite limited. A more severe issue is that the energy still contains arbitrary parameters. Lee and Gunzburger [10, 11] have investigated a finite element analysis of the Lagrangian multipliers method. The implementation by Chu et al. [5] is based on gradient decent, which does not contain any user-selected arbitrary parameters for regularization purpose, but the calculation of the variational gradient is not based on rigorous derivation. In this paper, we formulate a new algorithm with solid mathematical foundations, and apply it to a challenging brain image registration problem.

2. Description of the new algorithm

2.1. Optimal control approach Our optimal control approach is based on the deformation method of numerical grid generation for computational fluid dynamics [6, 13, 14, 16]. The deformation method generates a nonfolding grid with prescribed Jacobian determinant.

Let $f(x, y, z)$ be a positive function such that $\iint_\Omega (f - 1)\, dx = 0$. The following two steps will generate a grid transformation Φ with $J(\Phi) = f(\Phi)$. The first step is to generate a vector field $\mathbf{u}$ from

$$\begin{aligned} \operatorname{div} \mathbf{u} &= f - 1, \\ \operatorname{curl} \mathbf{u} &= \mathbf{g} \quad \text{on } \Omega, \\ \mathbf{u} &= \mathbf{0} \quad \text{on } \partial\Omega. \end{aligned} \tag{2.1}$$

https://doi.org/10.1017/S1446181114000091 Published online by Cambridge University Press

The second step is to form a grid node velocity field for s in [0,1]:

$$\mathbf{v} = \mathbf{u}/(s + (1 - s)f). \tag{2.2}$$

Then find the transformation $\Phi(x) = \phi(x, 1)$, where $\phi(x, s)$ is determined from the differential equation

$$\partial\phi(x, s)/\partial s = \mathbf{v}(\phi, s). \tag{2.3}$$

Liu [15] proved that the Jacobian determinant of transformation Φ equals f: $J(\Phi)(x) = f(\Phi(x))$.

The main idea of our optimal control approach to image registration is to find Φ by optimizing a similarity measure, subject to the above grid generation equations. Previously, we used f and $\mathbf{g}$ as the control functions. By adjusting f and $\mathbf{g}$, we use equations (2.1)–(2.3) to form the largest search space possible, which consists of all smooth and invertible transformations. However, the normalization requirement of f (that is, $\iint_\Omega (f - 1)\,dx = 0$) and the positivity requirement $f > 0$ are not easy to maintain during the optimization process. Moreover, the calculation of variational gradient is not based on a solid mathematical derivation. In this paper, we propose a new algorithm which uses a different set of control functions. The main advantage of the new algorithm is that both the positivity and the normalization requirement of f are automatically guaranteed.

2.2. New algorithm We now describe the new algorithm in two dimensions for simplicity. In this case, $\mathbf{g}$ is replaced by $g = \operatorname{curl}\mathbf{u}\cdot\mathbf{k}$, where $\mathbf{k}$ is the unit vector in the positive z-direction.

Given the template image T and the reference image R, we determine a registration transformation Φ by iteratively minimizing

$$\mathrm{SSD}(\Phi) = \frac{1}{2}\iint_\Omega \{T(\Phi(x)) - R(x)\}^2\,d\mathbf{x} \tag{2.4}$$

under constraints that are modified from equations (2.1)–(2.3).

The registration transformation Φ is updated from the current transformation by solving equation (2.2) with the one-step forward finite difference method. Namely,

$$\Phi(x) = \Phi_{\mathrm{old}}(x) + \frac{\mathbf{u}(x)}{1 + \operatorname{div}\mathbf{u}}\Delta t, \tag{2.5}$$

where we set $s = 0$ in (2.2) to get $\mathbf{v} = \mathbf{u}/f$ and then used $f = 1 + \operatorname{div}\mathbf{u}$ from (2.1). By differentiating equation (2.2) and then decoupling, we find that $\mathbf{u}$ satisfies the Poisson equations

$$\Delta\mathbf{u} = \mathbf{F} = (f_1, f_2), \quad \mathbf{u} = 0 \quad \text{on } \partial\Omega, \tag{2.6}$$

where f_1 and f_2 are the control functions defined by

$$f_1 = f_{x_1} - g_{x_2} \quad \text{and} \quad f_2 = f_{x_2} + g_{x_1}. \tag{2.7}$$

The image registration problem is formulated as follows: minimize $\mathrm{SSD}(\Phi)$ defined in equation (2.4) subject to the constraints in equations (2.5)–(2.7). The optimization is done by the method of gradient descent. A main technical achievement of this paper is the derivation of the gradient of SSD with respect to $\mathbf{F}$. This is done in the next section.

3. Derivation of the gradient

We now derive that

$$\partial\mathrm{SSD}/\partial\mathbf{F} = (\mathbf{a} + \nabla b)\Delta t.$$

Where vector $\mathbf{a}$ is determined by solving the Poisson equations

$$\begin{aligned} &\Delta\mathbf{a} = \mathbf{w} \quad \text{on } \Omega, \qquad \mathbf{a} = \mathbf{0} \quad \text{on } \partial\Omega, \\ \text{with} \quad &\mathbf{w} = \frac{T(\Phi(x)) - R(x)}{1 + \operatorname{div}\mathbf{u}}\nabla T(\Phi(x)), \end{aligned} \tag{3.1}$$

and b is determined by solving the Poisson equation

$$\begin{aligned} &\Delta b = h \quad \text{on } \Omega, \qquad b = 0 \quad \text{on } \partial\Omega, \\ \text{with} \quad &h = \frac{T(\Phi(x)) - R(x)}{(1 + \operatorname{div}\mathbf{u})^2}\nabla T \cdot \mathbf{u}. \end{aligned} \tag{3.2}$$

Let $\mathbf{F} = (f_1, f_2)$, and $\delta\mathbf{F} = (\delta f_1, \delta f_2)$ be the variations that vanish at the boundary. Then $\Delta\delta\mathbf{u} = \delta\mathbf{F} = (\delta f_1, \delta f_2)$ from equation (2.6). So we have

$$\delta\mathrm{SSD} = \iint_\Omega (T - R)\delta T \, d\mathbf{x} = \iint_\Omega (T - R)\nabla T \cdot \delta\boldsymbol{\Phi} \, d\mathbf{x}.$$

Since

$$\delta\boldsymbol{\Phi} = \delta\Big(\frac{\mathbf{u}}{1 + \operatorname{div}\mathbf{u}}\Big)\Delta t = \frac{\delta\mathbf{u}}{1 + \operatorname{div}\mathbf{u}}\Delta t - \frac{\mathbf{u}(\mathbf{div}(\delta\mathbf{u}))}{(1 + \operatorname{div}\mathbf{u})^2}\Delta t,$$

we get

$$\delta\mathrm{SSD} = \iint_\Omega \frac{(T - R)\nabla T \cdot \delta\mathbf{u}}{1 + \operatorname{div}\mathbf{u}}\Delta t \, d\mathbf{x} - \iint_\Omega \frac{(T - R)(\nabla T \cdot \mathbf{u})}{(1 + \operatorname{div}\mathbf{u})^2}\operatorname{div}(\delta\mathbf{u})\Delta t \, d\mathbf{x}.$$

Let $\mathbf{a}$ and b be the solutions to the Poisson equations (3.1) and (3.2), respectively. Then

$$\begin{aligned} \frac{1}{\Delta t}\delta\mathrm{SSD} &= \iint_\Omega \mathbf{w} \cdot \delta\mathbf{u} d\mathbf{x} - \iint_\Omega h \operatorname{div}(\delta\mathbf{u}) \, d\mathbf{x} \\ &= \iint_\Omega (\Delta\mathbf{a} \cdot \delta\mathbf{u} - \Delta b \operatorname{div}(\delta\mathbf{u})) \, d\mathbf{x} \\ &= \iint_\Omega (\Delta\mathbf{a} \cdot \delta\mathbf{u} - b \, \Delta(\operatorname{div}(\delta\mathbf{u}))) \, d\mathbf{x} \\ &= \iint_\Omega (\Delta\mathbf{a} \cdot \delta\mathbf{u} - b \operatorname{div}(\Delta\delta\mathbf{u})) \, d\mathbf{x} \\ &= \iint_\Omega (\mathbf{a} \cdot \Delta\delta\mathbf{u} + \nabla b \cdot \Delta\delta\mathbf{u}) \, d\mathbf{x} \\ &= \iint_\Omega ((\mathbf{a} + \nabla b)\delta\mathbf{F}) \, d\mathbf{x}. \end{aligned} \tag{3.3}$$

Hence, $\partial\mathrm{SSD}/\partial\mathbf{F} = (\mathbf{a} + \nabla b)\Delta t$.

https://doi.org/10.1017/S1446181114000091 Published online by Cambridge University Press

4. Implementation of the algorithm

We now describe an implementation of the algorithm in the following steps. The parameter Δt in (2.5) is not used explicitly in the algorithm. Instead, it is combined with the step size Δs in the gradient descent method into a single parameter tstep, namely tstep $= \Delta t \Delta s$. An initial tstep is chosen according to a preset number of iterations or by experience. The tstep at each iteration is automatically adjusted: if SSD improves, it is increased by a factor 1.1; if SSD fails to decrease, then tstep is reduced by a factor 0.9. A maximum value and a minimum value for tstep are set, based on experience.

- Step 0:
  Input the reference image R and the template image T.

- Step 1:
  Set $f_1 = f_2 = 0$; $\mathbf{a} = 0$; $b = 0$; $\Phi_{\mathrm{old}}(x) = x$.
  Calculate $\mathbf{a}, b$ from the Poisson equations (3.1) and (3.2), respectively, by fast Fourier transform (FFT).

- Step 2:
  Form the gradient $\partial \mathrm{SSD}/\partial \mathbf{F} = (\mathbf{a} + \nabla b)\Delta t$.

- Step 3:
  Update F by $F = F_{\mathrm{old}} + (\partial \mathrm{SSD}/\partial \mathbf{F})\Delta s = F_{\mathrm{old}} + (\mathbf{a} + \nabla b)\Delta t \Delta s = F_{\mathrm{old}} + (\mathbf{a} + \nabla b)$ tstep.

- Step 4:
  Determine $\mathbf{u}$ from the Poisson equations $\Delta \mathbf{u} = F$ by FFT.

- Step 5:
  Form $f = 1 + \operatorname{div} \mathbf{u}$. Check if f is positive, and maintain $f > 0$ by reducing tstep if necessary. A smaller tstep will reduce the difference between F and F_{old}, and thus maintain positivity of f. In practice, f is very close to 1 (>0.95) at each iteration step, and therefore no reduction of tstep is necessary.

- Step 6:
  Determine Φ from $\Phi(x) = \Phi_{\mathrm{old}}(x) + \mathbf{u}(x)/(1 + \operatorname{div} \mathbf{u})$ tstep.

- Step 7:
  Resample T by evaluating T at $\Phi(x)$.

- Step 8:
  Calculate SSD using the resampled T. If SSD is reduced, increase tstep by the factor 1.1 and go back to Step 1 for next iteration; if SSD fails to decrease, reduce

https://doi.org/10.1017/S1446181114000091 Published online by Cambridge University Press

tstep by the factor 0.9 and go back to Step 3 and continue until SSD decreases or tstep reaches its preset minimum value, which is 10^{-4} in the experiment. The maximum value for tstep is set to be 0.1.

- Step 9:
  Stop if SSD is reduced to a preset percentage tolerance, or if a preset minimum value of tstep is achieved, or if a preset number of iterations is reached.

Note that $f = 1 + \operatorname{div} \mathbf{u}$ after $\mathbf{u}$ is updated. By automatically reducing tstep if necessary, we can reduce $\mathbf{u}$ to ensure that $f > 0$. In the experiment, f is indeed very close to 1 (>0.95) at each iteration step without reducing tstep. Moreover, from $f - 1 = \operatorname{div} \mathbf{u}$, and since $\mathbf{u} = 0$ on $\partial\Omega$, it follows from the divergence theorem that $\iint_\Omega (f - 1)\, d\mathbf{x} = 0$.

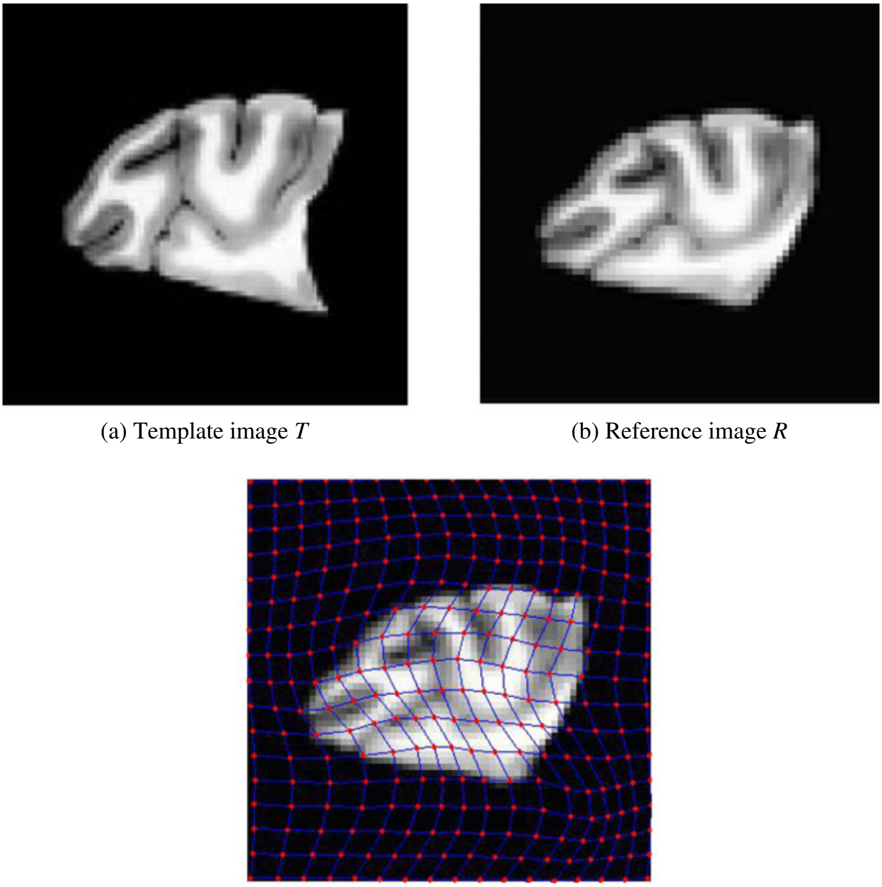
5. Brain image experiment

We now describe an experiment on a pair of brain images with large deformations. A Matlab program for the algorithm based on gradient decent is developed, which is used to repeat the well-known Experiment 5 of Beg et al. [3]. For more details of the images to be registered see the original description of the experiment [3].

The purpose of the experiment is to register image T in Figure 1(a) to image R in Figure 1(b). Note that image T has a sharp right-hand end, while image R has a round end. This large deformation is why the experiment is widely used to test nonrigid registration algorithms. In the Matlab program, the proposed algorithm is implemented in a multi-resolution strategy and is applied on the image pair of 64×64 pixels, which is padded with zeros to make images of 66×66 pixels. We may start on a coarse 8×8 grid by sampling the images at the nodes. We then pad the resized 8×8 images to images of 10×10 pixels by adding zeros on the four sides. Now we carry out the above algorithm on the 10×10 images with zero boundary condition on the Poisson equations by FFT. The calculated transformation is interpolated to the next level of 16×16 as the starting transformation. This process continues on a 32×32 grid and stops after the algorithm is completed on the finest grid of 64×64 for the full images. Since the original 64×64 images have been padded (with zeros) to make 66×66 images, we solve the Poisson equations for $\mathbf{u}$ with the zero boundary conditions on the 66×66 grid. In this experiment, we do not need multi-resolution. We directly register the images in the finest resolution. Our results, shown in Figure 1(c), are the resampled template image of 64×64 pixels. All of these pixels are free to deform under the registration transformation. That is why there are slight movements of the nodes on the four boundary portions on Figure 1(c). This is the intended result of padding the original images in order to use all information in the images in the registration process.

The new program produced data for Figure 1(c) in 17.8 seconds on a Dell Optix 780 PC with accurate results: the SSD value on the two images of 64×64 pixels was reduced from 2292.97 to 19.88 after 2639 iterations, a reduction by more than 99.1%. Visual comparison indicated that the resampled T in Figure 1(c) is much closer to R

https://doi.org/10.1017/S1446181114000091 Published online by Cambridge University Press



(a) Template image T

(b) Reference image R

(c) Resampled image

FIGURE 1. Experiment data.

than the results in the original study [3], especially at the lower right-hand end. The main part of computational cost of the algorithm was solving four Poisson equations on each iteration step. We used the FFT method for the Poisson equations. It is expected that the algorithm can be made more efficient by use of graphics processing units. This will be investigated in future work.

6. Conclusion

We describe a new method of nonrigid image registration. It uses a new set of control functions which automatically ensures the positivity and the normalization requirements. The derivation of the variational gradient is based on solid mathematical foundations. Numerical results on a pair of brain images with large deformations are presented, which confirm that the algorithm is capable of accurately registering

https://doi.org/10.1017/S1446181114000091 Published online by Cambridge University Press

images with large variability. The algorithm can be extended to and tested in three dimensions. More experiments and thorough studies of its efficiency are needed to assess its potential for clinical use.

Acknowledgements

The corresponding author acknowledges support from the National Science Foundation of the United States and the NSFC-Guangdong Joint Fund (U1135003, U1201252). The opinions expressed in this study are the authors' own.

References

[1] D. Auroux and J. Fehrenbach, "Identification of velocity fields for geophysical fluids from a sequence of images", INRIA, Rapport de recherche RR-6675, 2008.

[2] B. Avants, C. Epstein, M. Grossman and J. Gee, "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain", *Medical Image Analysis* **12** (2008) 26–41; doi:10.1016/j.media.2007.06.004.

[3] M. F. Beg, M. I. Miller, A. Trouve and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms", *Int. J. Comput. Vis.* **61** (2005) 139–157; doi:10.1023/B:VISI.0000043755.93987.aa.

[4] H.-M. C. Chen, C.-Y. Hsieh and G. Liao, "Nonrigid image registration using adaptive grid generation: Preliminary results", in: *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (2007), 580–583; doi:10.1109/ISBI.2007.356918.

[5] M.-Y. Chu, H.-M. Chen, C.-Y. Hsieh, T.-H. Lin, H.-Y. Hsiao, G. Liao and Q. Peng, "Adaptive grid generation based nonrigid image registration using mutual information for breast MRI", *J. Signal Proc. Syst.* **54** (2009) 45–63; doi:10.1007/s11265-008-0193-7.

[6] M. Grajewski, M. Köster and S. Turek, "Numerical analysis and implementational aspects of a new multilevel grid deformation method", *Appl. Numer. Math.* **60** (2010) 767–781; doi:10.1016/j.apnum.2010.03.017.

[7] H.-Y. Hsiao, H.-M. Chen, T.-H. Lin, C.-Y. Hsieh, M.-Y. Chu, G. Liao and H. Zhong, "A new parametric nonrigid image registration method based on Helmholtz theorem", in: *Medical Imaging 2008: Image Processing*, Volume 6914 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* (eds J. M. Reinhardt and J. P. W. Pluim), (2008).

[8] C.-Y. Hsieh, H.-M. Chen, T.-H. Lin, H.-Y. Hsiao, M.-Y. Chu, G. Liao and H. Zhong, "On the development of a new nonrigid image registration using deformation based grid generation", in: *Medical Imaging 2008: Image Processing*, Volume 6914 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* (eds J. M. Reinhardt and J. P. W. Pluim), (2008); doi:10.1117/12.769811.

[9] J. Kybic and M. Unser, "Fast parametric elastic image registration", *IEEE Trans. Image Proc.* **12** (2003) 1427–1442; doi:10.1109/TIP.2003.813139.

[10] E. Lee and M. Gunzburger, "An optimal control formulation of an image registration problem", *J. Math. Imaging Vision* **36** (2010) 69–80; doi:10.1007/s10851-009-0172-z.

[11] E. Lee and M. Gunzburger, "Analysis of finite element discretization of an optimal control formulation of the image registration problem", *SIAM J. Numer. Anal.* **49** (2011) 1321–1349; doi:10.1137/090767674.

[12] G. Liao, X. Cai, D. Fleitas, X. Luo, J. Wang and J. Xue, "Optimal control approach to data set alignment", *Appl. Math. Lett.* **21** (2008) 898–905; doi:10.1016/j.aml.2007.09.011.

[13] G. Liao, Z. Lei and G. de la Pena, "Adaptive grids for resolution enhancement", *Shock Waves* **12** (2002) 153–156; doi:10.1007/s00193-002-0149-y.

[14] G. Liao, F. Liu, G. C. de la Pena, D. Peng and S. Osher, "Level-set-based deformation methods for adaptive grids", *J. Comput. Phys.* **159** (2000) 103–122; doi:10.1006/jcph.2000.6432.

https://doi.org/10.1017/S1446181114000091 Published online by Cambridge University Press

[15] J. Liu, "New developments of the deformation method", Ph. D. Thesis, UT Arlington dissertation (2006).

[16] F. Liu, S. Ji and G. Liao, "An adaptive grid method and its application to steady Euler flow calculations", *SIAM J. Sci. Comput.* **20** (1999) 811–825; doi:10.1137/S1064827596305738.

[17] D. Rueckert, P. Aljabar, R. A. Heckemann, J. V. Hajnal and A. Hammers, *Diffeomorphic Registration Using B-Splines*, Volume 4191 of *Lect. Notes in Comput. Sci.* (Springer, Berlin-Heidelberg, 2006), 702–709; doi:10.1007/11866763_86.

https://doi.org/10.1017/S1446181114000091 Published online by Cambridge University Press