Non-rigid consistent registration of 2D image sequences

I Arganda-Carreras[1,5]**, C O S Sorzano**[1]**, P Thévenaz**[2]**,
A Muñoz-Barrutia**[3]**, J Kybic**[4]**, R Marabini**[1]**, J M Carazo**[1] **and
C Ortiz-de Solorzano**[3]

[1] Biocomputing Unit, National Centre for Biotechnology, CSIC, Darwin 3, Universidad Autónoma de Madrid, 28049 Madrid, Spain
[2] Biomedical Imaging Group, École polytechnique fédérale de Lausanne (EPFL), Switzerland
[3] Cancer Imaging Laboratory, Centre for Applied Medical Research, University of Navarra, Pamplona, Spain
[4] Center for Machine Perception, Czech Technical University, Prague, Czech Republic

E-mail: iarganda@cnb.csic.es

Abstract

We present a novel algorithm for the registration of 2D image sequences that combines the principles of multiresolution B-spline-based elastic registration and those of bidirectional consistent registration. In our method, consecutive triples of images are iteratively registered to gradually extend the information through the set of images of the entire sequence. The intermediate results are reused for the registration of the following triple. We choose to interpolate the images and model the deformation fields using B-spline multiresolution pyramids. Novel boundary conditions are introduced to better characterize the deformations at the boundaries. In the experimental section, we quantitatively show that our method recovers from barrel/pincushion and fish-eye deformations with subpixel error. Moreover, it is more robust against outliers—occasional strong noise and large rotations—than the state-of-the-art methods. Finally, we show that our method can be used to realign series of histological serial sections, which are often heavily distorted due to folding and tearing of the tissues.

1. Introduction

Registering two images consists in finding a deformation function that maps a given image *S* onto a target image *T*. Finding an optimal deformation function requires maximizing a

---

[5] Author to whom any correspondence should be addressed.

similarity criterion between the intensities of all pixels in both images and/or mapping selected reference—landmark—points. In an ideal situation, the deformation function should be unique and bijective. In other words, it should unequivocally link every pixel in the target image T with a pixel in the source image S. It should also have biological meaning appropriate for the particular image modality and source of misalignment. We refer to Zitová and Flusser (2003) for a detailed survey of basic registration methods.

The problem of registering pairs of images can be extended to registering sets of spatially, temporally or spectrally related images. The state-of-the-art approaches for this type of registration are as follows: (i) register all images to a reference image or (ii) pairwise register all the images starting from the first image of the sequence. In the first approach all the images in the group are registered to a reference image, which can be a selected image of the group (Marsland *et al* 2003, Malandain *et al* 2004), an average image (Bhatia *et al* 2004) or an iteratively calculated image model (Twining *et al* 2005). This is commonly used to locate an image in a normalized reference frame or coordinate system as part of the process of generating anatomical atlases (Bhatia *et al* 2004, Twining *et al* 2004). In the second approach, the images are registered as a sequence applying pairwise deformation functions. This is more appropriate for sequences of images that change in time, space, wavelength, etc. It has been used, for instance, for tracking cardiac motion in magnetic resonance images (MRI) (Radeva *et al* 1997, Huang *et al* 1999, Rueckert *et al* 1999, Ledesma-Carbayo *et al* 2005), registering blocks of histological tissue sections (Auer *et al* 2005, Wirtz *et al* 2005, Arganda-Carreras *et al* 2006) or registering multispectral fluorescence microscope images (Matula *et al* 2004).

Both approaches bias the registration by assigning special relevance to one of the images in the sequence. An alternative method consists of sequentially registering neighborhoods of images—i.e. subgroups of consecutive images—in order to increase the robustness of the registration and the smoothness of the deformations without assigning an arbitrary normative value to any of the images. For instance, Wirtz *et al* (2004) unidirectionally registered neighborhoods of three images. Yushkevich *et al* (2006) bidirectionally registered five-image neighborhoods, after choosing a reference image and removing what they considered bad slices in the sequence. Alternatively, Geng *et al* (2005) registered groups of three manifolds, i.e. contours or surfaces.

A key element of any registration algorithm is the method used to model the transformation between the source and target images. Some authors proposed using diffeomorphic deformation functions, which are invertible, differentiable and bijective (Avants *et al* 2006, Rueckert *et al* 2006, Cootes *et al* 2004, Grenander and Miller 1998). This method is fast, efficient and very appropriate when the transformation satisfies the diffeomorphic properties. However, this might not be the case for some large nonlinear transformations, such as the ones that can be found in series of manually processed histological sections, which are often heavily folded or torn. An alternative approach (Christensen and Johnson 2001, Thirion 1998) consists in jointly estimating the direct (S onto T) and reverse (T onto S) deformations and imposing as a constraint that one be as close as possible to the theoretical inverse of the other. This method in some applications seems to result in a more robust or accurate registration. Furthermore, while calculating the direct transformation g^{st} it provides a deformation field that is very similar to the ideal $(g^{st})^{-1}$.

Rogelj and Kovacic (2006) proposed an approach which does not impose consistency, but achieves reasonable one by establishing a symmetric image interdependence. The key idea of this method is to consider that images interact through forces. Information about the image transformation is thus gathered from both sides as the image similarity is measured in both registration directions. The experimental results seem to improve the correctness and the consistency over those of unidirectional methods and is computationally lighter than other

bidirectional methods since the computation of the inverse transformation is not required. Although the registration of image sequences is suggested as a possible extension, the method was only applied to the registration of image pairs.

In this work, we build upon and combine the concepts of neighborhood registration (Geng *et al* 2005) and consistency (Christensen and Johnson 2001) in the context of the registration of series of images (Csapo *et al* 2007). Namely, we use the general idea of transitivity from computational anatomy suggested by Grenander and Miller (1998), but we do not enforce the invertibility by definition of the anatomical transformations. Instead, we enforce it by the approach proposed by Christensen and Johnson (2001). This way, we can calculate and recover from complex non-invertible distortions, such as those commonly found in histological sections (e.g. tissue folding and tearing). Then, looking for a computationally reasonable compromise solution to the problem of registering image sequences, we use (Geng *et al* 2005) a tri-wise registration concept—originally developed to register triples of contours—apply it to triples of images and then extend it to register entire image sequences. The propagation of the information through the entire sequence is ensured by the overlap between the image triples and an iterative back-and-forth progression of the algorithm. To register each image triple—which is the building block of the registration of the entire sequence— we take advantage of the fact, proved by Csapo *et al* (2007), that the sequential composition of two transformations is equivalent to the transformation between the first and third member of the triple. This was used by Csapo *et al* (2007) to unidirectionally and non-consistently register sequences of images, and we here extend its use to the bidirectional and consistent registration of triples of images. In summary, we approach the problem of registering series of images by bidirectionally registering consecutive triples of images, to find a solution that minimizes the overall similarity and consistency errors while keeping acceptable computational speed. The consistency is ensured locally through the registration of groups of three images and globally by spreading the results backward and forward in the whole set of images.

In our implementation we estimate the deformation function by minimizing a cost functional made of two terms: a similarity term and a consistency term. We use an efficient Marquardt–Levenberg optimization method with Hessian estimation (Press *et al* 1992). We use B-splines to interpolate the images and model the deformation functions. B-splines are computationally light, differentiable, have good approximation properties and can be used to represent both linear and nonlinear transformations, providing close control of the level of details of the transformation. Finally, we use a multiresolution (iterative coarse-to-fine) implementation, which improves the convergence speed and robustness of the algorithm (Unser 1999). Although this method can be applied to many registration problems, we show in particular the benefit obtained by using it to register series of histological sections, which are often affected by strong deformations.

The structure of the paper is as follows. First we describe the previous work in section 2.1, and then we introduce the concept of consistent tri-wise registration in section 2.2. The two terms of our cost functional are described in sections 2.3 and 2.4. The image and deformation representations are presented in sections 2.5 and 2.6. Our optimization algorithm is then described in section 2.7 while the multiresolution implementation is presented in section 2.8 and the main outline of our consistent sequential registration (CSR) method is introduced in section 2.9. Next, we show experimental results, first with synthetic images in section 3.1, and also with the real sequences of biological images in section 3.2. We finish with conclusions in section 4.

2. Methods

Given three consecutive images of a sequence, our algorithm finds the deformation fields that relate them. This is done by minimizing an energy functional that has two terms: the similarity error $E_S^{T_i}$ between every triple T_i—local group of three consecutive images I_i, I_{i+1}, I_{i+2}—and the consistency error of the corresponding partial deformations $E_C^{T_i}$. The energy functional can be written as

$$E = \sum_{i=1}^{N_S-2} \left(w_s E_S^{T_i} + w_c E_C^{T_i} \right), \tag{1}$$

where N_S is the number of images in the sequence and w_s and w_c are the specific weights given to the similarity and consistency energy terms.

2.1. Pairwise registration

Our work builds on our previous work and the work of others. Namely, we develop on the idea of unidirectional pairwise registration (UPR) (Sorzano *et al* 2005, Kybic and Unser 2003) and of consistent pairwise registration (CPR) (Christensen and Johnson 2001, Arganda-Carreras *et al* 2006).

UPR can be formulated as finding the best deformation function $g(\mathbf{x}) : \mathbb{R}^2 \to \mathbb{R}^2$ that transforms a source image I_s into a target image I_t. Therefore, the deformed version of the source image, $I_s(g(\mathbf{x}))$, should closely resemble the target image $I_t(\mathbf{x})$.

CPR adds the idea of imposing a consistency constraint by simultaneously calculating the transformations g and h such that g maps I_s to I_t and h maps I_t to I_s subject to the constraint $g \simeq h^{-1}$. In CPR, the dimensionality of the nominal search space is twice as large as in UPR, but the consistency constraint defines a smaller embedded search space than the one corresponding to UPR. Consequently, it often allows for faster convergence in terms of the number of iterations when compared to a non-consistent registration, as will be shown in section 3.1.6.

2.2. Consistent tri-wise registration

We now extend the previous concepts—UPR and CPR—to the case of simultaneously registering three images. We call it consistent tri-wise registration (CTR). Let I_1, I_2 and I_3 be the images to be registered. We aim at finding the deformation functions

$$g^{ij}(\mathbf{x}) : \mathbb{R}^2 \to \mathbb{R}^2, \tag{2}$$

which map coordinates of I_i into coordinates of I_j. The explicit estimation of g^{13} and g^{31}, as described by Geng *et al* (2005), improves the consistency of the deformations at the expense of increasing the complexity of the optimizer. Instead, looking for a compromise between accuracy and computation time, we decided to use only pairwise transformations $(g^{12}, g^{23}, \ldots)$ to move through the sequence. This was originally proposed by Csapo *et al* (2007), who showed that the composition of the intermediate transformations g^{12} and g^{23} makes g^{13} and g^{31} implicitly consistent. Therefore, only g^{12}, g^{21}, g^{23} and g^{32} need to be calculated, since g^{13} and g^{31} are defined by composition, that is $g^{13} = g^{23} \circ g^{12}$ and $g^{31} = g^{21} \circ g^{32}$, i.e. $g^{13}(\mathbf{x}) = g^{23}(g^{12}(\mathbf{x}))$ and $g^{31}(\mathbf{x}) = g^{21}(g^{32}(\mathbf{x}))$. See figure 1 for a graphical description of the transformations involved.
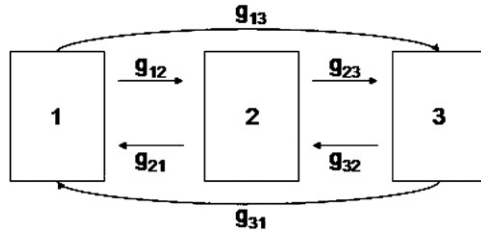
Figure 1. CTR scheme. The diagram shows a group of three images and their corresponding transformations. For every pair of images (i and j with $i, j = 1, 2, 3$ and $i \neq j$) there are two functions (g^{ij}, g^{ji}) that map pixel coordinates from image i to image j and vice versa.

2.3. Similarity error term

We calculated the similarity between the images as the sum of squared image differences. This is not a critical choice, and other similarity criteria can be used instead, such as mutual information (Pluim *et al* 2003, Skouson *et al* 2001), cross-correlation (Andronache *et al* 2006), etc.

The direct similarity error term E_S^{ij} between a source image I_i and a target image I_j is defined as

$$E_S^{ij} = \frac{1}{\#\Omega_{ij}} \sum_{\mathbf{x} \in \Omega_{ij}} (I_j(\mathbf{x}) - I_i(g^{ij}(\mathbf{x})))^2, \tag{3}$$

where Ω_{ij} defines a set of relevant pixels common to the images I_i and I_j that can be defined by masks, and where $\#\Omega_{ij}$ is the total number of pixels under the mask.

In the bidirectional case, the similarity error E_S between two images I_i and I_j, is defined as $E_S = E_S^{ij} + E_S^{ji}$. Finally, the similarity error for CTR, E_S, can be written as the sum of the direct, E_{DS}, and the inverse, E_{IS}, similarity errors: $E_S = E_{DS} + E_{IS}$ where $E_{DS} = E_S^{12} + E_S^{23} + E_S^{13}$ and $E_{IS} = E_S^{21} + E_S^{32} + E_S^{31}$.

Since this similarity measure is sensitive to linear transformations of the image gray values, we assume that all images use a common intensity range, e.g. 0–255.

2.4. Consistency term

The consistency error E_C^{ij} is the Euclidean distance between a point $\mathbf{x}$ of the source image I_i and the same point after a forward and backward transformation $g^{ji}(g^{ij}(\mathbf{x}))$. Therefore, the consistency error E_C between two images I_i and I_j is given by

$$E_C = E_C^{ij} + E_C^{ji}, \tag{4}$$

where

$$E_C^{ij} = \frac{1}{\#\Omega_{ij}} \sum_{\mathbf{x} \in \Omega_{ij}} \|\mathbf{x} - g^{ji}(g^{ij}(\mathbf{x}))\|^2. \tag{5}$$

The UPR energy functional does not have a consistency term.

We now extend the consistency term for groups of three images (CTR method). To this end, we do not need to consider the consistency of g^{13} and g^{31}, since they are defined by composition (see section 2.2) and their consistency is ensured by that of the intermediate deformations fields g^{12}, g^{23}, g^{21} and g^{32}. Therefore, the total consistency term is then given by $E_C^{12} + E_C^{23} + E_C^{21} + E_C^{32}$.

2.5. Image representation

The target image I_j is always evaluated at integer positions while the source image I_i has to be evaluated generally at non-integer coordinates, $g^{ij}(\mathbf{x})$. We opted for using cubic B-spline interpolation, since it offers a good trade-off between accuracy and speed (Unser *et al* 1991). Therefore, we represent the source image I_i as

$$I_i(x, y) = \sum_{k,l \in \mathbb{Z}^2} c_{k,l} \beta^3 \left(\frac{x}{h} - k \right) \beta^3 \left(\frac{y}{h} - l \right), \tag{6}$$

where β^3 is a cubic B-spline, $c_{k,l}$ are the B-spline coefficients and h is a parameter that controls the level of details of the representation.

2.6. Deformation representation

Taking advantage of the properties of the B-splines (Szeliski and Shum 1996) we also represent the deformation fields as a linear combination of B-splines:

$$g^{ij}(\mathbf{x}) = g^{ij}(x, y) = \left(g_1^{ij}(x, y) \ \ g_2^{ij}(x, y) \right)$$
$$= \sum_{k,l \in \mathbb{Z}^2} (c_{1,k,l}^{ij}, c_{2,k,l}^{ij}) \beta^3 \left(\frac{x}{s_x} - k \right) \beta^3 \left(\frac{y}{s_y} - l \right), \tag{7}$$

where s_x and s_y are the sampling steps that control the degree of detail of the deformation field.

2.7. Optimization

To optimize the energy functional, we use a variation of the robust Levenberg–Marquardt method (Thévenaz *et al* 1998). This method iteratively updates the deformation coefficients $\mathbf{c} = \mathbf{c}_{m,k,l}^{ij}$ so that $\mathbf{c}^{(n+1)} = \mathbf{c}^{(n)} + \Delta \mathbf{c}^{(n)}$, where $\Delta \mathbf{c}$ is the solution of $\tilde{H} \Delta \mathbf{c}^{(n+1)} = \nabla E(\mathbf{c}^{(n)})$, n is the iteration number; $\nabla E(\mathbf{c}^n)$ is the gradient of the energy with respect to the deformation coefficients evaluated at $\mathbf{c}^{(n)}$ and $\tilde{H}$ is a modified version of the Hessian matrix H with the diagonal components $\tilde{H}$ calculated as $[\tilde{H}]_{ii} = (1+\lambda)[H]_{ii}$. This algorithm achieves a gradual transition between quasi-Newton and gradient descent steps, controlled by the parameter λ. The parameter λ is adaptively modified according to the ability of $\mathbf{c}^{(n)}$ to minimize the energy functional. If the previous step succeeded to decrease the goal function, then the local model is considered appropriate and more weight is given to the second-order information provided by the Hessian. Consequently, the next iterative step of the optimization algorithm will be more Newton-like. If the previous iterative step did not decrease the criterion, the local model is considered inappropriate and the optimization algorithm will retry the last step, thus behaving more like a gradient descent. We compute a Broyden–Fletcher–Goldfarb–Shanno (BFGS) approximation of the Hessian (Press *et al* 1992) using the algorithm described by Sorzano *et al* (2005). The BFGS approximates the Hessian as a semi-definite matrix, thus assuring the stability of the quasi-Newton optimization algorithm. The Hessian estimated this way is initialized by a diagonal approximation and is updated only on successful estimates $\mathbf{c}^{(n)}$ and when the positive definite quality of the estimated is assured.

Our optimizer makes extensive use of the derivatives of the different energy terms with respect to the deformation coefficients. Since the images and the deformation fields are expressed in B-splines basis, these derivatives can be computed explicitly and efficiently (see appendices A.1 and A.2).
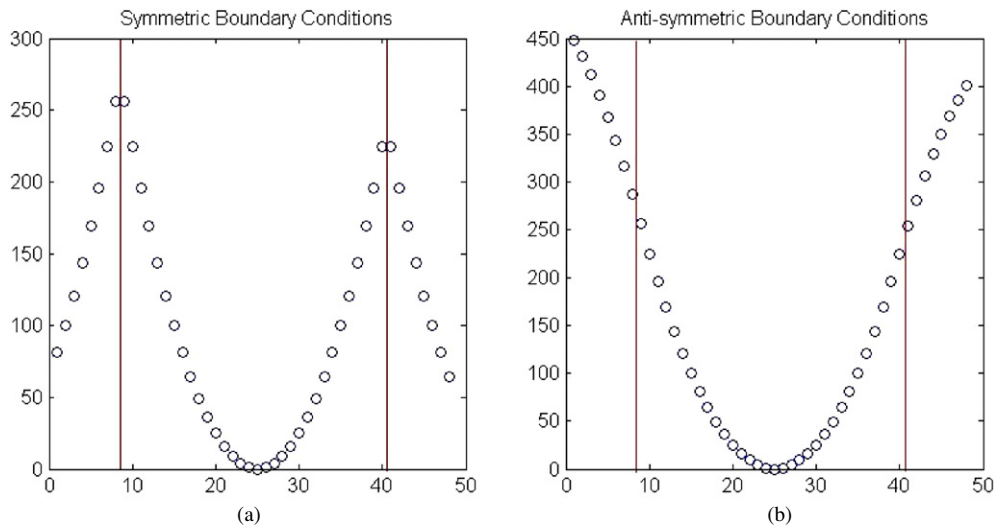
Figure 2. Example of symmetric (a) and anti-symmetric (b) boundary conditions for a quadratic function. The vertical lines represent the boundaries.

2.8. Multiresolution

All images and deformations are represented using multiresolution pyramids (Unser *et al* 1993). The number of B-spline coefficients at the pyramid level L is $2^L \times 2^L$ for $L = 0$, 1, 2, ... The algorithm starts using low-resolution versions of the images and deformations, finds the minimum of the energy functional and then moves to the next pyramid level, using the deformation obtained in the lower level as the starting point of the next—upper—level. This is repeated until the highest resolution level of the pyramid is reached.

To move between pyramid levels, customized reduction and expansion operators were used (Unser *et al* 1993) (see appendix A.3 for a technical explanation), resolved at the image and deformation borders using the appropriate boundary conditions. For the images we used standard symmetric boundary conditions. For the deformation fields, symmetric mirror conditions do not properly account for the deformation trend at the boundaries, and thus we applied anti-symmetric boundary conditions (figure 2). See appendix B for a practical justification of the choice of boundary conditions.

2.9. Consistent sequential registration

We have further extended the idea of CTR to do CSR. To this end, at each resolution level, triples of images are iteratively CTR-registered forward and backward—with a two-image overlap with respect to the previous iteration—until either the desired degree of accuracy or a maximum number of iterations is reached. Registering backward and forward iteratively refines the intermediate results incorporating contributions of all the members of the sequence to the registration process. See algorithm 1 for a formal description.

Algorithm 1. Consistent Sequential Registration CSR method

Require: $S \leftarrow$ A sequence of images I_i, $i = 1, \ldots, N_S$

 1: Initial transformations: $g_{i,i+1}$, $g_{i+1,i} \leftarrow$ Identity $\forall i \in [1, N_S - 2]$
 2: Auxiliary transformations: g_{12}, g_{21}, g_{23}, $g_{32} \leftarrow$ Identity
 3: **for all** pyramid level **do**
 4:     **while** not (convergence **or** max number of iterations) **do**
 5:         **for** $i := 0$ to $N_S - 2$ **do** {Forward Registration}
 6:             $g_{12} \leftarrow g_{i+1,i+2}, g_{21} \leftarrow g_{i+2,i+1}, g_{23} \leftarrow g_{i+2,i+3}, g_{32} \leftarrow g_{i+3,i+2}$
 7:             Do tri-wise registration of I_i, I_{i+1}, I_{i+2} with initial transformations g_{12}, g_{21}, g_{23}, g_{32}
 8:             Save $g_{i,i+1}$, $g_{i+1,i}$, $g_{i+1,i+2}$, $g_{i+2,i+1}$
 9:         **end for**
10:         **for** $i := N_S - 3$ down to 0 **do** {Backward Registration}
11:             $g_{12} \leftarrow g_{i+1,i+2}, g_{21} \leftarrow g_{i+2,i+1}, g_{23} \leftarrow g_{i+2,i+3}, g_{32} \leftarrow g_{i+3,i+2}$
12:             Do tri-wise registration of I_i, I_{i+1}, I_{i+2} with initial transformations g_{12}, g_{21}, g_{23}, g_{32}
13:             Save $g_{i,i+1}$, $g_{i+1,i}$, $g_{i+1,i+2}$, $g_{i+2,i+1}$
14:         **end for**
15:     **end while**
16: **end for**

---

3. Experimental results

3.1. Validation of the method using synthetic images

3.1.1. Selection of consistency and similarity weights.　　Since the terms of the energy function have different units, the choice of the weights plays an important role on the algorithm. The similarity term represents the mean squared error (MSE) of pixel intensities (0–255 for our 8-bit test images). Moreover, the consistency term relates to the mean geometrical distance (in pixels) between the forward and backward transformations. In our experiments, the similarity error usually lies around the hundreds, while the consistency error is in the order of the tens. To study the sensitivity of the registration results to the values of the weights, we performed the following experiment (see figure 3): we used the first three images from a sequence of consecutive coronal slices of *Macaca fascicularis* stained with cholinesterase. The images were grayscale, 300×282 pixel size. We then fixed the similarity weight to 1.0 and observed the evolution of both errors when registering the triplet using gradually increasing values of the consistency weight. For low consistency values—up to 10.0—both errors decrease at a similar rate. Then, the consistency error starts to decrease while the similarity error increases, indicating that a too strong consistency constraint was used. This experiment led us to use 1.0 similarity and around 10.0 consistency weights in all the registration of all sequences described in the rest of the paper.

3.1.2. Accuracy.　　We first tested the performance of our CSR algorithm using synthetic image sets. To this end, we applied 20 known deformations g^* to an image and then performed the registration of the image and all its warped versions. Figure 4 shows the image used, a standard histology section of human breast epithelium stained with hematoxylin and eosin (H&E), captured at $40\times$ magnification.
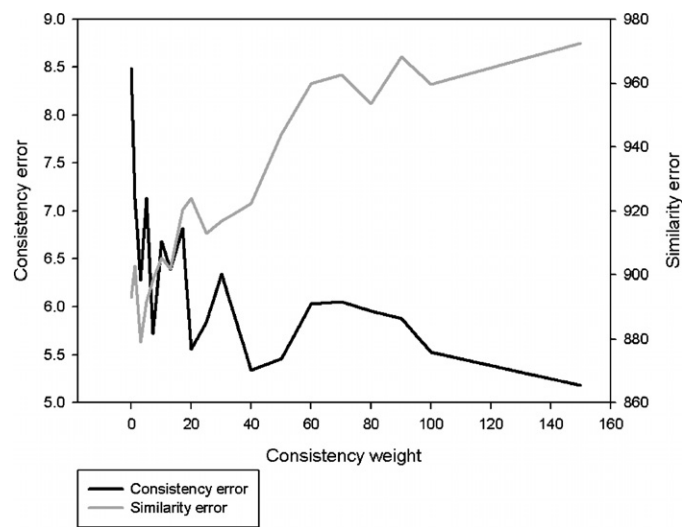
Figure 3. Evolution of the similarity and consistency error with increasing values of the consistency weight. These error values correspond to the registration of the first image triple used in section 3.1.3.
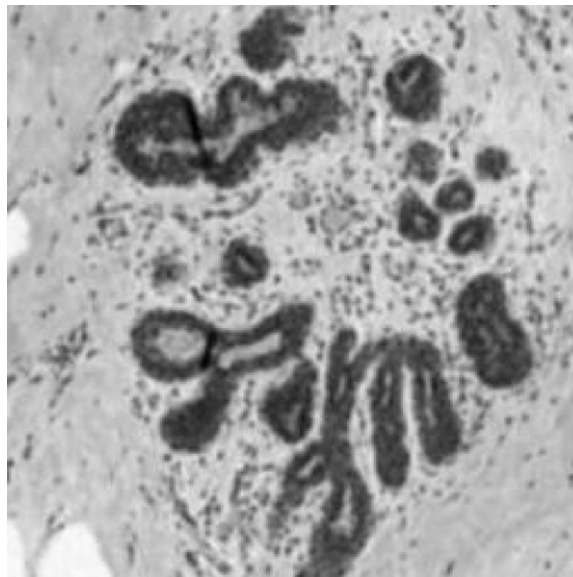

Figure 4. H&E stained section of human breast epithelium.

To deform the images, we used the barrel/pincushion distortion, which is characteristic of image acquisition and displaying devices such as cameras and monitors. Note that this choice does not favor the use of B-splines, since these distortions cannot be exactly represented in a B-spline space. Zero-mean Gaussian noise with standard deviation 0.05 to k_1 and k_2 was added to the images. To register the images we used a four-level deformation pyramid and
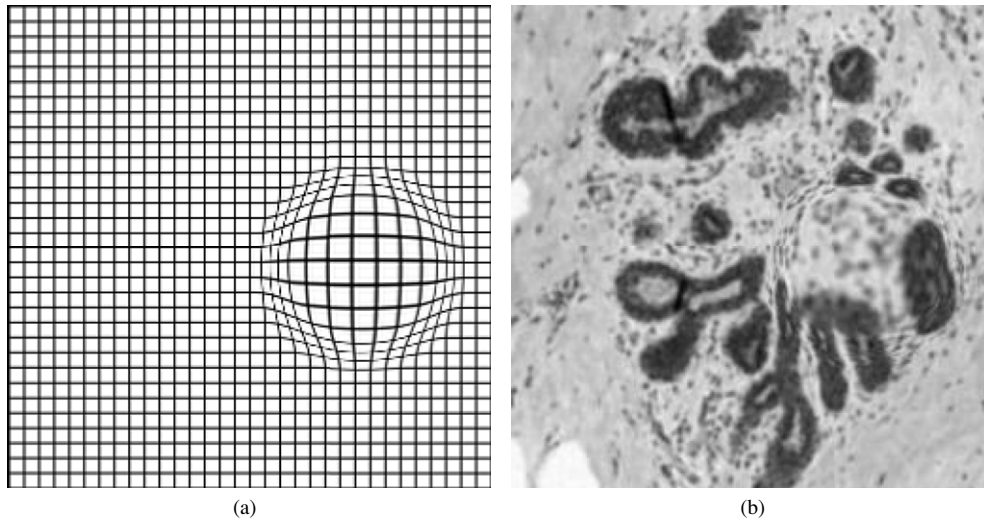
(a)   (b)

Figure 5. Example of fish-eye transformation on (a) a grid image and (b) the histological section.

a five-level image pyramid, i.e. a grid of 8×8 B-spline coefficients in the maximum detail level of the deformation pyramid and 16×16 B-spline coefficients in the highest resolution level of the image pyramid. We used 1.0 as a similarity weight w_s and 20.0 as a consistency weight w_c.

Let g be the transformation function obtained using the CSR algorithm. We measured the accuracy of the registration using the standard *warping index* (Thévenaz *et al* 1998) defined as

$$\varpi = \sqrt{\frac{1}{\|R\|} \sum_{\mathbf{x} \in R} \|g(\mathbf{x}) - g^*(\mathbf{x})\|^2}, \tag{8}$$

where $\|R\|$ is the number of image pixels. The warping index measures the geometric error— in pixels—between the deformation applied to the image and the transformation calculated by our algorithm, averaged out for all the triplets of the sequence. The original average warping index of the unregistered images was 4.50 ± 3.50 pixel. The average warping index of the 20 images after our sequential registration was 0.07 ± 0.05 pixel, i.e. the error was reduced to subpixel level.

The barrel/pincushion deformations affect mainly the border of the images. However, in practice, boundary features may be of less interest than central features. To test the performance of the algorithm in center-based distortions, we used 20 concatenated and randomly located fish-eye transformations of 60 pixel diameter and 3.0 magnification on the same histology image as the previous experiment. For an example of fish-eye transformation see figure 5(a) for a grid image and figure 5(b) for the histological section. The original average warping index of the unregistered images was 3.15 ± 0.11 pixel. The average warping index after applying CSR with the same parameters as before was 0.58 ± 0.37 pixel, i.e. the error was also reduced to subpixel level.

3.1.3. Registration of cyclical data. The iterative, bidirectional—forward and backward— way our CSR algorithm aims at the optimum registration transformation allows information from the whole set of images to be used in the registration process. This makes the algorithm
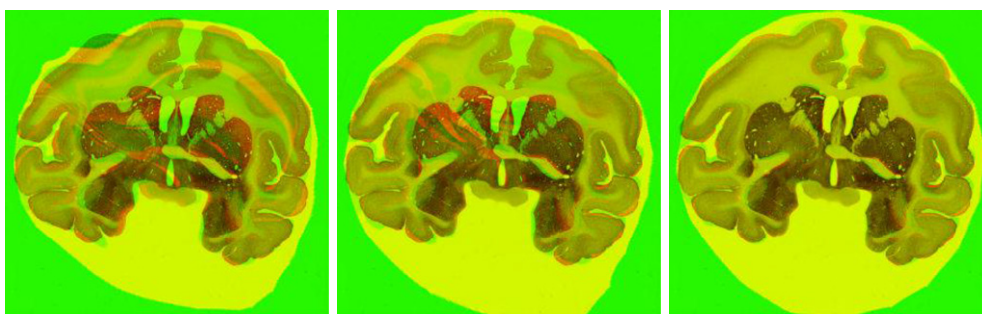
Figure 6. Visual estimation of registration accuracy of UPR (left), CPR (middle) and CSR (middle). A cyclical sequence—18 rotations of 20°—of a *M. fascicularis* brain section was registered using the three methods. The original image is shown in green and the final transformed image is in red. Therefore, the yellowish areas are regions of perfect overlap while the red and green areas represent incorrect alignment.

Table 1. Mean squared error (MSE), standard deviation (Std), minimum (Min) and maximum (Max) intensity difference between the original and the result after the combined rotations from the experiment described in section 3.1.5 for the new algorithm (CSR) and the previous methods (UPR and CPR).

	UPR	CPR	CSR
MSE	1494.38	1264.22	**715.24**
Std	2745.29	2559.83	**1499.46**
Min	0	0	**0**
Max	31684	29929	**22201**

suitable for registering sequences of images, especially if they are cyclical. In order to test this, we used a *M. fascicularis* brain section stained with cholinesterase. We applied 18 rotations to the image at 20° intervals, the last image of the sequence therefore being equal to the first one. Then we evaluated the results of the UPR, CPR and CSR algorithms by registering all 18 images and comparing the first and last images. We used $w_s = 1.0$ and $w_c = 10.0$. As expected, our sequential CSR method was more efficient for this type of sequences than the two other algorithms (table 1). Note that the MSE and the standard deviation (STD) obtained using CSR are half the ones obtained using either UPR or CPR. Furthermore, the CSR maximum intensity difference (MAX) is almost one fourth of the values obtained using UPR or CPR. Figure 6 shows the registration error visually for all three methods. We calculated and compared the percentage of correctly aligned pixels. For this, we converted the RGB images to HSV and set an acceptance threshold value of $\pm \frac{\pi}{25}$ that corresponds to a yellow hue value. CSR outperformed the other two methods: CSR produced 79.28% overlap compared to 55.99% (UPR) and 67.30% (CPR).

In figure 7, we show the MSE for each method and rotation angle. A neperian logarithmic scale was used in the graph. It can be observed that the CSR error remains consistently low throughout all angles by virtue of reusing sequence information. UPR causes high error values in all angles and CPR gets trapped in local minima for some of the angles, thus affecting the global registration.
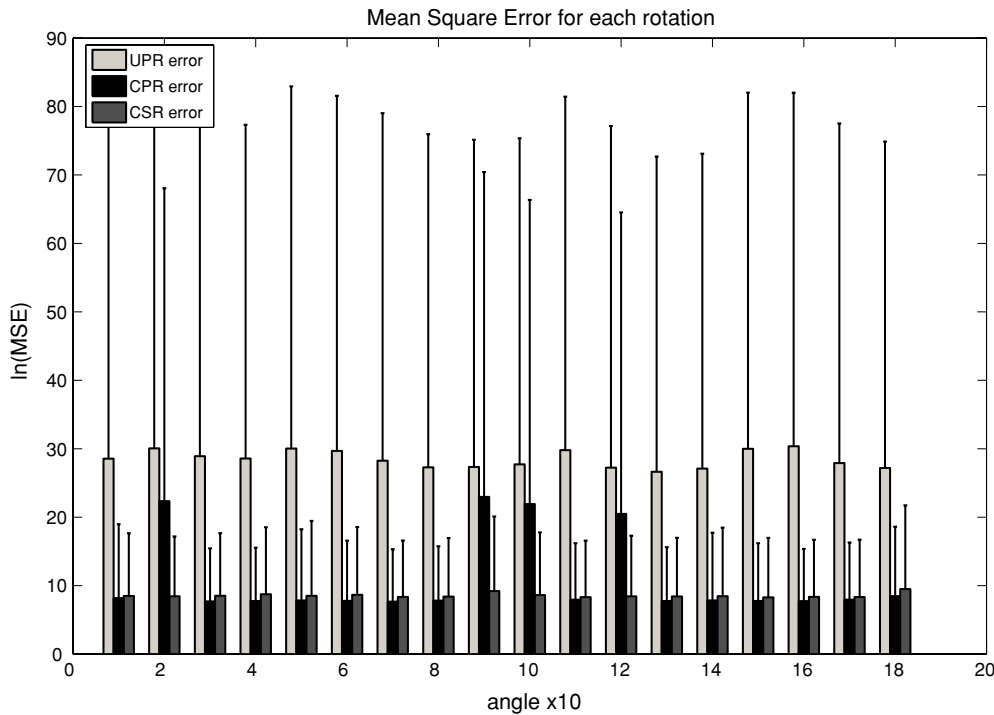
Figure 7. Error committed by each registration method on every angle of rotation in the cyclical experiment. A neperian logarithmic scale is used to facilitate the visualization of the error.

Table 2. Average and standard deviation of the MSE for the direct/inverse similarity (DSE/ISE) for the synthetic experiment of non-consistent sequential registration (non-CSR) versus CSR.

	Non-CSR	CSR
DSE	60.75 ± 143.50	32.48 ± 4.00
ISE	142.72 ± 255.78	33.72 ± 8.91

3.1.4. Use of neighbor information. Next we showed the benefits of using the consistency information in the registration process versus not using it. To this end we used the same sequence of monkey brain slices of the previous section and registered the sequence in groups of three images, first without consistency and then using our complete CSR algorithm ($w_c = 10.0$). In both cases, we used a four-level multiresolution image pyramid and a three-level deformation pyramid, with $w_s = 1.0$. As can be seen in table 2, our method substantially improved the registration results. To estimate the invertibility of the transformations calculated by both methods, we computed the final (direct and inverse) consistency errors, being 2243.50 ± 7387.42 and 2421.08 ± 6678.07 respectively for the non-consistent method and 1.69 ± 1.39 and 1.58 ± 1.16 for our CSR method.

3.1.5. Robustness against outliers. To further proof the benefit of using CTR as the building block of CSR, versus using UPR (Sorzano *et al* 2005) or CPR (Arganda-Carreras *et al* 2006) we used groups of three images where one of the members had been heavily distorted. This
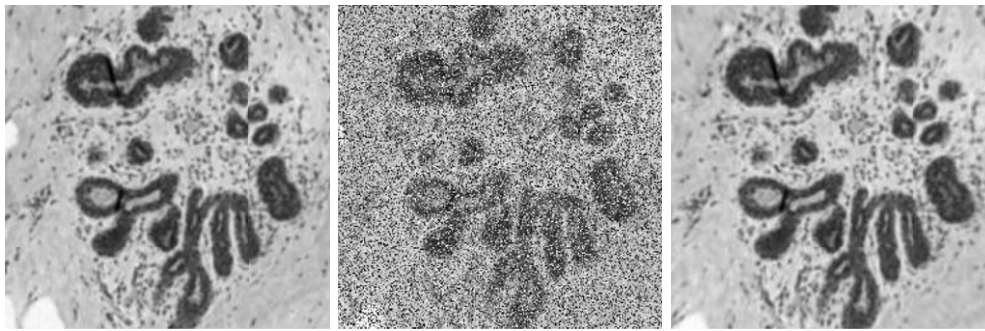
Figure 8. Example of triple histology images used in the validation test. A significant amount of random noise was added to the middle image.

way we showed the improvement obtained when combining consistency, bidirectionality and groupwise—in this case tri-wise—registration.

First, we used seven triples of images from the sequence of 20 used in section 3.1.2 and added random noise (with 50% of randomization, i.e. percentage of affected pixels) to one of the images of each triple (see figure 8). Then we registered triples using UPR, CPR and CTR with a four-level image pyramid and a five-level deformation pyramid. We used $w_s = 1.0$ and $w_c = 20.0$. The average warping indexes obtained were 14.96 ± 18.72 (UPR), 5.12 ± 5.33 (CPR) and 3.67 ± 3.77 (CTR). This shows the improvement obtained by using a consistent method compared to a non-consistent one, as well as the benefit of a tri-wise method over a pairwise method when one of the elements in the series is heavily distorted.

Next, we tested the robustness of our algorithm against independent, non-correlated rotations. This is an interesting problem from a practical point of view because it is very common in sequences of manually processed histological sections. We used ten groups of three images taken from a sequence of *M. fascicularis* brain sections stained with cholinesterase (grayscale, 300×282 pixel size). We then rotated one of the images of the triple a random number of degrees, uniformly distributed between 20 and 30. See figure 9 for an example. We used $w_s = 1.0$ and $w_c = 10.0$. The registration results, shown in table 3, confirm that CTR is more robust against outliers than UPR and CPR. Our CTR method achieves the lowest similarity error values while keeping moderate consistency error levels. CPR gives intermediate values for the similarity error while obtaining the best values for the consistency error. This is explained by the fact that CTR jointly minimizes the similarity error of three images (six dissimilarity measures) and the consistency error of two direct transformations with their corresponding inverse in both directions (four consistency measures), while in the two equivalent CPRs of the same images, the optimizer minimizes separately the dissimilarity of both pairs of images (four dissimilarity measures) and their corresponding direct-inverse transformations (four consistency measures).

3.1.6. Effect of the consistency term. Next, we have studied the behavior of the tri-wise algorithm in terms of the similarity error and the impact of our consistency term compared to using the previously described regularization term by Sorzano *et al* (2005). To this end, we used the first triple of the sequence of images described in the previous section (monkey brain section rotated 0, 20 and 40°). We then applied a tri-wise image registration using a four-level multiresolution image pyramid, a three-level deformation pyramid and either
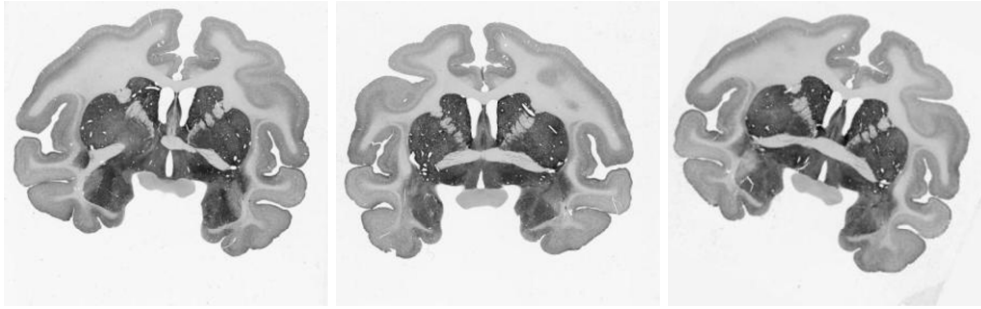
Figure 9. Example of a triple of *M. fascicularis* brain sections, stained with cholinesterase. The third image of the trio was artificially rotated 25°.

Table 3. Average and standard deviation of the MSE for the direct/inverse similarity (DSE/ISE) and the direct/inverse consistency (DCE/ICE) for the synthetic rotation experiment described in section 3.1.5 for the new algorithm (CTR) and the previous methods (UPR and CPR).

	UPR	CPR	CTR
DSE	543.21 ± 23.10	497.23 ± 97.84	**398.48 ± 66.29**
ISE	570.86 ± 30.67	489.65 ± 81.66	**450.29 ± 51.15**
DCE	15.28 ± 2.08	**2.15 ± 1.38**	8.59 ± 6.86
ICE	14.84 ± 1.28	**2.01 ± 1.53**	10.57 ± 4.09

 (i) only the similarity term of our energy functional (1), or
 (ii) the similarity term plus a vector-spline regularization with two sub-terms, one based on the gradient of the divergence of the deformation fields and one based on the gradient of their curl (Sorzano *et al* 2005), or
(iii) the similarity term plus our consistency term.

The results of the convergence of the three methods are shown in figure 10. Adding a regularization or consistency term to the energy function increases the complexity but in return forces the deformations to be smooth, which usually helps reducing the global similarity error E_S. When comparing both scenarios, we observe that using consistency instead of regularization produces an improvement in the final value of the similarity error, with the additional advantage of also obtaining the pseudo-inverse transformations. Additionally, in our case the number of deformation coefficients is sufficiently low so as to not need to be regularized. This is therefore the choice when no prior knowledge about the deformation exists. However, if *a priori* information about the deformation is known, adding an appropriate regularization term to the energy functional can greatly improve the results (Sorzano *et al* 2005).

3.1.7. Segmentation-based evaluation. A proper evaluation of the registration results cannot only be based on the final image dissimilarity or the recovery from specific deformation models and should take into account that the final goal of the registration, for instance in the case of histological sections, is to correct for unwanted deformations, and align the images while preserving the shape and connectivity of the structures of interest. An evaluation based on a known well-defined ground truth is required. For this reason, we performed the following experiment to quantitatively evaluate the results in a way which is independent of the algorithm.
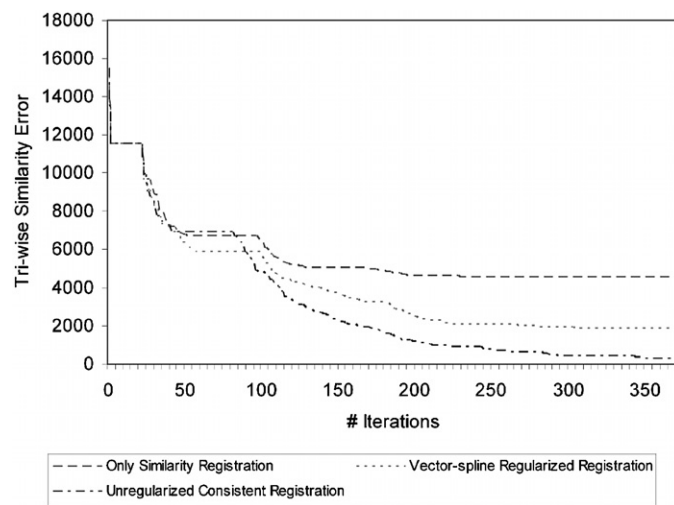
Figure 10. Evolution of the similarity error with the number of iterations for the non-regularized non-consistent, regularized non-consistent and non-regularized consistent tri-wise registration of the first image triple used in section 3.1.3.
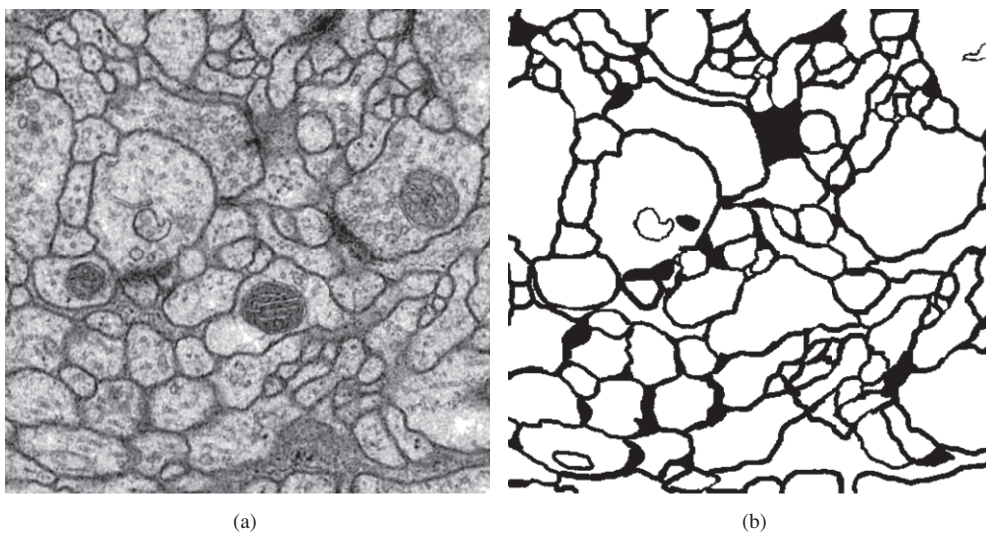

(a)   (b)

Figure 11. TEM section from Drosophila first instar larva ventral nerve cord (a) and corresponding binary labeling (b). The membranes are labeled in black, while the neurites and glia are labeled in white.

Namely, we used a 3D image stack composed of 30 serial transmission electron microscopy (TEM) sections, 512×512 pixel each. The dataset contains a Drosophila first instar larva ventral nerve cord. These sections, which have been manually segmented, are part of a public dataset kindly provided by Dr Albert Cardona, from the Institute of Neuroinformatics in Zurich, Switzerland (http://www.ini.uzh.ch/~acardona/data.html). See figure 11 for an example of TEM section and its binary labeling. The image stack is a $2 \times 2 \times 1.5$ μm volume,

(a)   (b)

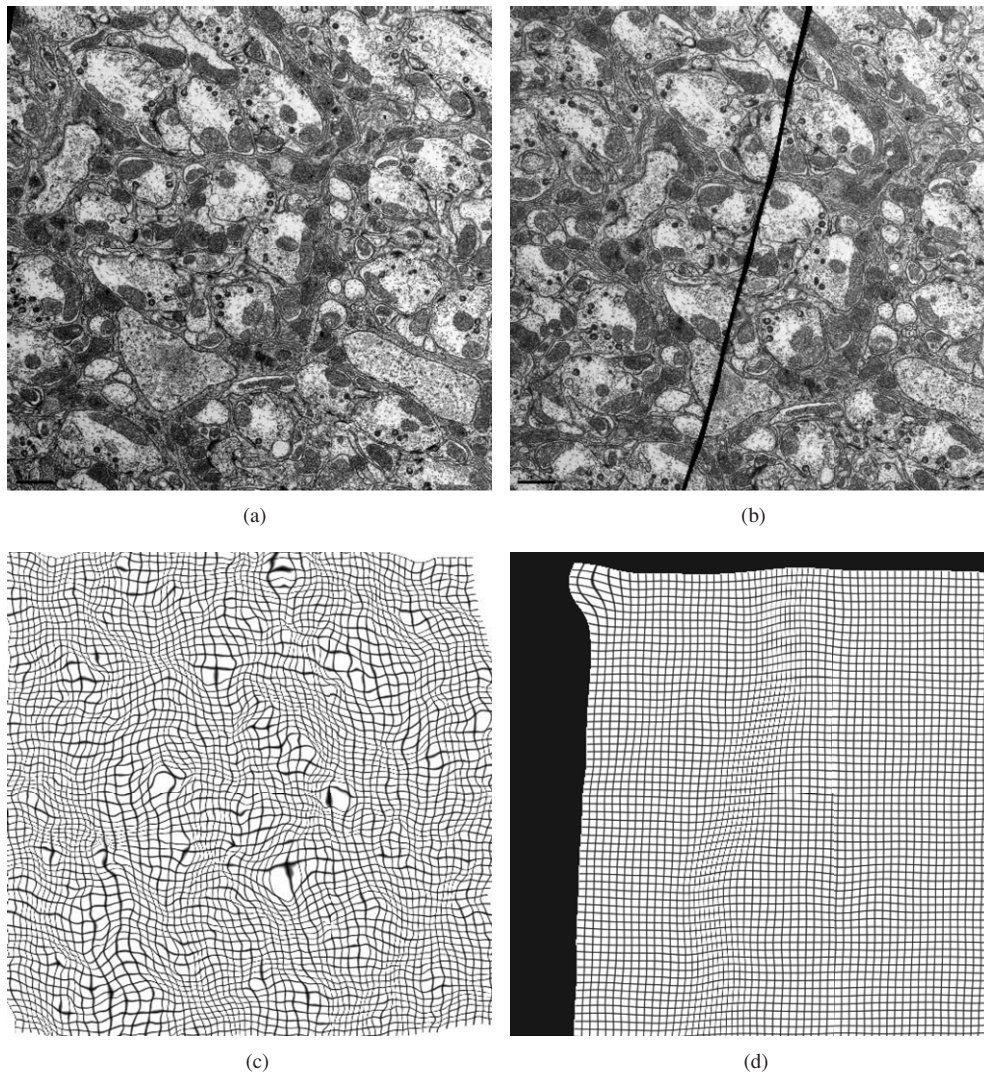(c)   (d)

Figure 12. Correction of heavy distortion. Two consecutive TEM tissue sections (a) where the second one (b) is folded in the middle. (c) Deformation grid after diffeomorphic registration (Vercauteren's method). (d) Deformation grid after CSR.

with a resolution of $4 \times 4 \times 50$ nm/pixel. On top of the already present minor misalignments and tissue distortions produced by the sectioning process and the slice thickness, we have applied a set of known elastic distortions to one every four images in the sequence. The elastic deformations were created by applying a grid of 8×8 B-spline coefficients, each of them following a Gaussian distribution with mean 30. These distortions simulate well the elasticity properties of the histological tissue, providing a proper test for the algorithm.

The registration was initialized with a rigid alignment before applying our CSR method with four levels in the multiresolution image pyramid, 8×8 deformation coefficients at the finest level, $w_s = 1.0$ and $w_c = 10.0$. To evaluate the results, we binarized the labels provided
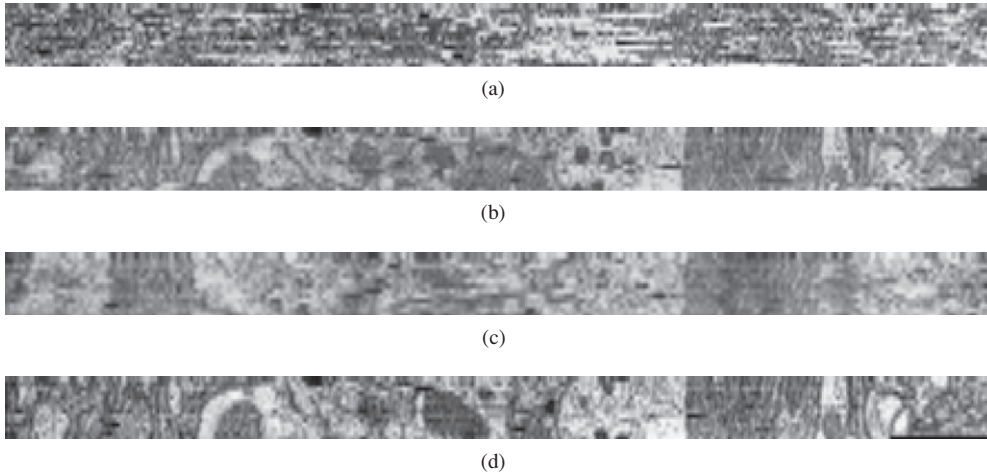
(a)

(b)

(c)

(d)

Figure 13. Detail of center section after re-slicing the sequence of TEM images from the top of the stack: (a) original misaligned sequence, (b) UER results, (c) ESR results, and (d) CSR results.

with the manual segmentation of the tissue, i.e. setting membranes to black and neurites and glia to white. We then quantified the misalignment between the original labels with the pre- and post-registration labels. The simplest way to do it is by measuring the amount of overlap between the labels. In summary, the pre-registered labels had $67.85 \pm 2.77\%$ overlap while the registered labels overlapped $78.53 \pm 2.17\%$. A different way of comparing the pre- and post-aligned segmentations is using the Rand error (Unnikrishnan *et al* 2007), which measures the agreement between two segmentations. For two different set of labels, S and T, the Rand error is defined as

$$R(S, T) = \binom{N}{2}^{-1} \sum_{i \neq j} |\delta(S_i, S_j) - \delta(T_i, T_j)|, \tag{9}$$

where the sum is over all pairs of distinct pixels i, j, N is the total number of pixels and $\delta(S_i, S_j)$ is 1 if i and j belong to the same object and 0 if they belong to different objects. In our case the Rand error of the pre-registered images was 0.0215 ± 0.0043, being reduced by 40% (0.0133 ± 0.0031) after CSR.

3.2. Experiments with real image sequences

We now show the results of using CSR to align real image sequences.

3.2.1. Robustness against non invertible transformations.
We first show how CSR can efficiently correct heavy distortions, such as those caused by folding and tearing in histological tissue sections. As a reference, we compare our CSR algorithm with Vercauteren's diffeomorphic demon (VDD) method (Vercauteren *et al* 2007), a non-parametric registration algorithm that generalizes Thirion's diffeomorphic demons (Thirion 1998). We only used the non-rigid registration algorithm, without previous rigid alignment of the images. We used three sections from a set of 14 consecutive 50 nm thick TEM sections of Lamina tissue from *Drosophila melanogaster*. All images were acquired at $3500\times$ magnification, 512×512 pixel
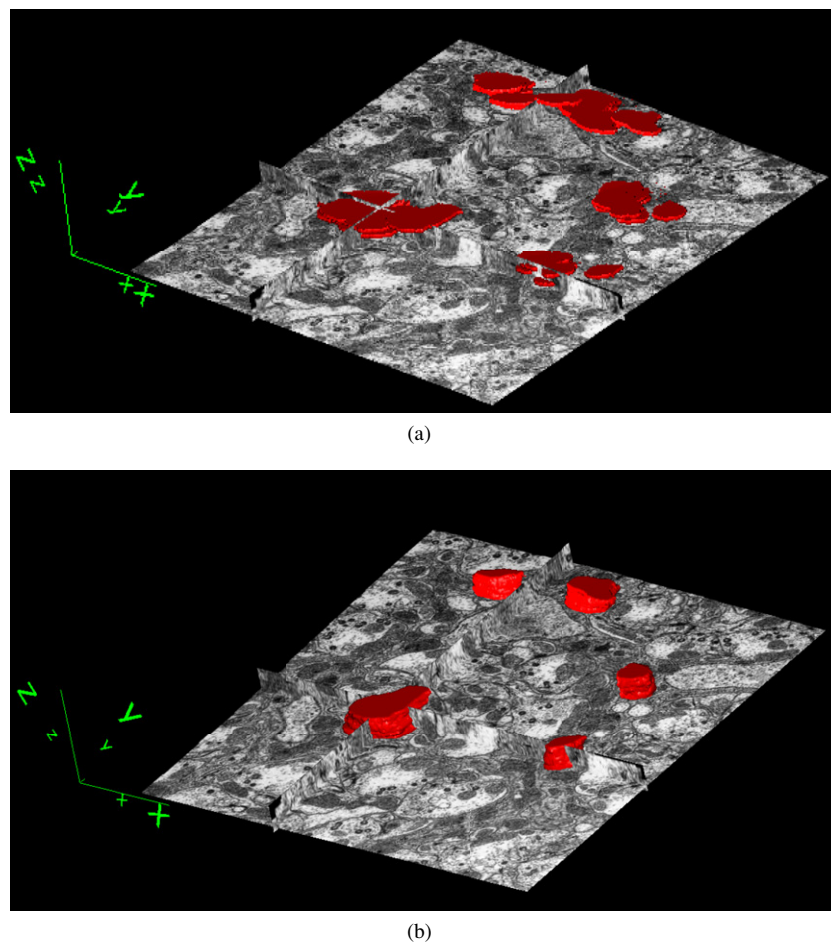
(a)

(b)

Figure 14. 3D reconstruction of registered lamina TEM sections with some segmented cells before (a) and after CSR alignment (b). The tissue is displayed in an orthoslice fashion to better appreciate the isosurfaces of the segmented cells (in red).

size. These thin tissue sections are often folded or torn, making the registration a difficult task. An example of this is shown in figure 12, which shows two consecutive TEM sections, one of them (figure 12(b)) affected by a folding area—black stripe in the middle of the section. We applied our CSR registration—adding a third section to complete a triple—and then VDD, obtaining a direct similarity error at convergence of 3306.38 (CSR) and 4546.82 (VDD). The corresponding deformation maps (figures 12(c) and (d)) show how the CSR method is able to model the tissue fold located in the middle of the image and capture the displacement between the structures caused by the cutting process, while the VDD algorithm is not. The demon method tries to locally adjust the images while the CSR method profits from the neighbor's information to better characterize the transformation. Then, to rule our the possibility that the better performance of CSR versus VDD was due to the initialization—note that VDD is known to be more effective at correcting local versus global changes than CSR—we modified the previous experiment and used an initial rigid alignment of the first pair of sections of the
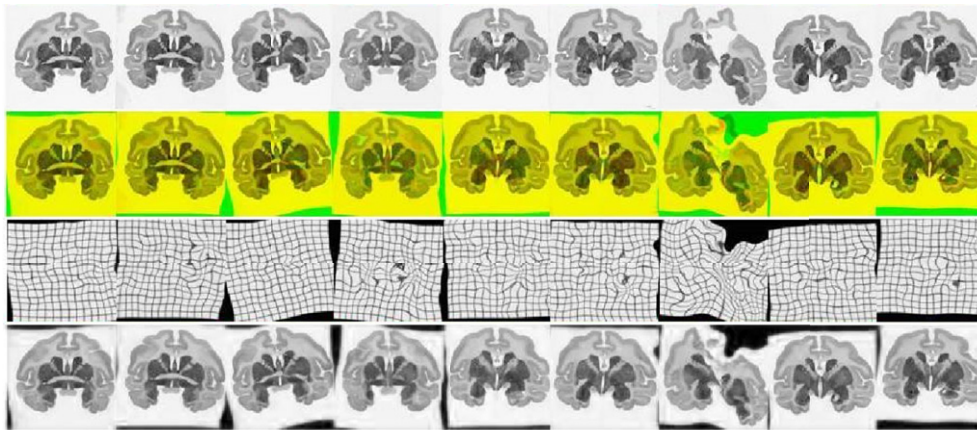
Figure 15. Sequence of *M. fascicularis* brain sections stained with cholinesterase. From top to bottom: original histology sections, RGB images showing some direct and inverse CSR results (yellow color meaning perfect superposition and red and green colors pointing out the misalignment regions) and their corresponding deformation grids and transformed versions.

triplet to have the best possible initial condition of the elastic registrations in both cases. Then the images were registered pair-wise only (CPR versus VDD). In this experiment, as expected, the image dissimilarities are closer, even though CPR still performs better than VDD: 3274, 59 (CPR) and 3652, 12 (VDD). However both results over-fit the alignment and consequently, many details of the source image are eliminated.

3.2.2. Performance versus other elastic methods. To further compare the performance of CSR with other standard registration methods, we used the whole set of 14 TEM lamina sections (512×512 pixel size) described in the previous section. The TEM image sequence was registered first using CSR, then with a combined method consisting of an initial affine registration followed by unidirectional elastic registration with conjugate gradient optimizer (UER), implemented in the open-source toolkit *elastix* and finally using the elastic sequential registration (ESR) developed by Ledesma-Carbayo *et al* (2005) with a previous affine alignment too. All methods used the same initial conditions: four levels of multiresolution image pyramid, 8×8 deformation coefficients at the finest level and the same stopping threshold (0.01), i.e. desired level of precision (absolute error difference between last and previous level registration). To evaluate the results, several cells of the tissue were manually segmented by an expert. The average cell overlap after registration was $78.48 \pm 17.77\%$ (UER), $81.22 \pm 6.91\%$ (ESR) and $83.91 \pm 6.06\%$ (CSR). The result can be visualized by re-slicing the sequence of sections and comparing it with the original alignment (figure 13(a)), the UER result (figure 13(b)), the ESR result (figure 13(c)) and the CSR result (figure 13(d)). Here we observe that in this case the ESR method minimizes the registration error by slightly over-aligning, since several 3D structures are missing after the alignment. In the UER and CSR re-slice results those structures are better preserved. Figure 14 shows the segmented cells—in red—and the sequence of 14 CSR-aligned TEM sections from the previous experiment. This way, we can see the segmented cells in a completely aligned 3D environment.
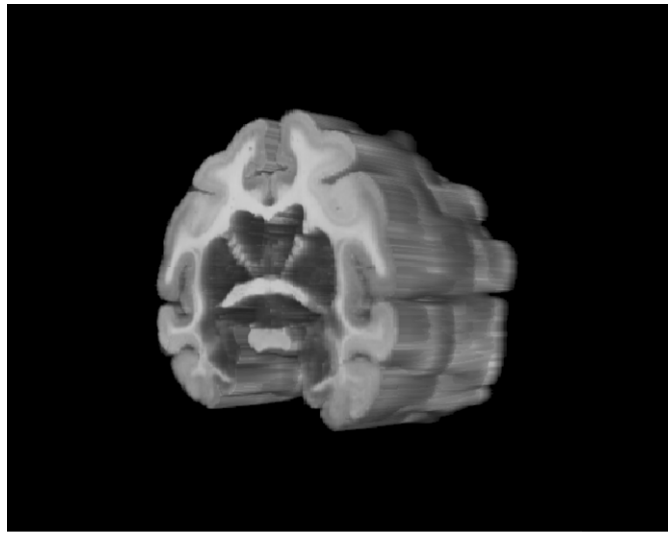
Figure 16. 3D reconstruction of registered *M. fascicularis* brain sections stained with cholinesterase.

These results confirm in real image series our previous deduction: CSR is especially beneficial when one or several images of the sequence are heavily distorted.

3.2.3. Registration of brain histological sections. We finally registered nine consecutive coronal cuts of *M. fascicularis*, including the triplet used in section 3.1.1. The registration algorithm used a five-level multiresolution image pyramid and a four-level deformation pyramid, with 1.0 as similarity weight and 10.0 as consistency weight. We set the stopping threshold to 0.01 empirically, and the maximum number of iterations (i.e. whole sequence registration in the corresponding level) to 10.

Our method achieved a satisfactory result, as can be seen in figure 15. Note that, even though the 7th slice is badly torn and the registration of that slice is far from accurate, that fact does not affect the registration of the neighboring sections, thus highlighting the benefit of using global information in the registration process. The initial dissimilarity error 2122.00 ± 426.36 was reduced to 375.27 ± 88.97, while VDD and UER only achieved 617.93 ± 348.73 and 610.34 ± 159.32, respectively. This significant difference in the minimization of the similarity error with respect to the previous example is due to the higher level of pyramid resolution and especially to the flat light background of the brain images. The registration time was 90 min on an Intel Pentium M, 1.59 GHz, with 2 GB of RAM memory, under Linux SUSE 10.0. The contours of the structures of interest were satisfactorily aligned, as can be seen in figure 16.

4. Conclusions

We have introduced a novel, automatic landmark-free method for the registration of 2D image sequences. Our method uses consistency, elasticity and full bidirectionality to register series of related images in a multiresolution fashion. We have taken advantage of the B-spline theory to implement both images and deformations with multiresolution spline pyramids, which produces a general, flexible and computationally efficient solution to the registration problem.

Furthermore, we chose a very powerful optimizer with the BFGS estimation of the Hessian, which combines steepest-descent and quasi-Newton steps in order to speed up convergence. In principle, the deformations represented by B-splines are not invertible; however and thanks to the enforced consistency of the transformation fields, we obtain deformations that are quasi-inverse of each other. Moreover, the method spreads the information from the neighbors through the whole set of images enforcing the consistency of the whole sequence.

As shown in section 3, this algorithm reduces the warping index of synthetic images to sub-pixel level, even when the deformation belongs to a space that cannot be exactly represented using B-splines. These results are better than those obtained using classic pairwise unidirectional or bidirectional (consistent) registration algorithms. Besides, we have shown that our algorithm is more robust than those methods against noise or large rotations. We believe that this increase in robustness is due to the combined use of consistency and global image information, granted by the iterative, back and forth tri-wise registration algorithm, which assigns equal relevance to all the images of the sequence.

Finally, although our algorithm allows using both landmarks and deformation regularization, we show that registering using the consistency constraint provides satisfactory, unbiased results without requiring either of them. This is especially helpful when the regularization term cannot be customized due to the lack of *a priori* information about the deformation and the number of deformation coefficients is sufficiently low. Based on these improvements, we show the benefit of applying iteratively tri-wise bidirectional consistent registration to sequences of images and in particular to series of TEM and optical microscopy histological images.

The extension of the proposed algorithm to the registration of 3D images is mathematically straightforward. It would require adding the third—z—dimension to the equations included in section 2 and in the appendices A and B, and the reformulation of those equations to include the derivatives with respect to the third dimension. However, this extension would involve a considerable increase of the computational load in terms of time and memory consumption. A popular solution to this limitation is the use of a reduced and randomly sampled set of voxels instead of the complete image stacks in the registration process. If properly distributed, these selected voxels will lead the algorithm to an optimal solution.

This 2D registration method is available from the author on request as an ImageJ/Fiji plugin (Rasband 1997–2009).

Acknowledgments

Appendix A. Operators and explicit derivatives

We calculate the derivatives of the similarity and consistency terms with respect to the deformation coefficients c in appendices A.1 and A.2. Because B-splines were used to represent the images and the deformations, the derivatives can be calculated explicitly. Then we define the reduction and expansion operators needed to move through the resolution pyramids in appendix A.3.

A.1. Data term derivatives

The derivative of the similarity term E_S^{ij} between images I_i and I_j with respect to the coefficients c of the deformation field $c_{a,k,l}^{r,s}$ can be easily inferred from the corresponding definitions of E_S^{ij} (see (3)) and $c_{a,k,l}^{r,s}$ (see (7)). Thus,

$$\frac{\partial E_S^{ij}}{\partial c_{a,k,l}^{rs}} = -2 \sum_{\mathbf{x} \in \Omega_{ij}} \left[(I_j(\mathbf{x}) - I_i(g^{ij}(\mathbf{x}))) \frac{\partial I_i(g^{ij}(\mathbf{x}))}{\partial c_{a,k,l}^{rs}} \right], \tag{A.1}$$

where $a \in \{1, 2\}$, $i, j, r, s \in \{1, 2, 3\}$ and $k, l \in \mathbb{Z}^2$.

The last term in the expression above equals

$$\frac{\partial I_i(g^{ij}(\mathbf{x}))}{\partial c_{a,k,l}^{rs}} = \frac{\partial I_i(g^{ij}(\mathbf{x}))}{\partial g^{ij}(\mathbf{x})} \cdot \frac{\partial g^{ij}(\mathbf{x})}{\partial c_{a,k,l}^{rs}}$$

$$= \frac{\partial I_i(g^{ij}(\mathbf{x}))}{\partial g_1^{ij}(\mathbf{x})} \frac{\partial g_1^{ij}(\mathbf{x})}{\partial c_{a,k,l}^{rs}} + \frac{\partial I_i(g^{ij}(\mathbf{x}))}{\partial g_2^{ij}(\mathbf{x})} \frac{\partial g_2^{ij}(\mathbf{x})}{\partial c_{a,k,l}^{rs}}, \tag{A.2}$$

where $\cdot$ denotes the scalar product. Finally, we need to evaluate $\frac{\partial g_b^{ij}(\mathbf{x})}{\partial c_{a,k,l}^{rs}}$.

If $i = r$ and $j = s$ ($ij = \{12, 21, 23, 32\}$), we can write

$$\frac{\partial g_b^{ij}(\mathbf{x})}{\partial c_{a,k,l}^{ij}} = \frac{\partial g_b^{ij}(x, y)}{\partial c_{a,k,l}^{ij}}$$

$$= \begin{cases} \beta^3\left(\frac{x}{s_x} - k\right)\beta^3\left(\frac{y}{s_y} - l\right) & \text{if} \quad a = b \\ 0 & \text{if} \quad a \neq b \end{cases} \tag{A.3}$$

with $b \in \{1, 2\}$.

In other cases, that is, if $i \neq r$ and $j \neq s$ and $|i - j| = 2$ (i.e. $ij = \{13\}$ and $rs = \{12, 23\}$ or $ij = \{31\}$ and $rs = \{21, 32\}$), the expressions are slightly more complicated. For instance, for $ij = \{13\}$ and $rs = \{12\}$, we get

$$\frac{\partial g_b^{13}(\mathbf{x})}{\partial c_{a,k,l}^{12}} = \frac{\partial g_b^{23}\left(g_b^{12}(x, y)\right)}{\partial c_{a,k,l}^{12}}$$

$$= \frac{\partial g_b^{23}}{\partial x}\bigg|_{(x', y')} \frac{\partial g_b^{12}}{\partial c_{a,k,l}^{12}}\bigg|_{(x, y)} + \frac{\partial g_b^{23}}{\partial y}\bigg|_{(x', y')} \frac{\partial g_{|b-2|+1}^{12}}{\partial c_{a,k,l}^{rs}}\bigg|_{(x, y)}, \tag{A.4}$$

where $(x', y') = g_b^{12}(x, y)$ and $\frac{\partial g_b^{12}(\mathbf{x})}{\partial c_{a,k,l}^{12}}$ is calculated as given in (A.3).

For $ij = \{13\}$ and $rs = \{23\}$, we can write

$$\frac{\partial g_b^{13}(\mathbf{x})}{\partial c_{a,k,l}^{23}} = \frac{\partial g_b^{23}\left(g_b^{12}(x, y)\right)}{\partial c_{a,k,l}^{23}}$$

$$= \frac{\partial g_b^{23}}{\partial c_{a,k,l}^{23}}\bigg|_{(x', y')} + \frac{\partial g_{|b-2|+1}^{23}}{\partial c_{a,k,l}^{23}}\bigg|_{(x', y')}, \tag{A.5}$$

where $(x', y') = g_b^{12}(x, y)$ and $\frac{\partial g_b^{23}(\mathbf{x}')}{\partial c_{a,k,l}^{23}}$ is calculated as in (A.3).

The corresponding derivatives for $ij = \{31\}$ and $rs = \{21, 32\}$ are obtained by symmetry from the ones already calculated.

A.2. Consistency term derivatives

The consistency energy functional E_C^{ij} needs to be evaluated only for the intermediate deformation fields, which simplifies the number of the derivatives to calculate. We have

$$\frac{\partial E_C^{ij}}{\partial c_{a,k,l}^{rs}} = -2 \sum_{\mathbf{x} \in \Omega_{i,j}} (\mathbf{x} - g^{ji}(g^{ij}(\mathbf{x}))) \left(\frac{\partial g^{ji}(g^{ij}(\mathbf{x}))}{\partial c_{a,k,l}^{rs}} \right). \tag{A.6}$$

For $i = r$ and $j = s$, we can write for the last term

$$\begin{aligned}
\frac{\partial g^{ji}(g^{ij}(\mathbf{x}))}{\partial c_{a,k,l}^{ij}} &= \frac{\partial g^{ji}(g^{ij}(\mathbf{x}))}{\partial g^{ij}(\mathbf{x})} \cdot \frac{\partial g^{ij}(\mathbf{x})}{\partial c_{a,k,l}^{ij}} \\
&= \frac{\partial g^{ji}(g^{ij}(\mathbf{x}))}{\partial g_1^{ij}(\mathbf{x})} \frac{\partial g_1^{ij}(\mathbf{x})}{\partial c_{a,k,l}^{ij}} + \frac{\partial g^{ji}(g^{ij}(\mathbf{x}))}{\partial g_2^{ij}(\mathbf{x})} \frac{\partial g_2^{ij}(\mathbf{x})}{\partial c_{a,k,l}^{ij}} \\
&= \left. \frac{\partial g^{ji}}{\partial x} \right|_{(x',y')} \left. \frac{\partial g_1^{ij}}{\partial c_{a,k,l}^{ij}} \right|_{(x,y)} + \left. \frac{\partial g^{ji}}{\partial y} \right|_{(x',y')} \left. \frac{\partial g_2^{ij}}{\partial c_{a,k,l}^{ij}} \right|_{(x,y)},
\end{aligned} \tag{A.7}$$

where $(x', y') = g_1^{ij}(x, y)$ and $\frac{\partial g_b^{ij}}{\partial c_{a,k,l}^{ij}}$ is calculated from (A.3).

For $i = s$ and $j = r$, the evaluation is straightforward also from (A.3):

$$\begin{aligned}
\frac{\partial g^{ji}(g^{ij}(\mathbf{x}))}{\partial c_{a,k,l}^{ji}} &= \frac{\partial g^{ji}(\mathbf{x}')}{\partial c_{a,k,l}^{ji}} \\
&= \frac{\partial g_1^{ji}(\mathbf{x}')}{\partial c_{a,k,l}^{ji}} + \frac{\partial g_2^{ji}(\mathbf{x}')}{\partial c_{a,k,l}^{ji}},
\end{aligned} \tag{A.8}$$

where $(x', y') = g^{ij}(x, y)$ and $\frac{\partial g_b^{ji}}{\partial c_{a,k,l}^{ji}}$ is calculated from (A.3).

A.3. Reduction and expansion operators

The expansion operator maps a coarser level of the pyramid onto a finer grid, and the reduction operator makes the complementary action. Following Unser *et al* (1993), we define the expansion and reduction operators needed to implement the image and deformation pyramids. When the image dimensions are not a power of 2, we divide by 2 and truncate to the closest integer. We can write the expansion operator as

$$c_{-1}(k) = \left(u_2^3 * \uparrow_2 (c_0) \right)(k), \tag{A.9}$$

where c_0 are the initial coefficients, u_2^3 is the binomial filter defined by

$$u_2^3(k) = \begin{cases} 2^{-3} \begin{pmatrix} 4 \\ k+2 \end{pmatrix} & |k| \leqslant 2 \\ 0 & \text{other,} \end{cases} \tag{A.10}$$

the $*$ operator represents a discrete convolution such that $(u * v)(k) = \sum_{l \in \mathbb{Z}} u(l) v(k - l)$ and the operator $\uparrow_2 (\cdot)$ upsamples its argument by a factor of 2.

The reduction operator can be expressed as

$$c_1(k) = \left(\tfrac{1}{2}(b^7)^{-1} * \downarrow_2 \left(u_2^3 * b^7 * c_0 \right) \right)(k), \tag{A.11}$$

where $b^n(k) = \beta^n(x)|_{x=k}$ for $n = 3, 7$, and the operator $\downarrow_2 (\cdot)$ downsamples its argument by 2. Since u_2^3 and b^7 are FIR filters, the main difficulty to evaluate (A.11) stems from the fact that $(b^7)^{-1}$ is an IIR filter.

As images are always finite, some data extrapolation is necessary to calculate the infinite sum involved in filtering with $(b^7)^{-1}$. We have chosen to extend the images using the widely extended mirror-on-bounds symmetric boundary conditions. If $f_0(k)\forall k \in [0, \ldots, N-1]$ is the 1D signal we want to process and $f(k)$ is the mirrored signal with symmetric boundary conditions, then

$$f(k) = \begin{cases} f_0(-k) & k < 0 \\ f_0(k) & k \in [0, \ldots, N-1] \\ f_0(2N-2-k) & k > N-1, \end{cases} \quad \text{(A.12)}$$

Let us next give the definition of the mirror-on-bounds anti-symmetric boundary conditions in 1D. Let $f(k)$ be the mirrored signal with anti-symmetric boundary conditions $\forall x \in \mathbb{R}$:

$$\begin{cases} f(x) - f(0) = f(0) - f(-x) \\ f(x+N-1) - f(N-1) = f(N-1) - f(N-1-x). \end{cases} \quad \text{(A.13)}$$

The direct B-spline filter $(b^7)^{-1}$ is an all-pole system that can be implemented efficiently using a cascade of first-order causal and anti-causal recursive filters (Unser *et al* 1993). Next, we describe the explicit procedure for the calculation using the same reasoning as described in Unser (1999).

By sampling the seventh degree B-spline at integers values we find that

$$b^7(z) = (z^3 + 120z^2 + 1191z + 2416 + 1191z^{-1} + 120z^{-2} + z^{-3})\left(\frac{1}{5040}\right). \quad \text{(A.14)}$$

Thus, the filter to implement is

$$(b^7)^{-1}(k) \leftrightarrow \frac{1}{b^7(z)} = 5040 \prod_{i=1}^{3} \left(\frac{1}{z_i z^{-1}}\right)\left(\frac{-z_i}{1 - z_i z}\right), \quad \text{(A.15)}$$

with $z_1 = -0.535\,28$, $z_2 = -0.122\,558$ and $z_3 = -0.009\,148\,69$ (see table A1). This three-pole filter can be implemented as three sequential one-pole filters. Given the input signal values $\{f(k)\}_{k=0,\ldots,N-1}$ and defining $c^-(k) = \frac{c(k)}{5040}$. The right-hand side factorization leads to the following recursive algorithm:

$$c^+(k) = f(k) + z_i c^+(k-1), \qquad (k = 1, \ldots, N-1), \quad \text{(A.16)}$$

$$c^-(k) = z_i c^-(k+1) + (1 - z_i)^2 c^+(k), \qquad (k = N-2, \ldots, 0), \quad \text{(A.17)}$$

where z_i is the corresponding pole. Note that the first filter is casual, running from left to right and the second filter is anticausal running from right to left. Therefore, we need to specify the appropriate starting values for both recursions, i.e. $c^+(0)$ and $c^-(N-1)$. To ensure that the procedure is reversible, we impose the requirement that $f(k)$ can be recovered exactly by convolving $c(k)$ with b^7 using the same type of boundary conditions. The starting values to be used for symmetric mirror boundary conditions are given by Unser (1999). We have calculated the corresponding values for anti-symmetric mirror boundary conditions using the same reasoning. The initial value for the first recursion is given by

$$c^+(0) = \left(\frac{1}{1 - z_i^{2N-2}}\right)$$

Table A1. Poles of the septimic B-spline.

$$\theta = \tfrac{1}{3} \arccos \left(-\frac{738}{\sqrt{556\,549}} \right)$$
$$c = \sqrt{301} \cos(\theta)$$
$$s = \sqrt{903} \sin(\theta)$$
$$\lambda_1 = 20 - 2c$$
$$\lambda_2 = 20 + c - s$$
$$\lambda_3 = 20 + c + s$$
$$z_1 = \sqrt{\lambda_1^2 - 1} - \lambda_1$$
$$z_2 = \sqrt{\lambda_2^2 - 1} - \lambda_2$$
$$z_3 = -\left(\sqrt{\lambda_3^2 - 1} + \lambda_3 \right)^{-1}$$

$$\left(\left(\frac{1 + z_i}{1 - z_i} \right) \left(f\left(0 + z_i^{N-1} f(N-1)\right) \right. \right.$$
$$\left. \left. + \sum_{n \in [1, \ldots, N-2]} \left(z_i^{(2N-2-n)} - z_i^n \right) f(n) \right) \right). \tag{A.18}$$

Note that here $f(k)$ indicates the results of the previous recursions and not the original samples as it was above in (A.12) and (A.13). Correspondingly, the initialization for the second recursion is

$$c^-(N-1) = c^+(N-1) - z_i c^+(N-2). \tag{A.19}$$

Appendix B. Selection of boundary conditions

The following straightforward experiment shows the benefit of using anti-symmetric over symmetric boundary conditions to represent the deformation fields. We first applied a barrel/pincushion (Ma *et al* 2003) deformation to a 256×256 pixel image of a uniform grid (figure B1). This type of deformation is very relevant in this context because it preferentially affects the borders of the image. If the input coordinate **x** is normalized to lie in $[1, -1] \times [1, -1]$, and if r_{in} is the radius of its polar expression in this coordinate system, then the output radius r_{out} produced by the distortion is $r_{out} = 1 + k_1 r_{in}^2 + k_2 r_{in}^4$. Thus, the type of barrel pincushion deformation is given by the signs of k_1 and k_2. Thus, $k_1 < 0$ and $k_2 < 0$ ($k_1 > 0$ and $k_2 > 0$) generate a barrel (pincushion) deformation. We represented the deformation using 16×16 B-spline coefficients with k_1 and k_2 set to 0.10. Then, we reduced the deformation field to 8×8 coefficients using both symmetric and anti-symmetric boundary conditions and subsequently expanded them back. To compare the effect of the boundary conditions, we calculated the difference between the original and the reduced and then expanded deformation fields, for both the symmetric and anti-symmetric case. Figure B2 shows the details of the upper-left corner of the difference image, where the effect of the boundary conditions should be more pronounced. It is clear from the image that anti-symmetric mirror conditions preserve the deformation field better than their symmetric counterpart. Note also that the anti-symmetric boundary conditions force the second derivatives to be zero at the boundary, which is convenient when working with distortions that have small high-order derivatives.
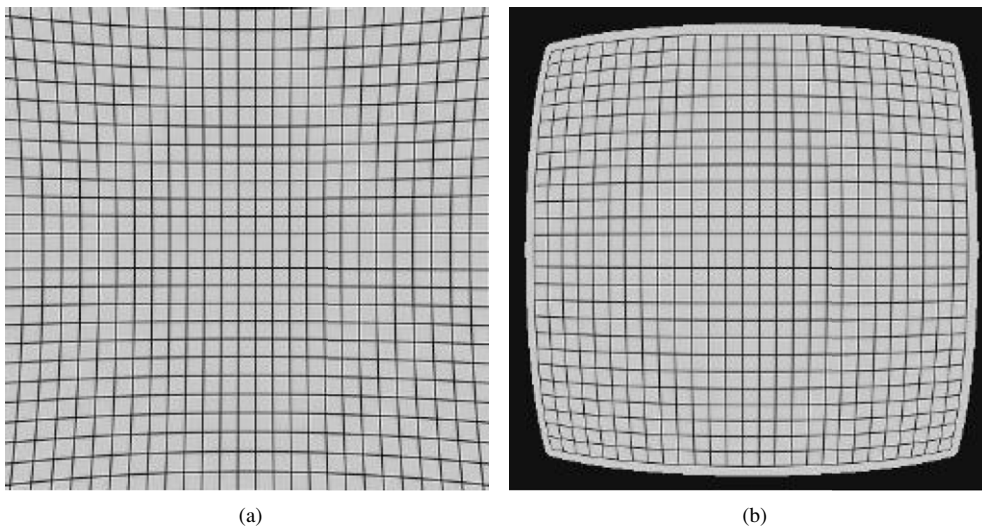
(a)   (b)

Figure B1. (a) Pincushion distortion ($k_1 = 0.10$ and $k_2 = 0.05$) of a uniform grid and (b) its inverse barrel transformation ($k_1 = -0.10$ and $k_2 = -0.05$).
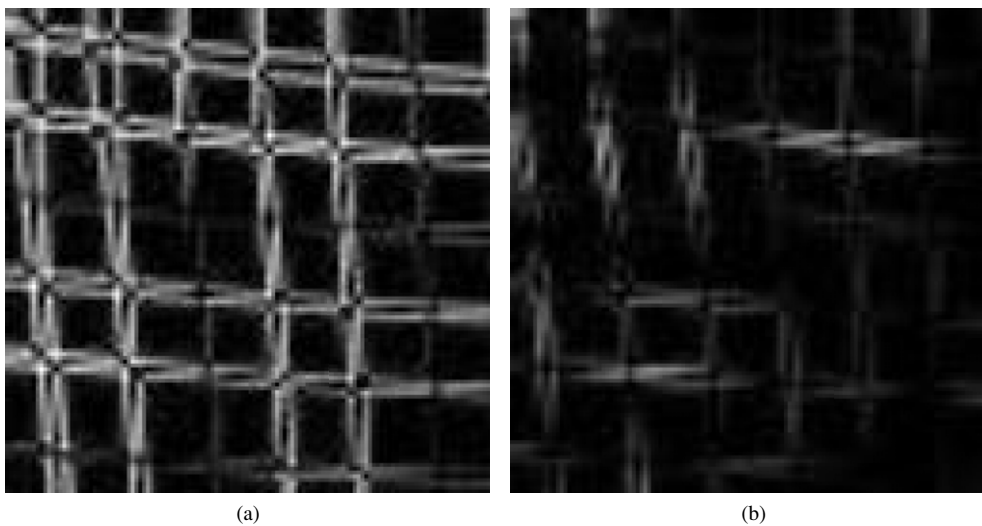

(a)   (b)

Figure B2. Upper-left corner of the difference image between 16×16 B-spline coefficients barrel/pincushion deformation and its 8×8 coefficients reduced and then expanded version using symmetric boundary conditions (a) and anti-symmetric boundary conditions (b).

References

Andronache A, Cattin P and Szekely G 2006 Local intensity mapping for hierarchical non-rigid registration of multi-modal images using the cross-correlation coefficient *Biomedical Image Registration (Lecture Notes in Computer Science* vol 4057*)* (Berlin: Springer) pp 26–33

Arganda-Carreras I, Sorzano C O S, Marabini R, Carazo J M, Ortiz-de Solorzano C and Kybic J 2006 Consistent and elastic registration of histological sections using vector-spline regularization *Computer Vision Approaches to Medical Image Analysis (Lecture Notes in Computer Science* vol 4241*)* (Berlin: Springer) pp 85–95

Auer M, Regitnig P and Holzapfel G A 2005 An automatic nonrigid registration for stained histological sections *IEEE Trans. Image Process.* **14** 475–86

Avants B B, Schoenemann P T and Gee J C 2006 Lagrangian frame diffeomorphic image registration: morphometric comparison of human and chimpanzee cortex *Med. Image Anal.* **10** 397–412

Bhatia K K, Hajnal J V, Puri B K, Edwards A D and Rueckert D 2004 Consistent groupwise non-rigid registration for atlas construction *Proc. IEEE Int. Symp. on Biomedical Imaging: Nano to Macro* vol 1 pp 908–11

Christensen G E and Johnson H J 2001 Consistent image registration *IEEE Trans. Med. Imaging* **20** 568–82

Cootes T, Marsland S, Twining C, Smith K and Taylor C 2004 Groupwise diffeomorphic non-rigid registration for automatic model building *Computer Vision—ECCV 2004: Part 4 (Lecture Notes in Computer Science* vol 2034) ed T Pajdla and J Matas (Berlin: Springer) pp 316–27

Csapo I, Holland C M and Guttmann C R 2007 Image registration framework for large-scale longitudinal MRI data sets: strategy and validation *Magn. Reson. Imaging* **25** 889–93

Geng X, Kumar D and Christensen G E 2005 Transitive inverse-consistent manifold registration *Information Processing in Medical Imaging (Lecture Notes in Computer Science* vol 3565*)* pp 468–79

Grenander U and Miller M 1998 Computational anatomy: an emerging discipline *Q. Appl. Math.* **56** 617–94

Huang J, Abendschein D, Dávila Román V G and Amini A A 1999 Spatio-temporal tracking of myocardial deformation with a 4D B-spline model from tagged MRI *IEEE Trans. Med. Imaging* **18** 957–72

Kybic J and Unser M 2003 Fast parametric elastic image registration *IEEE Trans. Image Process.* **12** 1427–42

Ledesma-Carbayo M J, Kybic J, Desco M, Santos A, Sühling M, Hunziker P R and Unser M 2005 Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation *IEEE Trans. Med. Imaging* **24** 1113–26

Ma L, Chen Y and Moore K L 2003 Flexible camera calibration using a new analytical radial undistortion formula with application to mobile robot localization *Proc. IEEE Int. Symp. on Intelligent Control* pp 799–804

Malandain G, Bardinet E, Nelissen K and Vanduffel W 2004 Fusion of autoradiographs with an MR volume using 2-D and 3-D linear transformations *Neuroimage* **23** 111–27

Marsland S, Twining C and Taylor C 2003 Groupwise non-rigid registration using polyharmonic clamped-plate splines *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2003 (Lecture Notes in Computer Science* volume 2879*)* (Berlin: Springer) pp 771–9

Matula P, Kozubek M and Matula P 2004 Applications of image registration in human genome research *Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis (Lecture Notes in Computer Science* vol 3117*)* ed M Sonka, I A Kakadiaris and J Kybic (Berlin: Springer) pp 376–84

Pluim J P W, Maintz J B A and Viergever M A 2003 Mutual information based registration of medical images: a survey *IEEE Trans. Med. Imaging* **22** 986–1004

Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes in C* 2nd edn (Cambridge: Cambridge University Press)

Radeva P, Amini A A and Huang J 1997 Deformable B-solids and implicit snakes for 3D localization and tracking of SPAMM MRI data *Comput. Vis. Image Underst.* **66** 163–78

Rasband W 1997–2009 *ImageJ* US National Institutes of Health, Bethesda, Maryland, USA (available online: http://rsbweb.nih.gov/ij/)

Rogelj P and Kovacic S 2006 Symmetric image registration *Med. Image Anal.* **10** 484–94

Rueckert D, Aljabar P, Heckemann R A, Hajnal J V and Hammers A 2006 Diffeomorphic registration using B-Splines *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2006* (Berlin: Springer) pp 702–9

Rueckert D, Sonoda L I, Hayes C, Hill D L G, Leach M O and Hawkes D J 1999 Non-rigid registration using free-form deformations: application to breast MR images *IEEE Trans. Med. Imaging* **18** 712–21

Skouson M B, Guo Q and Liang Z P 2001 A bound on mutual information for image registration *IEEE Trans. Med. Imaging* **20** 843–6

Sorzano C O S, Thévenaz P and Unser M 2005 Elastic registration of biological images using vector-spline regularization *IEEE Trans. Biomed. Eng.* **52** 652–63

Szeliski R and Shum H Y 1996 Motion estimation with quadtree splines *IEEE Trans. Pattern Anal.* **18** 1199–210

Thévenaz P, Ruttimann U E and Unser M 1998 A pyramid approach to subpixel registration based on intensity *IEEE Trans. Image Process.* **7** 27–41

Thirion J P 1998 Image matching as a diffusion process: an analogy with Maxwell's demons *Med. Image Anal.* **2** 243–60

Twining C, Cootes T, Marsland S, Schestowitz R, Petrovic V and Taylor C 2005 A unified information-theoretic approach to groupwise non-rigid registration and model building *Information Processing in Medical Imaging (Lecture Notes in Computer Science* vol 3565*)* ed G Christensen and M Sonka (Berlin: Springer) pp 1–14

Twining C, Marsland S and Taylor C 2004 Groupwise non-rigid registration: the minimum description length approach *Proc of the British Machine Vision Conf.* pp 417–26

Unnikrishnan R, Pantofaru C and Hebert M 2007 Toward objective evaluation of image segmentation algorithms *IEEE Trans. Pattern Anal. Mach. Intell.* **29** 929–44

Unser M 1999 Splines: a perfect fit for signal and image processing *IEEE Signal Process. Mag.* **16** 22–38

Unser M, Aldroubi A and Eden M 1991 Fast B-spline transforms for continuous image representation and interpolation *IEEE Trans. Pattern Anal.* **13** 277–85

Unser M, Aldroubi A and Eden M 1993 The L_2-polynomial spline pyramid *IEEE Trans. Pattern Anal.* **15** 364–79

Vercauteren T, Pennec X, Perchant A and Ayache N 2007 Non-parametric diffeomorphic image registration with the demons algorithm *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2007 (Lecture Notes in Computer Science* vol 4792*)* pp 319–26

Wirtz S, Fischer B, Modersitzki J and Schmitt O 2004 Super-fast elastic registration of histologic images of a whole rat brain for three-dimensional reconstruction *Proc. SPIE* **5370** 328–34

Wirtz S, Papenberg N, Fischer B and Schmitt O 2005 Robust and staining-invariant elastic registration of a series of images from histologic slices *Proc. SPIE* **5747** 1256–62

Yushkevich P A, Avants B B, Ng L, Hawrylycz M, Burstein P D, Zhang H and Gee J C 2006 3D mouse brain reconstruction from histology using a coarse-to-fine approach *Int. Workshop on Biomedical Image Registration (Lecture Notes in Computer Science* vol 4057*)* pp 230–7

Zitová B and Flusser J 2003 Image registration methods: a survey *Image Vis. Comput.* **21** 977–1000