

Piecewise Affine Registration of Biological Images for Volume Reconstruction

Alain Pitiot^{a,b,c,*} Eric Bardinet^{b,d} Paul M. Thompson^c
Grégoire Malandain^b

^a*Mirada Solutions, Ltd., Level 1, 23-38 Hythe Bridge Street, Oxford, OX1 2EP, United Kingdom*¹

^b*EPIDAURE Laboratory, INRIA, Sophia Antipolis, France*

^c*Laboratory of Neuro Imaging, UCLA School of Medicine, Los Angeles, USA*

^d*CNRS, UPR640-LENA, Paris, France*

Abstract

This manuscript tackles the reconstruction of 3D volumes via mono-modal registration of series of 2D biological images (histological sections, autoradiographs, cryosections, etc.). The process of acquiring these images typically induces composite transformations that we model as a number of rigid or affine local transformations embedded in an elastic one. We propose a registration approach closely derived from this model.

Given a pair of input images, we first compute a dense similarity field between them with a block matching algorithm. We use as a similarity measure an extension of the classical correlation coefficient that improves the consistency of the field. A hierarchical clustering algorithm then automatically partitions the field into a number of classes from which we extract independent pairs of sub-images. Our clustering algorithm relies on the Earth mover's distribution metric and is additionally guided by robust least-square estimation of the transformations associated with each cluster. Finally, the pairs of sub-images are, independently, affinely registered and a hybrid affine/non-linear interpolation scheme is used to compose the output registered image.

We investigate the behavior of our approach on several batches of histological data and discuss its sensitivity to parameters and noise.

Key words: registration, clustering, reconstruction, histology, MRI

* Corresponding author.

Email address: apitiot@loni.ucla.edu (Alain Pitiot).

¹ *Fax:* +44 1865 265 501

1 Introduction

A key component of medical image analysis, image registration essentially consists of bringing two images, acquired from the same or different modalities, into spatial alignment. This process is motivated by the hope that more information can be extracted from an adequate merging of these images than from analyzing them independently. For instance, mono-modal registration of a population's MRIs can be used to build anatomical atlases (Collins et al., 2003; Thompson et al., 2000), while mono- or multi-modal registration of the same patient's data can help determine the nature of an anomaly (Jolesz et al., 2001) or monitor the evolution of a tumor (Haney et al., 2001) or other disease process (Rey et al., 2002).

In particular, pair-by-pair registration of a series of 2-D biological images (histological sections or autoradiographs) enables the reconstruction of a 3D biological image. Subsequent fusion with 3D data acquired from tomographic imaging modalities (*e.g.* MRI) then allows the tissue properties to be studied in an adequate anatomic framework, using *in vivo* reference data.

More formally, given two input images, registering the floating (*i.e.*, movable) image to the reference (*i.e.*, fixed) one entails finding the transformation that minimizes the dissimilarity between the transformed floating image and the reference. As such, it can be decomposed into three elements:

- a transformation space, which describes the set of admissible transformations from which one is chosen to apply to the floating image;
- a similarity criterion, which measures the discrepancy between the images; and
- an optimization algorithm, which traverses the transformation space, in search of the transformation that will minimize the similarity criterion.

A large variety of transformation spaces have been discussed in the literature: among others, one finds linear transformations (rigid, affine) and non-linear transformations (polynomial (Woods et al., 1998), polyaffine (Feldmar and Ayache, 1996; Arsigny et al., 2003), elastic (Davatzikos, 1997; Gee et al., 1993) or fluid (Christensen, 1999)). Similarly, many similarity criteria have been presented: Studholme et al. (1996) use normalized mutual information, Collins et al. (1995) cross-correlation, Roche et al. (2000) the correlation ratio, Ashburner and Friston (1999) the squared intensity difference, etc. Optimization algorithms range from the straightforward Powell method (Collignon et al., 1995) to sophisticated multi-scale Levenberg-Marquardt techniques (Taubin, 1993) or stochastic search (Wells et al., 1996) (please refer to Maintz and Viergever (1998), van den Elsen et al. (1993) or Cachier (2002, Chapitre 1)

for thorough reviews of medical image registration).

1.1 Registration algorithms for 3-D reconstruction

By registering each pair of consecutive slices in the stack of biomedical images, we can recover a geometrically coherent 3-D alignment of the 2-D images. This problem essentially consists of the mono-modal registration of similar objects (even though non-coherent distortions may occur between consecutive slices). A number of techniques have been proposed in the literature.

manual registration: Classically, people have relied on their anatomical expertise to register images manually (Deverell et al., 1993; Rydmark et al., 1992), a time-consuming, operator-dependent and poorly reproducible process.

fiducial markers: The use of fiducial markers helps increase the reliability of the registration process (Ford-Holevinski et al., 1991; Goldszal et al., 1996, 1995; Humm et al., 1995). In histology, physical markers (straight rigid needles for instance) can be inserted into the organ to be processed, prior to the slicing step. They are later tracked in the images. A least square minimization process then attempts to recover the original geometrical configuration.

While obviously a lot less operator-independent and more reproducible than manual registration, the biases introduced by the non-orthogonality between the needle's axes and the cutting planes and the geometrical deformations undergone by the needle holes during the histological preparations (the chemical baths can collapse the holes) may significantly hamper the ability of the least square minimization to recover the correct transformation.

geometrical approaches: These require that feature elements be extracted *a priori*: those elements are usually geometrically remarkable points such as lines, curvature extrema or contours. Hibbard and Hawkins (1988) for instance use principal axes to align digital autoradiograms. Better performances in terms of precision can be achieved by matching contours (Cohen et al., 1998; Zhao et al., 1993), edges (Kay et al., 1996; Kim et al., 1995) or points (Rangarajan et al., 1997).

While it could be argued that these techniques enable a better control over the registration process, the segmentation of feature elements can sometimes be as difficult a problem as the overall registration process itself, with similar drawbacks: non-reproducibility if the feature are extracted manually, lack of precision for fully automated segmentation, etc.

iconic approaches: They rely on the comparison of the intensities associated to voxels in the input images. Usually, they can be casted into a similarity-minimization framework where the transformation space is traversed in search for the transformation which minimizes the intensity dis-

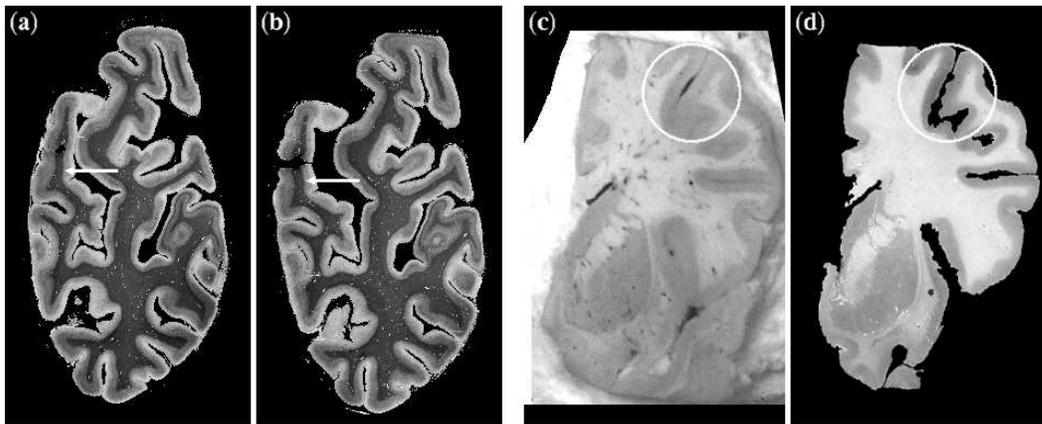


Fig. 1. Two consecutive myelin-stained histological sections of the human brain (a, b); Human brain cryosection (c) and its associated Nissl-stained section (d). White arrows and circles indicate moving gyri.

similarity between the transformed floating image and the reference image (Davatzikos, 1997; Christensen, 1999; Studholme et al., 1996; Wells et al., 1996, and many others). In Malandain et al. (2004), the reconstruction process was additionally guided by an external anatomical reference (MRI) to further approximate the anatomical ground truth.

1.2 Application-tailored registration

A careful study of the above-mentioned registration methods and reviews highlights the lack of specificity of most of the approaches introduced in the literature. All too often, these techniques are presented in a very general context and advertised as generic, even though they ought to be applied to specific registration problems. This is all the more surprising that specific approaches abound in closely related fields, like medical image segmentation for instance, where the combined use of atlases and of a variety of shape and appearance constraints (see (McInerney and Terzopoulos, 1996)) helps tailor the segmentation methods to the segmentation problems at hand. Furthermore, again with few remarkable exceptions (for iconic methods, Roche et al. (2000) derived optimal similarity measures from an analysis of the expected relationships between the input images, with different hypotheses leading to different measures), very limited *a priori* medical expertise is used in the design of the registration algorithms.

Yet, medical image registration is an ill-posed problem (in the sense of Hadamard). This primarily stems from the difficulty of characterizing the objectives and requirements of the registration process, as these may quite substantially vary from one application to the next and often depend on the images to be registered. For instance, the registration precision required in a post-mortem quan-

titative study of drug effect depends on the position in the volume: clearly, it must be maximum in or around the structures targeted by the drug and does not really matter elsewhere. However, when building anatomical atlases, most, if not all, structures and organs should be correctly matched. Furthermore, even though a given image transformation may adequately put the floating and reference image in correspondence, this transformation may not reflect the actual physical transformation that took place, if the latter only exists (for multi-patient registration for instance, it might be difficult to establish the existence of such physical transformation). Finally, a number of transformations, and not only one, may give very similar result in terms of visual correspondence, most especially when the number of degrees of freedom of the allowed transformations is large. Evaluating the quality of a registration process then becomes particularly difficult when the images are considered independently from their medical context, that is, when the problem to be solved is not that of putting in correspondence two views of an underlying medical truth but that of registering two images, taken as mere sets of voxels with associated intensity values.

We submit that much benefit is to be gained from the use of medical expertise in the design of a registration methodology. We have consequently developed a new registration paradigm, the piecewise approach, adapted to the registration of 2-D biomedical images, which we articulate below and detail in the following Section 2.

1.3 Piecewise Registration

As described above, a registration algorithm basically consists of three elements: a similarity measure, a transformation search space, and an optimization algorithm, which must all three conform to *a priori* medical expertise.

1.3.1 Adapted similarity measure

In a fashion similar to Roche’s tailorization of the similarity measure to the expected relationships between the input images, we have extended a classical similarity measure, the correlation coefficient, with the goal to allow a better modeling of the combined transformation and intensity relationships via the building of a more coherent similarity map of the input floating image (see Section 2.1.2).

Special care was also taken to distinguish the background from the actual pieces of tissues to be registered, as treatments are only applied to the latter (see Section 2.1.1).

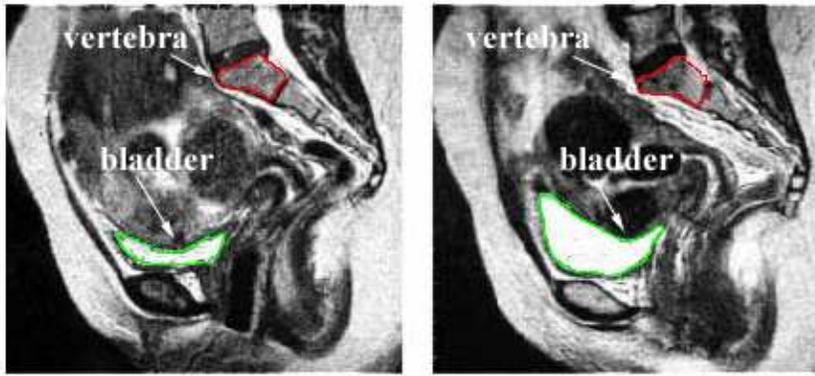


Fig. 2. Two abdominal MRIs of the same patient, with corresponding vertebra (red) and bladder (green) outlined.

1.3.2 Histology-tailored transformation space

A priori knowledge about the acquisition process for biological images should also allow the transformation space to be modeled more accurately. In our case, the cutting process, successive chemical treatments, and the glass mounting step that a slab of tissue undergoes during a histological preparation yield a fairly flexible global transformation that is however locally affine for some identifiable components of the section (even though the chemical baths introduce non-linearities, an affine model remains a good approximation). In brain sections for instance, each gyrus (compare white arrows in Fig. 1(a) & (b), and white circles in (c) & (d)) undergoes a transformation (due to successive manipulations) relatively independent from those of other gyri: in (b) the operator introduced an artificial rotation to the marked sulcus when he positioned the tissue on the glass slide; holes and tears appeared in (d) during the slicing and chemical bath steps. In spite of the large variety of available transformation spaces in the literature, their functional form may not reflect our specific needs.

Discussions with neuro-anatomists and histologists prompted us to consider the input biological images as a set of independent components, subject to linear transformations. We then model the composite transformation yielded by the chain of physical histological processes as a number of affine or rigid transformations applied to carefully delimited areas, with non-linear transformations interpolated in between.

Note that the utility of this transformation model extends beyond biological images (our primary motivation here) to medical images as well. For instance, abdominal or torso MRIs (see Fig. 2) often include rigid structures such as bones (ribs, vertebrae, etc.), deformable organs (liver, heart, etc.), and elastic tissues. Two abdominal MRIs of the same patient are then linked by a complex transformation which can be rigid in some regions (for bones) but potentially exhibits large local dilations (in deformable organs). Global rigid

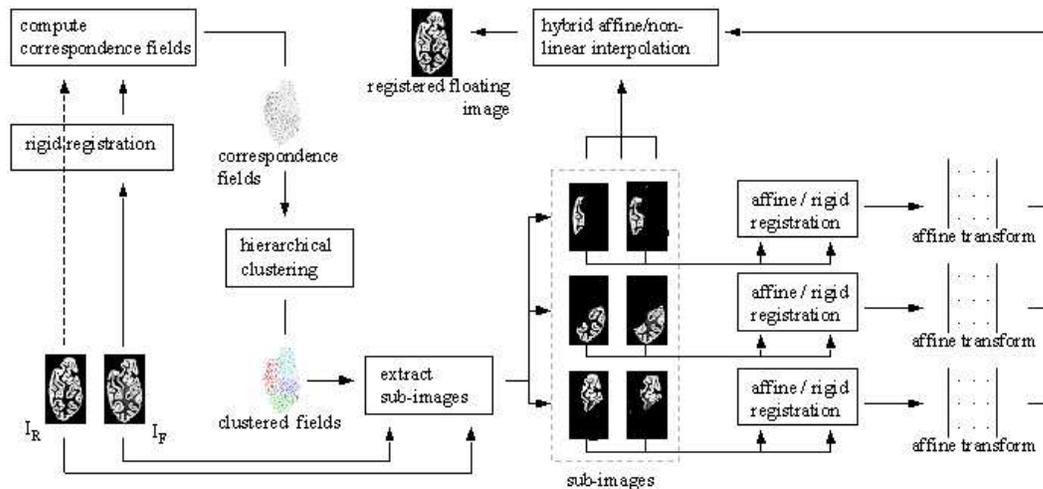


Fig. 3. Detailed overview of the piecewise registration approach.

or affine transformations cannot adequately handle such a case. Also, a single rigid transformation could not correctly register all the vertebrae along the spinal column simultaneously. Furthermore, high degree of freedom (*e.g.*, fluid) transformations could correctly map one image onto the other, but they may not guarantee that specific components (*e.g.*, bones) will be only rigidly transformed.

1.3.3 Optimization algorithm

We selected a robust block-matching/least square estimation approach (see Sections 2.1.1 and 2.3). Here also, prior segmentation of the tissues helps discard background blocks from the chain of treatments. A number of tests on the blocks (intensity variance, number of active pixels, etc. see (Ourselin et al., 2001)) further adapt the search for an adequate transformation towards anatomically meaningful ones.

1.3.4 Piecewise registration

Fig. 3 displays an overview of the piecewise approach. Briefly, given the input floating and reference images, I_F and I_R respectively, we first rigidly register I_F to I_R , before computing a dense similarity map (or correspondence field) between them with a block matching algorithm (Section 2.1.1). A hierarchical clustering algorithm then partitions the correspondence field into a number of classes from which we extract independent pairs of sub-images (Section 2.2). Our clustering algorithm relies on a distribution metric (the Earth mover’s distance) to agglomerate blocks, and uses the estimated transformations associated with each cluster to guide the grouping process. The pairs of sub-images are then, separately, rigidly or affinely registered (Section 2.3). Finally, an hy-

brid affine/non-linear interpolation scheme is used to compose the registered floating image (Section 2.4).

Note that even though we specifically tailored the various modules of our approach to the problem at hand, namely the reconstruction of a 3-D histological volume, most of them could be independently optimized or replaced on a need-for basis. For instance, the block-matching affine registration algorithm we chose for the sub-images could be traded for another affine registration approach from the literature. In that, the volume reconstruction system we propose here can also be seen as a generic framework. However, care would have to be taken not to compromise the overall homogeneity and robustness. Good global robustness requires the selection or the development of robust modules. Homogeneity ensures that the best overall performances can be achieved from individual contributions.

We detail our method in Section 2 and discuss the reconstruction of two histological volumes in Section 3 along with the sensitivity of our algorithm to noise conditions and parameters.

2 Method

The first step of our approach consists of *automatically* partitioning the input floating and reference images (I_F and I_R) into a number of pairs of corresponding sub-images, where each sub-image is associated with an independent (in terms of transformation) image component.

We approach this segmentation issue as a process of partitioning a correspondence field computed from I_F to I_R . Our method is motivated by the following observation. When both images are composed of pairs of independent components, where each component is subject to some linear transformation, the associated correspondence field should exhibit rather homogeneous characteristics within each component, and heterogeneous ones across them. Consequently, by clustering the fields with a criterion based on local characteristics, we hope to extract from them the desired independent components.

2.1 Computing the initial correspondence field

We use a block-matching algorithm (Jain, 1981) to compute the correspondence field. Block-matching was favored over classical non-linear registration

approaches as it offers better spatial independence between neighboring correspondences, as opposed to standard techniques whose regularization schemes induce spatial correlations. Indeed, given the unpredictable nature of the input images (which may contain tears, holes, rotated sulci, etc.), we must allow for independent variations of the correspondence field.

2.1.1 Block-matching algorithm

We associate with I_F and I_R two rectangular lattices $L_F = \{(i, j) \in [1, \dots, w_F] \times [1, \dots, h_F]\}$ and $L_R = \{(i', j') \in [1, \dots, w_R] \times [1, \dots, h_R]\}$ respectively, whose sites correspond to pixels in the input images. We may choose to associate a site to each pixel of the input images, in which case w_F, h_F and w_R, h_R are the width and height of I_F and I_R , respectively. We could also consider a sparser regular or non-regular site distribution. In our case, we use sparse regular lattices and discard, for histological sections, sites which lie on the background. A prior segmentation step is then required to identify these background blocks: depending on the modality and quality (in terms of signal to noise ratio, and contrast) of the input images, we either use simple intensity thresholding (the case for the myelin-stained human brain section of Fig. 1), or a more sophisticated segmentation algorithm (the neural classifier introduced in (Pitiot et al., 2002) for instance).

The block-matching algorithm works as follows (see Fig. 4): for each site (i, j) in L_F , we consider a neighborhood $b_F^{i,j}$ in I_F of the pixels associated with (i, j) (usually a square neighborhood of constant size b_{size} called a “block”, whose centroid is denoted by $p_F^{i,j}$). We then compute the similarity measures (given a similarity metric sim) between block $b_F^{i,j}$ and every block $b_R^{k,l}$ in I_R associated with sites (k, l) in the corresponding neighborhood $N_R^{i,j}$ of (i, j) in L_R (the “exploration neighborhood”). For every site (i, j) in L_F , we then get a 2-D spatial similarity distribution (the values $sim(b_F^{i,j}, b_R^{k,l})$ defined in the neighborhood $N_R^{i,j}$ of (i, j)). We also record the “arg max” displacement $d^{i,j}$ defined by $d^{i,j} = p_R^{(k,l)_{max}} - p_F^{i,j}$ where $(k, l)_{max}$ is the site of L_R that is associated with the block $b_R^{k,l}$ in $N_R^{i,j}$ which is the most similar to $b_F^{i,j}$, *i.e.* $(k, l)_{max} = \arg \max_{k,l} sim(b_F^{i,j}, b_R^{k,l})$. This displacement field will serve to estimate transformations on clusters (see Section 2.2.1).

The quality of both the similarity map and the displacement field is essentially determined by three parameters: the size of the blocks, the similarity metric and the size of the exploration neighborhood in L_R .

- The similarity metric and the size of the blocks must reflect the expected relationship between the intensity distributions of blocks in the floating and reference images, and the scale of the features of interest within those blocks

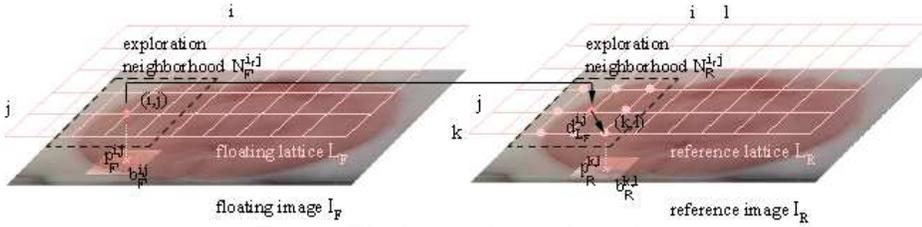


Fig. 4. Block-matching algorithm

respectively (see (Collins et al., 2003) and (Dengler, 1991) for details).

- The size of the exploration neighborhood is linked to the expected magnitude of the residual displacements after global alignment. It conditions the extent to which our registration algorithm can recover large deformations: the further apart corresponding components are, the larger the size of the neighborhood must be.

Section 3.2 comments on the impact of our algorithm’s parameters on the overall registration quality. Please refer to Table 1 for some standard values.

As a pre-processing step, we first rigidly register I_F to I_R to remove from the subsequently computed correspondence fields the global rigid transform that uniformly affects all components. We use the fully automated intensity-based registration algorithm presented in Ourselin et al. (2001), where a *robust* multi-scale block-matching strategy was introduced. Not accounting for this would only degrade the quality of the field and affect the efficiency of the clustering.

2.1.2 Constrained correlation coefficient

A ubiquitous choice for image registration (Roche et al., 2000), the correlation coefficient represents, in the context of block-matching, a measure of the affine dependency between the block of interest $b_F^{i,j}$ in the floating images and every block $b_R^{k,l}$ in the corresponding exploration neighborhood in the reference image. It is written:

$$\text{cor} \left(b_F^{i,j}, b_R^{k,l} \right)^2 = \frac{\text{cov} \left(b_F^{i,j}, b_R^{k,l} \right)^2}{\text{var} \left(b_F^{i,j} \right) \cdot \text{var} \left(b_R^{k,l} \right)} = \frac{\left(\sum_{u,v} \left(b_F^{i,j} (u,v) - \mu_F^{i,j} \right) \cdot \left(b_R^{k,l} (u,v) - \mu_R^{k,l} \right) \right)^2}{\sum_{u,v} \left(b_F^{i,j} (u,v) - \mu_F^{i,j} \right)^2 \cdot \sum_{u,v} \left(b_R^{k,l} (u,v) - \mu_R^{k,l} \right)^2} \quad (1)$$

where $\mu_F^{i,j}$ and $\mu_R^{k,l}$ are the mean intensities of $b_F^{i,j}$ and $b_R^{k,l}$ respectively. To make the affine dependency clearer, equation 1 can be rewritten (Roche et al.,

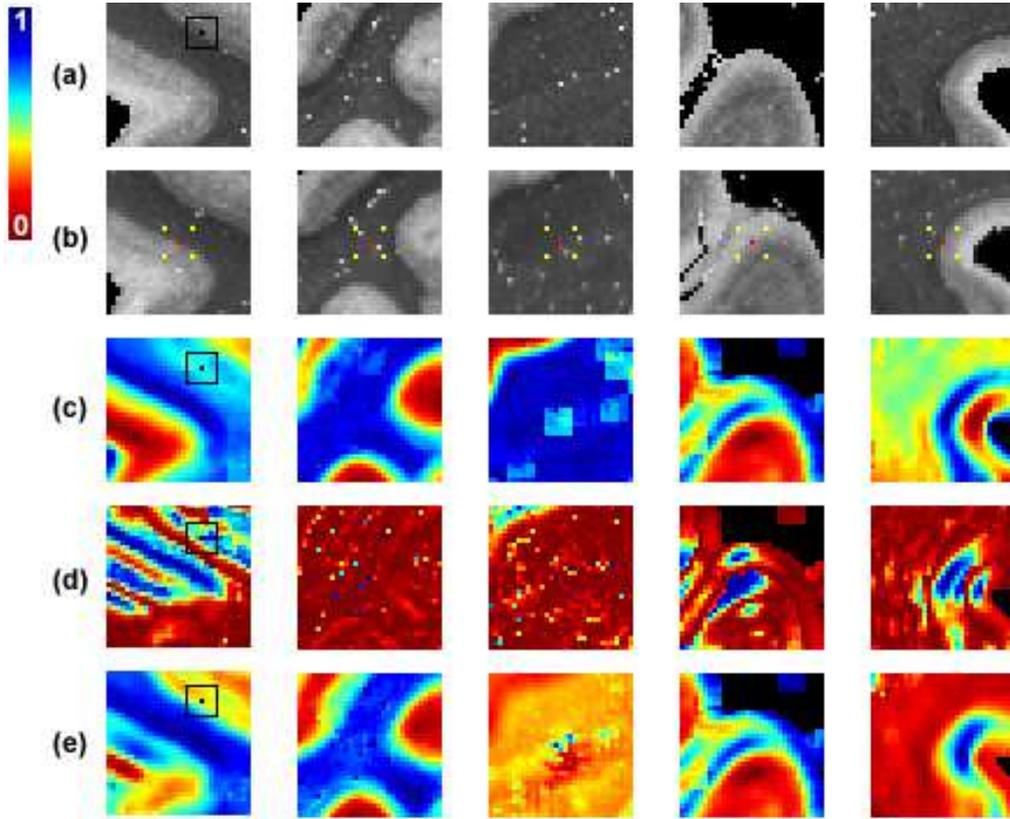


Fig. 5. Similarity maps induced by three different similarity measures: (a) input reference sub-images; (b) input floating sub-images with considered floating blocks in red and yellow; (c) maps induced by the SSD measure; (d) maps induced by the classical correlation coefficient; (e) maps induced by the constrained correlation coefficient.

2000):

$$1 - \text{cor}(b_F^{i,j}, b_R^{k,l})^2 = \frac{\min_{A,B} E [b_F^{i,j} - A \cdot b_R^{k,l} - B]^2}{\text{var}(b_F^{i,j})} \quad (2)$$

where E is the statistical expectation and A and B represent the affine coefficients of the intensity dependency function between $b_F^{i,j}$ and $b_R^{k,l}$.

This affine formulation reflects the assumption that each block contains at most two different tissues, a reasonable hypothesis when dealing with small image windows. A variety of studies have documented the effectiveness of the correlation coefficient in not only mono- but also multi-modal registration applications (Roche et al., 2000; Ourselin et al., 2001).

Yet, when building the similarity map of $b_F^{i,j}$ (and thus, also when computing the “arg max” displacement field), different *implicit* A and B are used with every block in the reference exploration neighborhood. Comparing similarity values, both within the similarity distribution associated with a single floating block and across the distributions associated to different floating blocks, when obtained under those conditions then becomes questionable. A simple way to alleviate this issue consists of *explicitly* estimating A and B over the entire neighborhoods $N_R^{i,j}$ (the “exploration” neighborhood of (i, j) in L_R) and $N_F^{i,j}$ (the corresponding, same size, neighborhood in L_F), to keep them constant during the computation of the similarity distribution of a given floating block. Equation 1 then becomes:

$$\text{ccor} \left(b_F^{i,j}, b_R^{k,l} \right)^2 = \frac{\sum_{u,v} \left(b_F^{i,j}(u,v) - M_F^{i,j} \right) \cdot \left(b_R^{k,l}(u,v) - M_R^{i,j} \right)}{\sum_{u,v} \left(b_F^{i,j}(u,v) - M_F^{i,j} \right)^2 \cdot \sum_{u,v} \left(b_R^{k,l}(u,v) - M_R^{i,j} \right)^2} \quad (3)$$

where $M_F^{i,j}$ and $M_R^{i,j}$ are the mean intensities of $N_F^{i,j}$ and $N_R^{i,j}$ in the floating and reference image respectively.

- By replacing $\mu_R^{k,l}$ by $M_R^{i,j}$ in equation 3, we validate the comparison of the similarity measures within the exploration neighborhood in I_R of the floating block $b_F^{i,j}$, and thus the choice of the “arg max” displacement vector $d^{i,j}$.
- By replacing $\mu_F^{k,l}$ by $M_F^{i,j}$, we homogenize by propagation (as the exploration neighborhoods $N_F^{i,j}$ of adjacent floating blocks overlap in I_F) the computation of the similarity field (and of the displacement field) over the entire image. This also justifies the computation of distances between our similarity distributions (see Section 2.2.1).

Note that by estimating A and B on a larger neighborhood, we also make a stronger hypothesis. Whereas, with the classic correlation coefficient, we assume an affine relation between small blocks (equation 1), we here extend that assumption to larger areas: the exploration neighborhoods (equation 3). As it is, this extended hypothesis is still reasonable in our context (better experimental results were obtained with the constrained coefficient). It should however be carefully re-considered for multi-modal registration applications.

Fig. 5 compares the behavior of the constrained correlation coefficient, the classical correlation coefficient and the sum of square difference (SSD) measure on a few sub-images extracted from the myelin-stained histological sections of Fig. 1. We selected a number of sub-images from the histological sections, and computed the similarity map of each measure, where the exploration neighborhoods coincided with the entire sub-images. We used 7 by 7 pixel blocks in a 15 by 15 pixel neighborhood. For each pair (floating sub-image, reference

sub-image), we show the associated similarity map in color. The comparison of the three maps demonstrates the good compromise achieved by the constrained coefficient between the regularity of the SSD and the good precision of the classical correlation coefficient. The benefits of the constrained formulation over the classical one are particularly obvious in the third and fourth column, where the correlation coefficient cannot manage a background composed of low variance blocks (high noise). Finally, comparison between maps (d) and (e) in the left-most column shows that while a high similarity value (dark blue) is obtained with the classical correlation coefficient for a pair of blocks with inverted tissues (white matter in the bottom-left corner/ gray matter in the top-right corner of the floating block and the other way around for the reference block, in black), a lower one (yellow) is obtained with the constrained coefficient.

Fig. 6 displays the similarity distributions for both the constrained and the classical correlation coefficient and the “arg max” displacement field (for the constrained coefficient) for the two consecutive histological sections of the brain of Fig. 1 ($60 \mu m$ myelin stained coronal sections through the occipital cortex). For every site of the floating lattice, a color square shows the similarity measures between the corresponding floating block and the reference blocks in its neighborhood (middle column). The optimal “arg max” displacement field is rendered with arrows whose length and direction are those of the optimal displacement vector associated to the lattice site at which the arrow originates. For visualization purposes, only half of the “similarity squares” are rendered in the similarity maps. Obviously, the similarity squares associated to the constrained coefficient present clear patterns (much clearer than those nonetheless exhibited by the classical coefficient), and, more importantly, conspicuous differences in patterns across the image, that will help the clustering algorithm segment the input images. Additionally, the “arg max” displacement field, although globally rather chaotic, tends to present more homogeneous patterns at a local scale, from which transformations can be estimated in a robust fashion. This will help both cluster the input images and register the extracted sub-images.

2.2 *Extracting the image components*

A dense correspondence field has been computed as described in the previous section. We further assume that the input floating image consists of components that share similar transformation characteristics. We describe here the way the correspondence field is clustered, and how we extract sub-images from the clustered sites. Those sub-images will be later used to estimate local transformations.

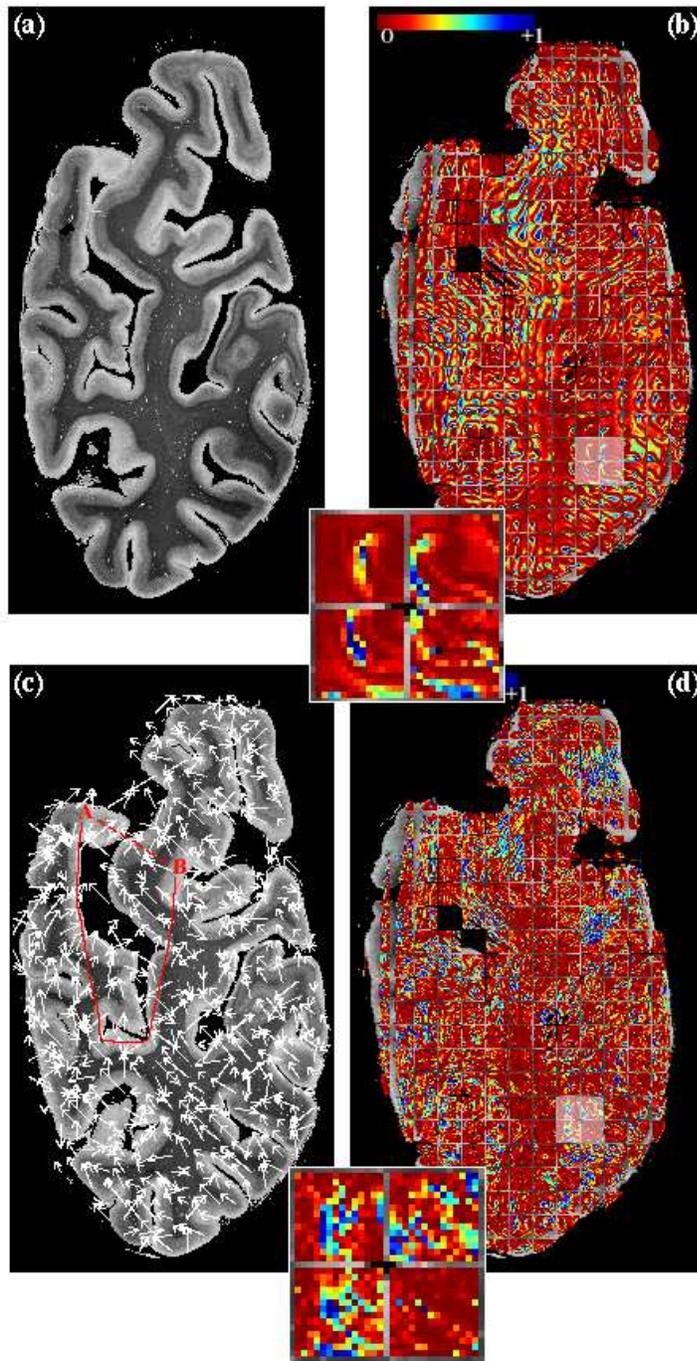


Fig. 6. Dense correspondence field for two consecutive myelin stained histological sections: (a) input reference image; (b) input floating image with superimposed similarity distribution with the *constrained* correlation coefficient; (c) floating image with superimposed arg max displacement field; (d) input floating image with superimposed similarity distribution with the *classical* correlation coefficient. The colorbars show the range of values of the similarity function, for each block. The red lines connecting A and B represent the geodesic (full) and Euclidean (dotted) paths.

2.2.1 Clustering the correspondence field

We are looking for a hierarchical clustering of L_F , that is, a sequence of partitions in which each partition is nested into the next partition in the sequence (Backer, 1995). Cluster analysis (unsupervised learning) essentially consists of sorting a series of multi-dimensional points into a number of groups (clusters) so as to maximize the intra-cluster degree of association and minimize the inter-cluster one. It is particularly well suited here as it behaves adequately even when very little is known about the category structure of the input set of points. That is, it does not require strong hypotheses to be formulated beforehand.

For simplicity's sake, we rewrite L_F as an ordered set of sites:

$$L_F = \{s \text{ such that } \exists!(i, j) \in L_F, s = (i, j)\}_{t=1}^{w_F \cdot h_F}.$$

Our clustering method is adapted from the standard agglomerative hierarchical clustering algorithm described in (Johnson, 1967):

- step 1:** initialize a cluster list by placing each site of L_F in an individual cluster, and let the distance between any two of those clusters be the distance (to be defined) between the sites they contain (the more similar the clusters, the smaller the distance).
- step 2:** find the closest pair of clusters, remove them from the cluster list, merge them into a new single cluster and add the new cluster to the cluster list.
- step 3:** compute the distances between the newly formed cluster and the other ones in the cluster list.
- step 4:** repeat steps 2 and 3 until the desired number of clusters have been reached.

In our case, the number of clusters is specified by the user. Pre-indicators like the Davies-Bouldin index (Davies and Bouldin, 1979) for instance, or the cophenetic correlation coefficient (Backer, 1995) could possibly assist this choice. Section 3.2 discusses the influence of this parameter over the registration quality.

To store the distances between any two clusters in the cluster list at each iteration, we maintain a variable-size distance matrix M which summarizes their proximity (or similarity). At each iteration, M is therefore a square symmetric matrix whose size is the number of clusters in the cluster list at that iteration. The computation of similarity matrix M is the pivotal element of the clustering algorithm. The distance measure between clusters should be consistent with both the model we chose for the input images and the relationships we expect between them.

To define a distance on clusters, we first need a distance on sites. This distance is defined as a linear combination of two distances, a distance between the centroids of the associated blocks and a distance between the associated similarity distributions:

$$D_{site} = \alpha D_{centroid} + (1 - \alpha) D_{distribution} \quad (4)$$

Distance between the centroids. To satisfy the model constraint, we have to ensure that close blocks are more likely to be clustered than blocks far apart. It appears that the Euclidean distance is not the most suitable here. Indeed, if the input images contain several pieces of tissues (*e.g.*, in histological images, they can easily be identified by thresholding) that are potentially non convex, a geodesic distance within each piece will be more convenient to define the proximity of two points from an anatomical point of view. We recall that the geodesic distance between two points is the length of the shortest path that connects these points within a component that must contain them. By convention, when two sites cannot be connected (when they belong to disjoint components), we define the geodesic distance as the Euclidean distance between their associated centroids plus the radius of the input image. Computation of the geodesic distances was done using a variant of the circular propagation algorithm introduced in (Cuisenaire, 1999) which achieves a good trade-off of precision over speed.

Given two sites t and u , their centroid distance is written: $D_{centroid}(t, u) = D_{geodesic}(p_F^t, p_F^u)$.

Distance between similarity distributions. The high expressivity of the similarity distributions described above (Section 2.1), which summarize the similarity landscapes associated with the neighborhoods of the blocks of interest, makes them remarkably well suited to capture the actual differences between those blocks, in spite of noise or decoys, and thus allows for a better discrimination. We use a normalized version ρ of these distributions to ensure that they all have the same overall unit mass (see (Singh and Allen, 1992) for a similar distributional approach in the context of image-flow computation).

Given a site t in L_F , the associated 2-D normalized distribution ρ^t is defined for sites u in the neighborhood N_t of t in L_R by $\rho^t(p_R^u - p_F^t) = \frac{sim(b_F^t, b_R^u)}{\sum_{v \in N_t} sim(b_F^t, b_R^v)}$. Such distributions are depicted in Fig. 6 (middle column).

As a distance between distributions, we chose the Earth mover’s distance (Rubner et al., 1998), a discrete solution to the discrete Monge-Kantorovich mass-transfer problem (Haker et al., 2003). Given the so-called “ground distance” (the distance between elements of the distribution, the Euclidean distance in our case), the Earth mover’s distance (EMD) between two distributions becomes the minimal total amount of work (= mass \times distance) it takes to transform one distribution into the other. As argued by Rubner et al. (1998), this boils down to a bipartite network flow problem, which can

be modeled with linear programming and solved by a simplex algorithm. Among other advantages, the EMD is a true metric, is not impaired by quantization problems (as opposed to histogram-based approaches for instance) and can handle variable-size distributions (our case here). For sites t and u , we obtain: $D_{distribution}(t, u) = D_{EMD}(\rho^t, \rho^u)$.

To summarize, given two sites t and u , their site distance is written:

$$D_{site}(t, u) = \alpha D_{geodesic}(p_F^t, p_F^u) + (1 - \alpha) D_{EMD}(\rho^t, \rho^u) \quad (5)$$

where α is a real-valued positive weight ($0 \leq \alpha \leq 1$).

Once we have a distance between sites, a cluster distance can be defined. We adapted the standard complete link distance (Backer, 1995) to additionally take into account the transformations that can be estimated on the already formed clusters.

Namely, when the size of a cluster reaches a given threshold (we usually take $\theta_{cluster}=20$, even though experiments showed that the value of that threshold does not really impact the quality of the clustering), a rigid or affine transformation can be estimated, in a robust fashion, from the associated set of “arg max” displacement vectors (via a least-square regression combined with an LTS (Least Trimmed Squares) estimator, see Section 2.4). The decision to merge two clusters can then be biased by the agreements between their associated estimated transformations, again as this might indicate that they belong to the same component. Incidentally, when the distance between a cluster with an associated transformation and another one without enough sites to have allowed an estimation must be computed, we choose to return 0. Although theoretically possible, such a case almost never occurs in practice as a hierarchical clustering algorithm tends to aggregate sites in small clusters at early stages before merging them into large ones in subsequent iterations, not leaving single sites un-aggregated very long (see (Backer, 1995) for details). This so-called “chaining effect” also motivates the use of transformation distances.

Given two transformations T^a and T^b , we use a standard symmetric distance:

$$D_{trsf}(T^a, T^b) = \begin{cases} \sum_{i,j} [T^a T^b^{-1} - Id]_{i,j}^2 + \sum_{i,j} [T^{a^{-1}} T^b - Id]_{i,j}^2 & \text{if both } T^a \text{ and } T^b \text{ are defined} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

(where i, j are matrix indices).

Finally, given two clusters of sites $C^a = \{a_1, \dots, a_{n_a}\}$ and $C^b = \{b_1, \dots, b_{n_b}\}$,

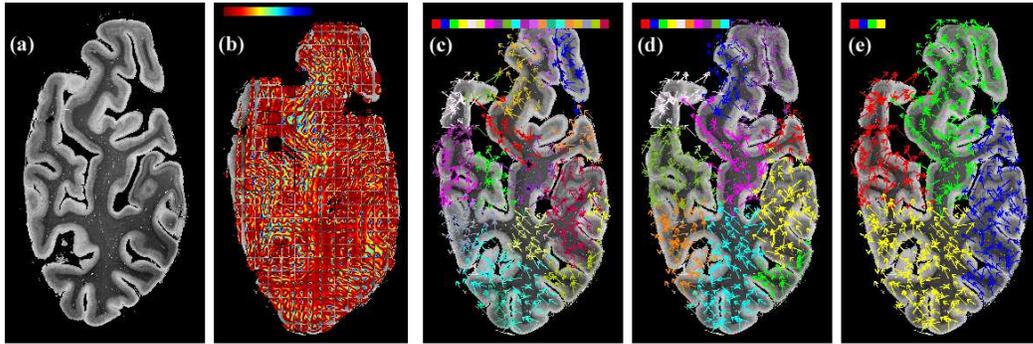


Fig. 7. Clustering the correspondence field: (a) input reference image; (b) input floating image with super-imposed similarity map; (c,d,e) floating image with super-imposed clustered "arg max" field for respectively 20, 10 and 4 clusters.

the cluster distance between them is the longest distance from any site in C^a to any site in C^b (complete-link) *plus* the "transformation distance" wherever it can be computed:

$$D_{cluster}(C^a, C^b) = \beta \max_{i,j} D_{site}(a_i, b_j) + (1 - \beta) D_{trsf}(T^a, T^b) \quad (7)$$

where β is a real-valued positive weight ($0 \leq \beta \leq 1$).

Fig. 7 shows the clustering process of the correspondence field for the two consecutive histological sections of Fig. 1 from 20 down to 4 clusters. It illustrates the nesting property of the hierarchical clustering approach: the clusters in (c) are merged to form the clusters in (d).

As an alternative, we could have used one of the numerous optical flow segmentation algorithms developed in the literature ((Weber and Malik, 1997) for instance) to segment the input images. However, a number of modifications would need to be made to allow for the registration of multi-modal images as they violate the principle of intensity conservation. Additionally, taking into account geodesic distances might also prove difficult. Finally, we believe that better results can be obtained by considering the *complete* similarity map associated with a block instead of choosing *a priori* a single displacement to perform the classification.

2.2.2 Extracting the sub-images

We have described above how we cluster the floating lattice L_F . We detail here how to extract, from the input floating and reference images, pairs of sub-images that will later be registered independently.

Let N_C be the final number of clusters, $C = \{C^1, \dots, C^{N_C}\}$ the cluster partition of L_F , and $\{c_1^i, \dots, c_{n_i}^i\}$ the n_i sites of the i^{th} cluster C^i . We want to build a set of N_C sub-images $\{I_F^i\}_{i=1}^{N_C}$, each of them associated with a single cluster. Given the partition of L_F , a partition of I_F can be built in many ways. For instance, one could compute a Voronoï diagram of the sites c_j^i (or equivalently of their centroids) and draw a partition of the pixels (x, y) of I_F from it. However, our clustering method does not ensure that the borders between clusters are sufficiently precise to adequately represent the sub-images' borders. Moreover, as we are going to use these sub-images to find local transformations, it is often better to choose larger supports to avoid boundary effects.

Consequently, rather than build a partition of I_F from the partition of L_F , we build a covering of I_F , *i.e.*, a set of sub-images that could overlap. To do so, we aggregate in I_F^i the pixels of I_F in the vicinity of the sites of the cluster C^i . We get:

$$I_F^i = \{(x, y) \in I_F \text{ such that } D((x, y), c_j^i) \leq \text{cover_radius} \text{ for some } c_j^i \in C^i\} \quad (8)$$

In practice we use the L_∞ distance. Then, with blocks of size b_{size} associated to the sites, taking $\text{cover_radius} = b_{\text{size}}/2$ we get $I_F^i = \bigcup_j b_{I_F^i}^{c_j^i}$. In our experiments, to ensure a large support, we chose $\text{cover_radius} = 3/4 b_{\text{size}}$.

The corresponding reference sub-images I_R^i are built identically, but with the centroids $p_R^{(k,l)\text{max}}$ of the most similar blocks (see Section 2.1):

$$I_R^i = \{(x, y) \in I_R \text{ such that } D((x, y), c_j^i + d^{c_j^i}) \leq \text{cover_radius}, \text{ for some } c_j^i \in C^i\} \quad (9)$$

Again, we use the L_∞ distance here, with $\text{cover_radius} = b_{\text{size}}$ (a larger extent than that of the floating sub-image) to ensure that I_F^i can be effectively registered against I_R^i .

2.3 Registering the sub-images

Once we have extracted the reference and floating sub-images, we use the robust affine block-matching algorithm described in (Ourselin et al., 2001) to register them, independently, pair by pair. Briefly, this algorithm first estimates a sparse “arg max” displacement field, using a block matching approach (our block-matching algorithm is closely derived from this approach, and we feed both of them the same parameters and similarity measure). From this field, a least square regression extracts a rigid or an affine transformation. As

an illustration, in the rigid case we are looking for R^* and t^* such that:

$$(R^*, t^*) = \arg \min_{R, t} \sum_{i, j} \left\| (p_F^{i, j} + d^{i, j}) - R \cdot p_F^{i, j} - t \right\|^2 \quad (10)$$

where $(p_F^{i, j} + d^{i, j}) - R \cdot p_F^{i, j} - t$ is the residual error and $\|\cdot\|$ the L_2 Euclidean norm.

However, given the rather noisy appearance of the displacement field, an LTS estimator (Least Trimmed Squares, see (Rousseeuw, 1984) for details) is used in place of the least square one to ensure a robust estimation of the transformation. At a glance, instead of minimizing the total sum of the squared residuals (equation 10), a LTS estimator will iteratively minimize the sum of the h smallest squared residuals (we take h at 50% of the number of residuals), to reduce the influence of outliers.

Finally, a better trade-off between robustness and registration precision is achieved via a multi-scale implementation. Note that even though this block-matching algorithm computes displacements (actually, translations) only *locally*, it is able to recover *global* rotations and translations, thanks to its iterative nature. For instance, a robustness study on rat brains sections presented in (Ourselin et al., 2001) demonstrated its ability to recover rotations up to 28 degrees.

Then, for each pair of sub-images $\{I_R^l, I_F^l\}$, $l \in 1 \dots N_C$, we obtain a rigid or an affine transform T^l . Note that since these registrations are robust, the sub-images do not need to perfectly correspond to the anatomically separate components.

2.4 Composing the final images

We selected the method of Little et al. (1997) to compose the final registered floating image. In their approach, a user selects a number of pairs of corresponding rigid structures in the input images along with associated linear transformations (also given by the user). A number of pairs of landmarks further constrain a hybrid affine/non-linear interpolation scheme that acts as a local registration algorithm. This technique then essentially applies user-provided affine transforms to user-defined structures and ensures a smooth interpolation in between them.

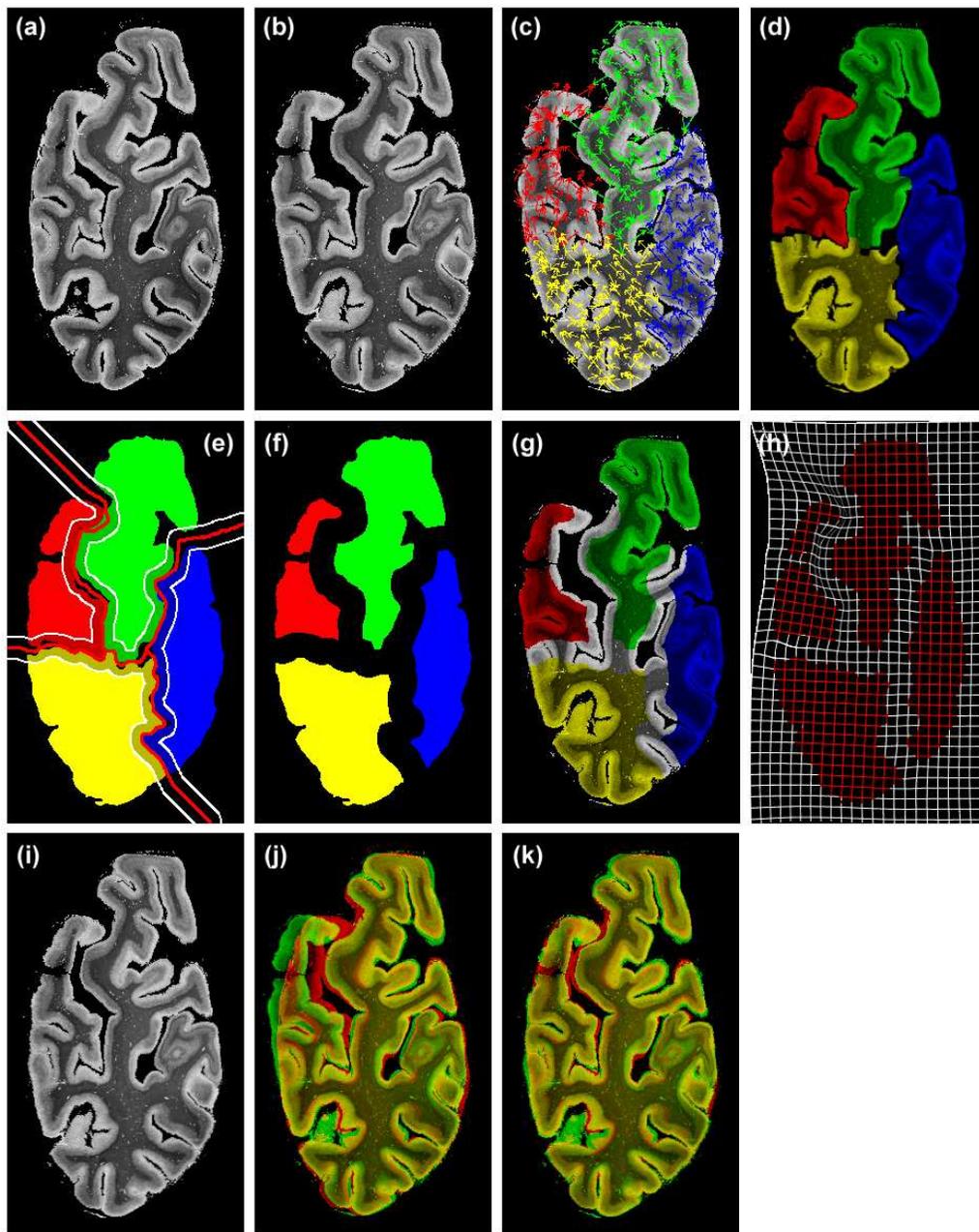


Fig. 8. Piecewise registration of two consecutive myelin-stained histological sections of the human brain: (a) input reference image; (b) floating image; (c) floating image with clustered "arg max" displacement field; (d) registered floating sub-images; (e) binarized sub-images with darkened eroded pixels; (f) eroded registered floating sub-images; (g) input floating image with super-imposed colored eroded floating sub-images; (h) image of a regular grid convected by the associated hybrid affine/non-linear transformation with superimposed transformed eroded floating sub-images (in red); (i) final composed locally registered floating image; (j) superposition of the reference image (red) and of the globally affinely registered floating image (green); (k) superposition of the reference image (red) and of the locally registered floating image (green).

In our application, the set of floating sub-images forms a covering of the input floating image, so we have to erode the sub-images to leave space for interpolation. Furthermore, the floating sub-images must be cut to ensure that they do not overlap, once transformed, as this may impair the interpolation scheme (note the overlap of the red and green sub-images, and the gap between the red and yellow ones in Fig. 8). This erosion algorithm works as follows:

- We first apply the transformations to the floating sub-images ($\forall l \in 1 \dots N_C$, $T^l(I_F^l)$ is the transformed floating sub-image), binarize them (zero for background, one for tissue) and fill in the holes.
- We superimpose the binarized transformed sub-images in a single image J and compute a distance map over the background of that image.
 - A series of morphological operations (erosion) first ensures (on a need for basis) that the $T^l(I_F^l)$ are disjoint.
 - A Euclidean distance map of the background of J is computed.
- A medial axis algorithm then extracts the skeleton of the background of J .
- We compute the distance map of this skeleton.
- We identify in J pixels whose corresponding distance to the skeleton is smaller than a given threshold ν . This ensures a minimum distance of 2ν between any two sub-images. Let N be the set of these pixels. We then remove from the floating sub-images their inverse transformed intersection with N : $\forall l \in 1 \dots N_C$, $\overline{I_F^l} = I_F^l - T^{l-1}(T^l(I_F^l) \cap N)$.

We choose as landmarks the corners of the original images, I_R and I_F (after the initial rigid registration), to further constrain the interpolation scheme, and use the modified Hardy multi-quadric recommended in (Little et al., 1997) as a basis function for interpolation, as this agrees with an affine transform at infinity.

Fig. 8 illustrates the full registration process on the myelin stained histological sections of Fig. 1. The pairs of images were locally rigidly registered with $\alpha = 0.5$, $\beta = 0.5$, $\theta_{cluster} = 20$ and standard parameters for the block matching algorithm (see Table 1 in Appendix) with an extended correlation coefficient and $N_C = 4$. The skeleton of the background of J is shown in red and the $2\nu = 20$ pixel wide band of eroded pixels in darkened colors with white borders in (e).

Clearly, our clustering algorithm adequately isolated in a separate sub-image the floating gyrus (red area in (c)) which was subsequently correctly registered to its counterpart in the reference image. Registering the sub-images affinely instead of rigidly would of course further decreased their discrepancies. However, in the general case, when one suspects only a rigid transformation

between sub-images, opting for an affine registration would only introduce unnecessary over-parameterization which, among other disadvantages, could substantially alter textures.

Note that the entire registration process could easily be included within an iterative multi-scale framework to achieve a better trade-off between accuracy and complexity. Such a framework could also be useful for handling both large-scale and small-scale components.

3 Results

We present here the various experiments we have conducted to assess the performances of our local registration approach. We detail two histological reconstructions (Section 3.1) before discussing in Section 3.2 the influence of the various components and parameters of our registration system on the quality of the registration.

3.1 Reconstruction of a 3-D histological volume

Even though the deformations recovered by our registration method may sometimes be rather subtle, as exemplified by the registration of the pair of myelin-stained sections (Fig. 8), they become a clear nuisance when entire stacks of sections must be aligned.

We aim here to reconstruct a 3-D volume from a series of histological images. Previous work (Ourselin et al., 2001; Malandain et al., 2004) showed that by registering (affinely or rigidly) each pair of consecutive slices in the stack we can recover a geometrically coherent 3-D alignment of the 2-D images and provide a satisfying 3-D reconstruction. However, local rigid/affine piece-wise transformations, as described in the Introduction Section, still impair this registration process and must be accounted for.

As an illustration of the benefits of our piece-wise approach, we first describe the reconstruction of a 3-D histological volume from a series of 70 images. These were $50\mu\text{m}$ thick myelin-stained histological sections of the human brain cut in the V1 area. Reconstruction was performed using the classic pair-wise approach described above. This process requires the choice of a reference section: if we let $Img(ref)$ be this reference section, with $1 < ref < 70$, the reconstruction algorithm is then as follows:

```

for i from ref+1 upto 70
  rigid piece-wise register Img(i) to Img(i-1)
for i from ref-1 downto 1
  rigid piece-wise register Img(i) to Img(i+1)

```

Consequently, the quality of the overall reconstruction depends on the characteristics of the reference section, which should therefore be selected with great care. Indeed, any hole, tear or arbitrary distortion in the reference image is bound to affect the reconstructed 3-D volume. However, in the absence of an external anatomical reference, such a 3-D volume can only be reconstructed upto the transformation associated with the reference image. When an external reference is available, anatomical information can be exploited to guide the registration process (Malandain et al., 2004).

We used here the same parameters as for the registration of the myelin-stained sections of Fig. 8: $\alpha = 0.5$, $\beta = 0.5$, $\theta_{cluster} = 20$ and standard parameters for the block matching algorithm with the constrained correlation coefficient and $N_C = 6$.

Fig. 9 compares the volume reconstructed with our piece-wise approach and that built with the robust rigid registration algorithm we use to register the sub-images (see Section 2.4). Note the greater regularity of the 3-D structures in both the sagittal and axial views of the piece-wise reconstructed volume, with respect to the global rigid volume. A better registration of the separate gyri, illustrated by the better superposition between the red edges and the underlying images, explains this smoother aspect. Visual inspection throughout the 3-D volume confirmed the enhanced continuity of the 3-D structures.

In a second experiment, we registered ten consecutive pairs of Nissl stained histological sections of a mouse brain. Here also, we used $\alpha = 0.5$, $\beta = 0.5$, $\theta_{cluster} = 20$, standard parameters for the block matching algorithm with the constrained correlation coefficient and $N_C = 5$. Fig. 10 shows the original sections (top) and the superposition of the successively registered pairs (bottom). Note how the piecewise approach successfully recovered the various foldings and dislocations, and how the cerebellum which was detached from the cerebrum in section #2 was correctly re-attached with minimal texture changes.

3.2 Sensitivity study

We discuss in this section the dependence of our piecewise registration approach on its various parameters. Standard values for most of these parameters are also reported in Table 1.

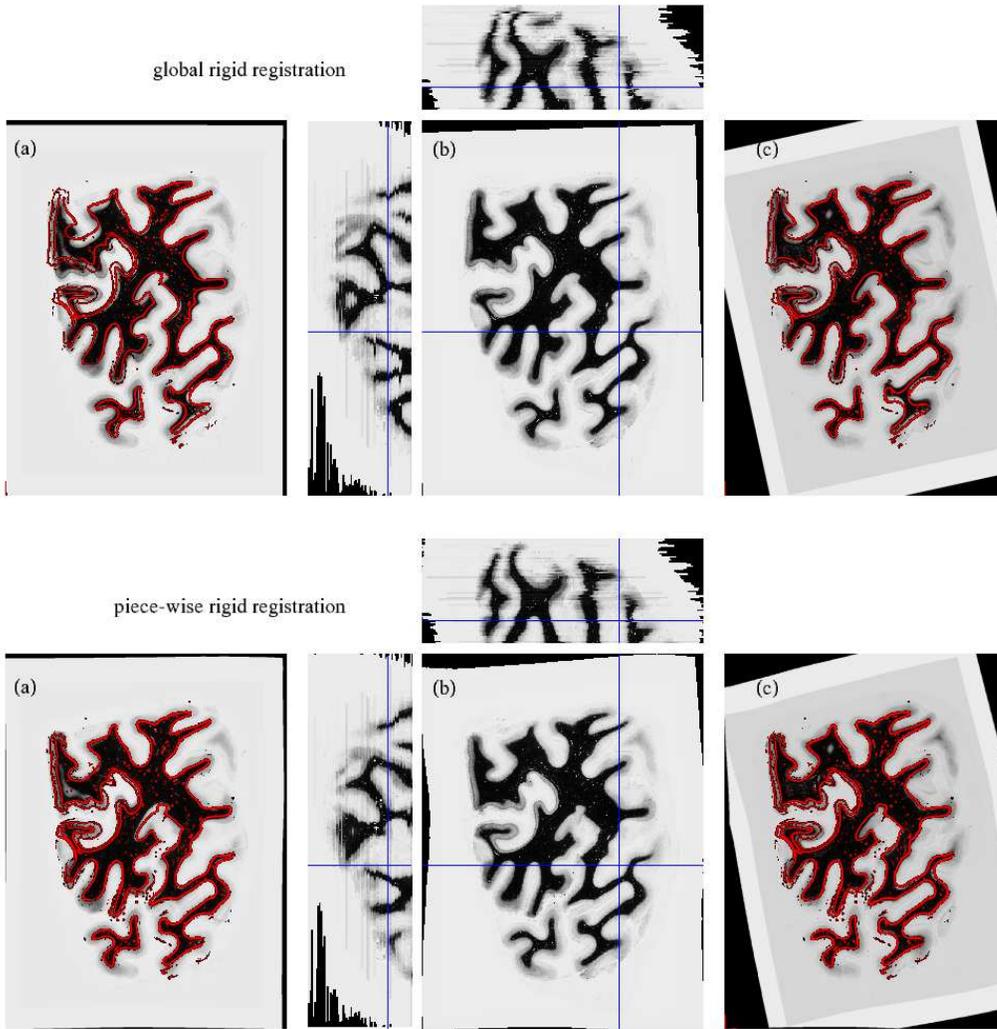


Fig. 9. Reconstruction of a 3-D histological volume with a globally rigid (top) and our piece-wise rigid (bottom) registration algorithm: (b) coronal view (middle) of the 3-D reconstructed volume corresponding to the 51st image of the stack with the associated axial (top) and sagittal (left) views, (a) 50th image (immediately preceding section) with edges of the 51st one superimposed in red, (c) 52nd image (immediately following section) with edges of the 51st image superimposed

3.2.1 Number of clusters

The number of clusters determine the number of degrees of freedom of the overall transformation, and consequently influences the quality of the final match. The middle column ($\nu = 10$) of Fig. 11 qualitatively illustrates the behaviour of the piecewise registration system when the number of clusters varies (between 2 and 8).

We observe that when the specified number of cluster increases above the number of actual components, we get sub-components that are correctly included in the components they come from. The associated transformations are also

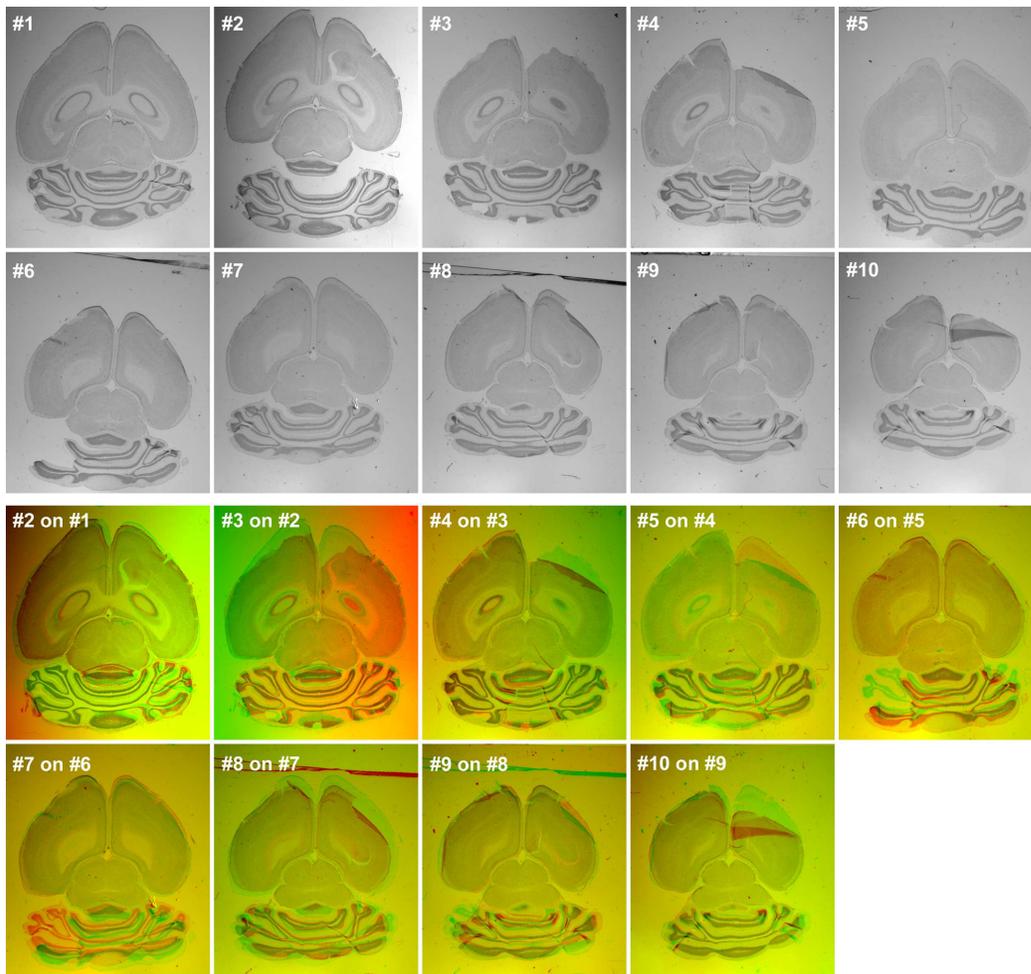


Fig. 10. Successive pair-by-pair registration of Nissl stained mouse brain sections: (top) original sections; (bottom) “#x on #y”: superposition of the previously registered #y image (red) with the image #x registered to it with our piecewise registration approach (green).

part of the transformation of the enclosing component (with minimal error, 2% on average).

This comes as no surprise. Indeed, in a hierarchical clustering, each partition is nested into the next partition in the sequence. Therefore, when the number of desired clusters increases above the actual number of components, the new sub-images (associated with the new clusters) are sub-parts of actual components. Since actual components are supposed to be rigid or affine by definition, affinely registering the new sub-images should produce transformations very similar to the transformations associated with the nesting sub-image. Conversely, when the specified number of cluster drops below the number of actual components, performances decrease and tend towards those of a robust global affine registration. In the limit where a large number of clusters are used, the piecewise registration tends to resemble a block-matching approach

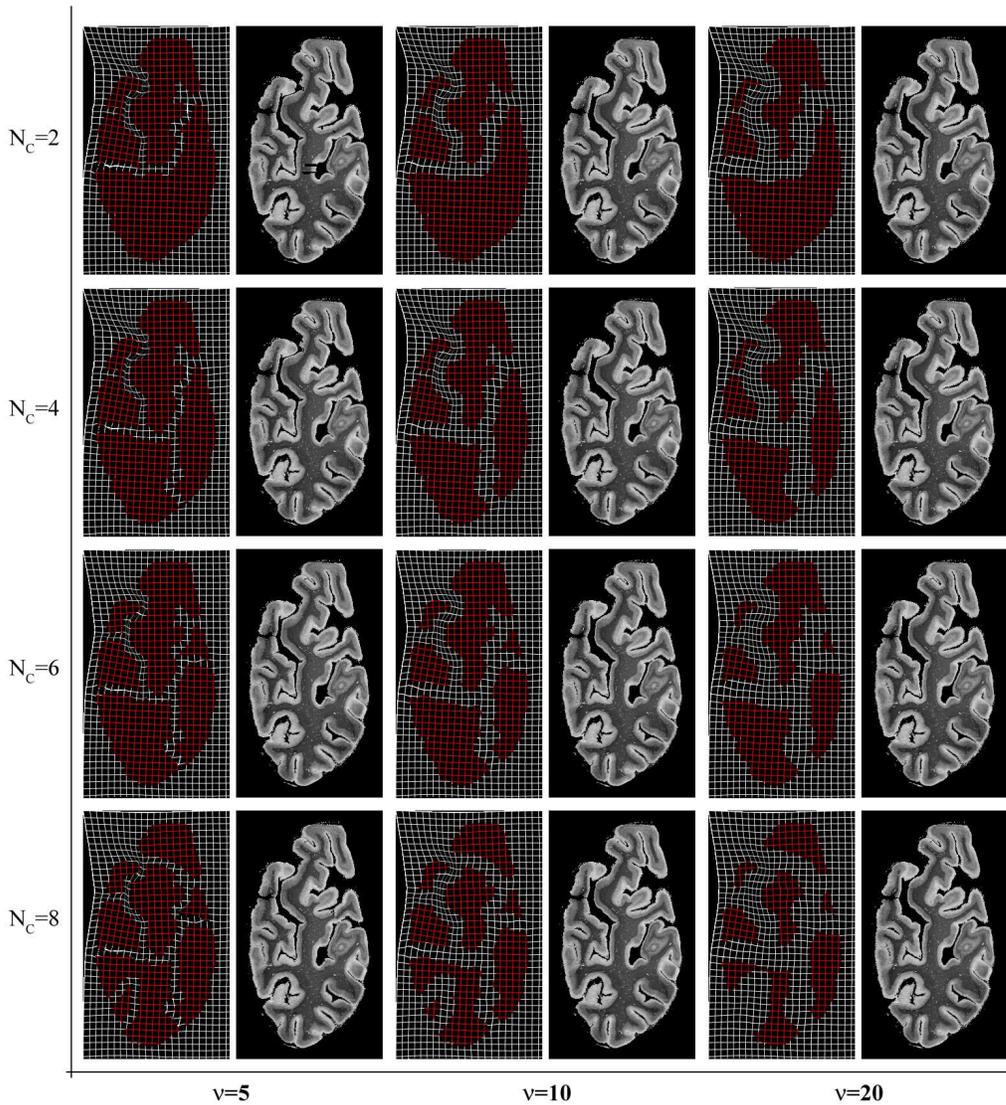


Fig. 11. Sensitivity of the piecewise affine registration method to the number of clusters and to the inter sub-image space. For each choice of parameters, we show the regular grid convected with the composed transformation with the superimposed transformed eroded sub-images (left) and the final composed registered floating image (right).

with the notable difference that complete affine (or rigid) transformations are estimated instead of simple translations at each block.

Note, that even though obtaining a *perfect* clustering may not actually be necessary, our clustering algorithm suffers from its inability to modify clusters that have already been created: namely, clusters can only be aggregated to form larger clusters, they cannot be re-cut or broken down. This is unfortunate as better cluster boundaries could probably be obtained based on the

associated cluster transformations, which are computed only after the clusters have reached a sufficient size to ensure a correct estimation. The use of stochastic clustering approaches or the introduction of uncertainty in the clustering process may alleviate this issue.

3.2.2 Block-matching parameters

As argued above, the quality of the similarity map computed between the floating and the reference image depends on the block matching parameters: similarity measure, size of the exploration neighborhood, step in that neighborhood, size of blocks, etc. We already studied in Section 2.1.2 how the use of a constrained correlation coefficient helped increase both the homogeneity and the precision of the similarity map. Our block-matching algorithm is similar to that detailed in (Ourselin et al., 2001), (Ourselin, 2002) and (Malandain et al., 2004) to which we report the reader for a detailed sensitivity investigation of the other parameters.

Note that the selection of the “arg max” displacement vector is clearly sub-optimal and somewhat arbitrary when many blocks in the reference exploration neighborhood have close associated similarity measures. Better estimation of the sub-image transformations might be obtained by taking into account the full spectrum of displacement vectors, together with their similarity measures.

3.2.3 Parameters of the registration algorithm for the sub-images

The robust affine block-matching algorithm used in Section 2.3 to register the sub-images too requires that a number of parameters be set. In addition to block matching parameters similar to ours, the cut-off of the robust estimator, the parameters controlling the multi-scale system within which it works, or the parameters of the various variance and intensity tests performed on the block to discard them from the robust estimation must be managed. Again, we report the reader to (Ourselin, 2002) for details about their influence on the registration performances.

3.2.4 Composition parameters

The choice of the amount of space to leave in between structures (2ν pixels) depends on the input images and should be set accordingly. However, there is no general prescription for selecting a good value for ν which would work well for all images and, within a single image, for all sub-images. Clearly, as the amount of space decreases, the band in between sub-images becomes more stretched (which might induce substantial textural changes).

Fig. 11 illustrates the relationships between the selected number of clusters and the space to leave between sub-images. As a rule of thumb, the greater the number of clusters is, the smaller ν should be, as the size of the components tend to decrease. Note that even though we restricted ourselves to a constant value for ν across the image in this study, a variable ν should increase the overall registration performances. Namely, for each pair of sub-images, ν should depend on the differences between the transformations associated to the sub-images. If the difference is large, then a large ν is to be used to prevent large distortions of the textures of the underlying tissues. Conversely, when the two sub-images share very similar transformations (even though these transformation might be large with respect to the identity), ν can be much smaller (for instance, for components A and B in Fig. 11).

The choice of ν could also depends on the nature of the underlying material: large deformations will not impair the quality of the registration if they occur over the black background. Use of tissue/background segmentation maps could then also help choose locally an optimal value for ν .

3.3 *Specificity versus genericity*

As argued in Section 1.2, the ill-posed nature of the medical image registration problem makes it difficult to evaluate *a posteriori* the performance of any given registration algorithm. This difficulty is essentially due to that of characterizing an appropriate measure of performance. We submit that since we cannot trust the registration result, we must at least have confidence in the registration method. A safe approach then consists of devising a registration model which mimics, at the desired scale, the actual transformation undergone by the imaged tissues. Prior medical knowledge is then pivotal in helping design such a specific registration method. This concept of trust becomes especially relevant in the context of high throughput registration (our case here since reconstructing a volume requires registering a large number of images) where the quality of each registration results affects that of the subsequent ones.

For histological reconstruction, our piecewise model adequately recovers the arbitrarily large transformations introduced during the glass mounting step, while allowing flexible deformations at the interface between anatomical regions. As mentioned before, a number of fluid or elastic approaches could be applied to the same histological reconstruction problem with most certainly very similar visual results once their parameters have been tuned. However, given their inherently generic nature, only somewhat indirect control could be exerted over the registration process. It is thus difficult to weight the contributions of the various parts of the image in the process and how they interact

with the registration algorithm. In particular, the resulting transformation may exhibit local deformations in unexpected locations. For instance, one might wonder if and how the background was discarded, or whether particularly bright lesions or tumors did not heavily bias the joint histogram from which mutual information or correlation ratio were computed, etc.

Incidentally, our piecewise approach is also more economical in terms of degrees of freedom than a typical fluid one.

4 Conclusion

We have presented an automated 3D volume reconstruction methodology based on the pair-by-pair registration of consecutive 2D images. Our method relies on the modeling of the actual transformation undergone by the imaged tissues in each section. In view of the ill-posed nature of image registration, we believe that specific registration techniques, which closely model the actual physical transformations, are more suitable than generic all-purpose algorithms for reconstruction.

In the case of histological sections, our system builds complex spatial transformations by elastically interpolating between rigid or affine transforms that are locally defined on pairs of sub-images. These sub-images represent geometrically, and often anatomically, coherent components. They are automatically extracted from an initial displacement field computed between the images to be registered (by contrast with other approaches (Little et al., 1997)).

The use of a hierarchical clustering approach and a similarity distribution distance proved very promising: while the distribution distance can effectively deal with noise and textural issues to discriminate between image blocks, our clustering algorithm manages to extract the expected sub-images.

Finally, preliminary experiments have also shown that the developed approach worked well on multi-modal registration cases (Pitiot et al., 2003b,a).

Acknowledgement

This work was partially funded by European project MAPAWAMO (ref. QL3-CT-2000-30161). Grant support for AP and PT was provided by a P41 Resource Grant from the National Center for Research Resources (RR13642), the National Library of Medicine (LM05639), the National Institute for Biomedical Imaging and Bioengineering (EB001561) and by a Human Brain Project

grant to the International Consortium for Brain Mapping, funded jointly by NIMH and NIDA (MH52176). AP was also supported by the INRIA associated team grant.

The authors would like to thank J. Annese from the LONI laboratory for providing the human brain sections, and Y.Z. Wadghiri, J. Blind and D. Turnbull from the Skirball Institute at NYU for the mouse brain images.

Appendix

Standard values

Table 1 reports standard values for the various parameters of the algorithms we use in our approach.

References

- Arsigny, V., Pennec, X., Ayache, N., 2003. Polyrigid and Polyaffine Transformations: A New Class of Diffeomorphisms for Locally Rigid or Affine Registration. In: Proc. of Medical Image Computing and Computer-Assisted Intervention (MICCAI'03). pp. 829–837.
- Ashburner, J., Friston, K. J., 1999. Nonlinear Spatial Normalization Using Basis Functions. *Human Brain Mapping* 7 (4), 254–266.
- Backer, E., 1995. *Computer-Assisted Reasoning in Cluster Analysis*. Prentice Hall.
- Cachier, P., 2002. *Recalage Non Rigide d'Images Mdicales Volumiques - Contribution aux Approches Iconiques et Gomtriques*. Ph.D. thesis, Ecole Centrale des Arts et Manufactures.
- Christensen, G. E., 1999. Consistent Linear-Elastic Transformations for Image Matching. In: Proc. of Information Processing in Medical Imaging (IPMI'99). pp. 224–237.
- Cohen, F., Yang, Z., Huang, Z., Nissanov, J., 1998. Automatic Matching of Homologous Histological Sections. *IEEE Transaction on Biomedical Engineering* 45 (5), 642–649.
- Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., Marchal, G., 1995. Automated multimodality image registration using information theory. In: Proc. of Information Processing in Medical Imaging (IPMI'95). pp. 263–274.
- Collins, D. L., Evans, A. C., Holmes, C., Peters, T. M., 1995. Automatic 3D Segmentation of Neuro-anatomical Structures from MRI. In: Proc. of Information Processing in Medical Imaging (IPMI'95). pp. 139–152.

- Collins, D. L., Zijdenbos, A. P., Paus, T., Evans, A. C., 2003. Use of Registration for Cohort Studies. In: Medical Image Registration.
- Cuisenaire, O., 1999. Distance Transformations: Fast Algorithms and Applications to Medical Image Processing. Ph.d. thesis.
- Davatzikos, C., 1997. Spatial Transformation and Registration of Brain Images using Elastically Deformable Models. *Computer Vision and Image Understanding* 66 (2), 207–222.
- Davies, D. L., Bouldin, D. W., 1979. A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1, 224–227.
- Dengler, J., 1991. Estimation of Discontinuous Displacement Vector Fields with the Minimum Description Length Criterion. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'91)*. pp. 276–282.
- Deverell, M., Salisbury, J., Cookson, M., Holman, J., Dykes, E., Whimster, F., 1993. Three-Dimensional Reconstruction: Methods of Improving Image Registration and Interpretation. *Analytical Cellular Pathology* 5 (5), 253–263.
- Feldmar, J., Ayache, N., 1996. Rigid, Affine and Locally Affine Registration of Free-Form Surfaces. *The International Journal of Computer Vision* 18 (2).
- Ford-Holevinski, T., Castle, M., Herman, J., Watson, S., 1991. Microcomputer-based three-dimensional reconstruction of in situ hybridization autoradiographs. *Journal of Chemical Neuroanatomy* 4 (5), 373–385.
- Gee, J. C., Reivich, M., Bajcsy, R., 1993. Elastically Deforming 3D Atlas to Match Anatomical Brain Images. *Journal of Computer Assisted Tomography* 17 (2), 225–236.
- Goldszal, A., Tretiak, O., Hand, P., Bhasin, S., McEachron, D., 1995. Three-Dimensional Reconstruction of Activated Columns from 2- ^{14}C deoxy-D-glucose Data. *Neuroimage* 2 (1), 9–20.
- Goldszal, A., Tretiak, O., Liu, D., Hand, P., 1996. Multimodality Multidimensional Image Analysis of Cortical and Subcortical Plasticity in the Rat Brain. *Annals of Biomedical Engineering* 24 (3), 430–439.
- Haker, S., Angenent, S., Tannenbaum, A., 2003. Minimizing Flows for the Monge-Kantorovich Problem. *SIAM Journal of Mathematical Analysis* 35 (1), 61–97.
- Haney, S., Thompson, P., Cloughesy, T., Alger, J., Toga, A., 2001. Tracking Tumor Growth Rates in Patients with Malignant Gliomas: A Test of Two Algorithms. *American Journal of Neuroradiology* 22 (1), 73–82.
- Hibbard, L., Hawkins, R., 1988. Objective Image Alignment for Three-Dimensional Reconstruction of Digital Autoradiograms. *Journal of Neuroscience Methods* 26 (1), 55–74.
- Humm, J., Macklis, R., Lu, X., Yang, Y., Bump, K., Beresford, B., Chin, L., 1995. The Spatial Accuracy of Cellular Dose Estimates Obtained from 3D Reconstructed Serial Tissue Autoradiographs. *Physics in Medicine and Biology* 40 (1), 163–180.
- Jain, A. K., 1981. Image Data Compression: A Review. *Proceedings of the IEEE* 69 (3), 349–389.

- Johnson, S. C., 1967. Hierarchical Clustering Schemes. *Psychometrika* 32, 241–254.
- Jolesz, F. A., Nabavi, A., Kikinis, R., 2001. Integration of Interventional MRI with Computer-Assisted Surgery. *Journal of Magnetic Resonance Imaging* 13 (1), 69–77.
- Kay, P., Robb, R., Bostwick, D., Camp, J., 1996. Robust 3-D Reconstruction and Analysis of Microstructures from Serial Histologic Sections, with Emphasis on Microvessels in Prostate Cancer. In: *Proc. of Visualisation in Biomedical Computing*. Hamburg (Germany), pp. 129–134.
- Kim, B., Frey, K., Mukhopadhyay, S., Ross, B., Meyer, C., 1995. Co-Registration of MRI and Autoradiography of Rat Brain in Three-Dimensions Following Automatic Reconstruction of 2D Data Set. In: *Proc. of Computer Vision, Virtual Reality and Robotics in Medicine*. pp. 262–266.
- Little, J. A., Hill, D. L. G., Hawkes, D. J., 1997. Deformations Incorporating Rigid Structures. *Computer Vision and Image Understanding* 66 (2), 223–232.
- Maintz, J. B. A., Viergever, M. A., 1998. A Survey of Medical Image Registration. *Medical Image Analysis* 2 (1), 1–36.
- Malandain, G., Bardinet, E., Nelissen, K., Vanduffel, W., 2004. Fusion of autoradiographs with an MR volume using 2D and 3D linear transformations. *Neuroimage* 23 (1), 111–127.
- McInerney, T., Terzopoulos, D., 1996. Deformable Models in Medical Image Analysis: A Survey. *Medical Image Analysis* 1 (2), 91–108.
- Ourselin, S., 2002. *Recalage d’Images Médicales par Appariement de Régions - Application à la Construction d’Atlas Histologiques 3D*. Ph.d. thesis, Université de Nice Sophia-Antipolis.
- Ourselin, S., Roche, A., Subsol, G., Pennec, X., Ayache, N., 2001. Reconstructing a 3D Structure from Serial Histological Sections. *Image and Vision Computing* 19 (1-2), 25–31.
- Pitiot, A., Bardinet, E., Thompson, P., Malandain, G., 2003a. Automated Piecewise Affine Registration of Biological Images. Research report RR-4866, INRIA.
- Pitiot, A., Malandain, G., Bardinet, E., Thompson, P., 2003b. Piecewise Affine Registration of Biological Images. In: *Proc. of Workshop on Biomedical Image Registration (WBIR’03)*. pp. 91–101.
- Pitiot, A., Toga, A., Ayache, N., Thompson, P., 2002. Texture based MRI segmentation with a two-stage hybrid neural classifier. In: *Proc. of World Congress on Computational Intelligence / INNS-IEEE International Joint Conference on Neural Networks WCCI-IJCNN’02*. pp. 2053–2058.
- Rangarajan, A., Chui, H., Mjolsness, E., Pappu, S., Davachi, L., Goldman-Rakic, P., Duncan, J., 1997. A Robust Point-Matching Algorithm for Autoradiograph Alignment. *Medical Image Analysis* 1 (4), 379–398.
- Rey, D., Subsol, G., Delingette, H., Ayache, N., 2002. Automatic Detection and Segmentation of Evolving Processes in 3D Medical Images: Application to Multiple Sclerosis. *Medical Image Analysis* 6 (2), 163–179.

- Roche, A., Malandain, G., Ayache, N., 2000. Unifying Maximum Likelihood Approaches in Medical Image Registration. *International Journal of Imaging Systems and Technology: Special Issue on 3D Imaging* 11 (1), 71–80.
- Rousseeuw, P., 1984. Least Median of Squares Regression. *Journal of the American Statistical Association* 79, 871–880.
- Rubner, Y., Tomasi, C., Guibas, L., 1998. A Metric for Distributions with Applications to Image Databases. In: *Proc. of International Conference on Computer Vision (ICCV'98)*. pp. 59–66.
- Rydmark, M., Jansson, T., Berthold, C., Gustavsson, T., 1992. Computer Assisted Realignment of Light Micrograph Images from Consecutive Sections Series of Cat Cerebral Cortex. *Journal of Microscopy* 165, 29–47.
- Singh, A., Allen, P., 1992. Image-Flow Computation: an Estimation-Theoretic Framework and a Unified Perspective. *Computer Vision, Graphics, and Image Processing: Image Understanding* 56 (2), 152–177.
- Studholme, C., Hill, D. L. G., Hawkes, D. J., 1996. Incorporating Connected Region Labelling into Automated Image Registration Using Mutual Information. In: *Proc. of IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA'96)*. pp. 23–31.
- Taubin, G., 1993. An Improved Algorithm for Algebraic Curve and Surface Fitting. In: *Proc. of International Conference on Computer Vision (ICCV'93)*. pp. 658–665.
- Thompson, P., Woods, R., Mega, M., Toga, A., 2000. Mathematical/Computational Challenges in Creating Deformable and Probabilistic Atlases of the Human Brain. *Human Brain Mapping* 9 (2), 81–92.
- van den Elsen, P., Pol, E., Viergever, M., 1993. Medical Image Matching - a Review with Classification. *IEEE Transactions on Engineering in Medicine and Biology* 12 (4), 26–39.
- Weber, J., Malik, J., 1997. Rigid Body Segmentation and Shape Description from Dense Optical Flow Under Weak Perspective. *IEEE transactions on Pattern Analysis and Machine Intelligence* 19 (2), 139–143.
- Wells, W. M., Viola, P., Atsumi, H., Nakajima, S., 1996. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis* 1 (1), 35–51.
- Woods, R. P., Grafton, S. T., Holmes, C. J., Cherry, S. R., Mazziotta, J. C., 1998. Automated Image Registration: I. General Methods and Intramodality Validation. *Journal of Computer Assisted Tomography* 22 (1), 141–154.
- Zhao, W., Young, T., Ginsberg, M., 1993. Registration and Three-Dimensional Reconstruction of Autoradiographic Images by the Disparity Analysis Method. *IEEE transactions on Medical Imaging* 12 (4), 782–791.

| Parameter | Algorithm | Name | Typical value | Comments |
|-----------------------------------|--------------------|--------------------|-------------------------------|---|
| block size | block matching | b_{size} | 6×6 pixel | values between 4 and 7 give similar results. |
| size of $N_R^{i,j}$ | block matching | $none_{\dot{z}}$ | 20×20 pixel | depends on maximal distance between corresponding components. |
| lattices step size | block matching | $none_{\dot{z}}$ | 5×5 pixel | little/no impact on the registration quality. |
| similarity measure | block matching | sim | cor or $ccor$ | depends on the modality of the images to be registered. |
| LTS cut-off | block matching | h | 50% | always 50% |
| centroid weight | clustering | α | 0.5 | values between 0.3 and 0.7 give similar results. |
| transformation distance weight | clustering | β | 0.5 | values between 0.3 and 0.6 give similar results. |
| transformation distance threshold | clustering | $\theta_{cluster}$ | 20 | values between 15 and 30 give similar results. |
| number of clusters | clustering | N_C | 6 | [see Section “Number of clusters”] |
| covering radius | extraction | $cover_{radius}$ | $\frac{3}{4}b_{size}=5$ pixel | little/no impact on the registration quality. |
| space between eroded structures | sub-images erosion | ν | 10 pixel | values between 5 and 20 give similar results. |

Table 1
Standard values for algorithm parameters