HAL
open science

Automated Piecewise Affine Registration of Biological Images

Alain Pitiot, Éric Bardinet, Paul M. Thompson, Grégoire Malandain

To cite this version:

Alain Pitiot, Éric Bardinet, Paul M. Thompson, Grégoire Malandain. Automated Piecewise Affine Registration of Biological Images. RR-4866, INRIA. 2003. inria-00071717

HAL Id: inria-00071717

https://inria.hal.science/inria-00071717v1

Submitted on 23 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Automated Piecewise Affine Registration of Biological Images

Alain Pitiot — Éric Bardinet — Paul M. Thompson — Grégoire Malandain

N° 4866

July 3, 2003

THÈME 3

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4866--FR+ENG

Automated Piecewise Affine Registration of Biological Images

Alain Pitiot* , Éric Bardinet† , Paul M. Thompson‡ , Grégoire Malandain

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projets Epidaure

Rapport de recherche n° 4866 — July 3, 2003 — 28 pages

Abstract:

This report tackles the registration of 2D biological images (histological sections or autoradiographs) to 2D images from the same or different modalities (*e.g.*, histology or MRI). 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. A hierarchical clustering algorithm then automatically partitions this 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 under a variety of conditions, and discuss examples using simulated and real medical images, including MRI, autoradiography, histology and cryosection data. We also detail the reconstruction of a 3-D volume from a series of 2-D histological sections and compare it against a reconstruction obtained with a global rigid approach.

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

* Alain Pitiot is also with the LONI laboratory, UCLA School of Medicine, Los Angeles.
† Éric Bardinet is now with CNRS UPR 640 - LENA, hôpital La Pitié-Salpétrière, Paris.
‡ Paul Thompson is with the LONI laboratory, UCLA School of Medicine, Los Angeles.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 —Télécopie : +33 4 92 38 77 65

Recalage automatique affine par morceaux des images biologiques

Résumé :

Ce rapport s'intéresse au recalage d'images biologiques en 2 dimensions (coupes histologiques ou autoradiographiques) vers des images de même dimension acquises suivant une modalité soit similaire (histologie) soit différente (IRM). Le processus d'acquisition de ces images induit typiquement des transformations *composites* que nous modélisons sous la forme d'un ensemble de transformations locales, rigides ou affines, plongées dans une transformation élastique. Nous proposons ici une méthode de recalage directement inspirée de ce modèle.

Etant donnée une paire d'images en entrée, nous calculons en premier lieu un champ dense de similarité via une méthode d'appariement par bloc. Un algorithme de partitionnement hiérarchique segmente ensuite automatiquement ces champs en un ensemble de classes d'où l'on extrait des paires indépendantes de sous-images. Notre algorithme de partitionnement s'appuie sur la métrique sur les distributions "Earth mover" et est par ailleurs guidé par l'estimation robuste aux moindres carrés des transformations associées à chaque classe. Ces paires de sous-images sont ensuite recalées, rigidement ou affinement, de façon indépendante. Enfin, un schéma d'interpolation hybride affine/non linéaire compose l'image recalée en sortie.

Nous analysons dans ce rapport le comportement de notre approche dans un ensemble de situations, et discutons d'exemples de recalage sur des images simulées et sur des images biologiques (IRM, autoradiographies, coupes histologiques et image cryogènes). Nous détaillons également la reconstruction d'un volume 3-D à partir d'un ensemble de coupes histologiques 2-D, volume que nous comparons à celui obtenu avec une méthode de recalage rigide globale.

Mots-clés : recalage, clustering, reconstruction, histologie, IRM

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 [9, 35], while mono- or multi-modal registration of the same patient's data can help determine the nature of an anomaly [20] or monitor the evolution of a tumor [17] or other disease process [27].

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 [5, 25, 4].

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 3 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 (see [24] for an extensive review of medical image registration): among others, one finds linear transformations (rigid, affine) and non-linear transformations (polynomial [38], polyaffine [14, 1], elastic [11, 15] or fluid [6]). Similarly, many similarity criteria have been presented: Studholme *et al.* [33] use normalized mutual information, Collins *et al.* [8] cross-correlation, Roche *et al.* [29] the correlation ratio, Ashburner *et al.* [2] the squared intensity difference, etc. Optimization algorithms range from the straightforward Powell method [7] to sophisticated multi-scale Levenberg-Marquardt techniques [34] or stochastic search [37].

Motivation For iconic (*i.e.*, intensity-based) methods, optimal similarity measures can be derived from a careful analysis of the expected relationships between the input images [29], with different hypotheses leading to different measures. Similarly, *a priori* knowledge about the acquisition process for biological images may 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

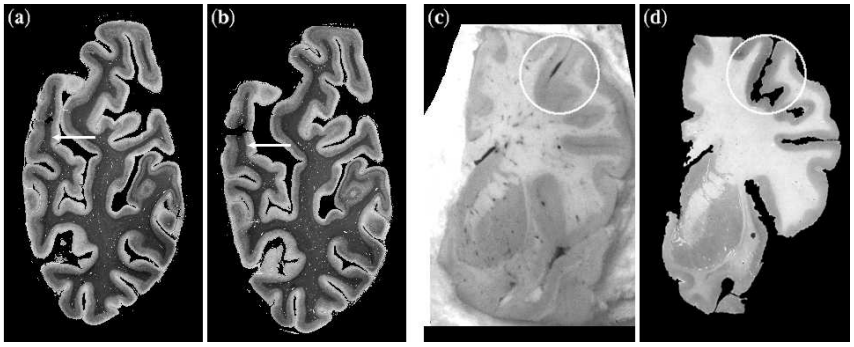

Figure 1: Two consecutive myelin-stained histological sections of the human brain (a & b); Human brain cryosection (c) and its associated Nissl-stained section (d). White arrows and circles indicate moving gyri.

section. In brain sections for instance, each gyrus (compare white arrows in Figure 1.a & b, and white circles in c & d) undergoes an affine transformation (due to successive manipulations) relatively independent from those of other gyri. Consequently, even though a large variety of transformation spaces have been discussed in the literature, their functional form may not reflect our specific needs.

Discussions with neuro-anatomists and histologists prompted us to model the composite transformation yielded by this chain of physical 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 to medical as well as biological images (our primary motivation here). For instance, abdominal or torso MRIs 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 or affine transformations cannot adequately handle such a case. Also, a single rigid transformation would 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 ensure that specific components (*e.g.*, bones) will be only rigidly transformed.

Prior work To alleviate these issues, a few authors have developed local registration techniques, where the input images are divided into a number of smaller sub-images, and a transformation is associated with each. An automatic hierarchical elastic image registration technique is presented in [22]. The initial 2-D images are partitioned into quad-tree structures. At each level of the quad-tree, the floating sub-images are independently registered to their counterparts in the reference image, before being merged via thin-plate spline interpolation. However, this technique cannot apply a transform selectively to a specific region

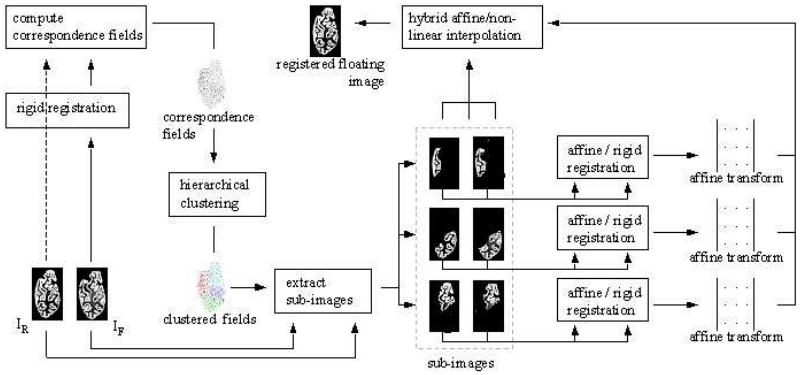

Figure 2: Overview of our local registration approach

whose boundary does not coincide with the quad-tree grid. In [23], Little *et al.* describe an approach where 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 method, even though it fits our needs, still relies on interactive specification of the components to be rigidly matched.

Proposed approach This paper addresses the problem of *automatically* registering two images, when the images consist of a number of independent components, subject to linear transformations. Figure 2 illustrates our 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. A hierarchical clustering algorithm then partitions the correspondence field into a number of classes from which we extract independent pairs of sub-images. 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. Finally, the hybrid affine/non-linear interpolation scheme described in [23] is used to compose the registered floating image.

Note that in contrast to that of Little *et al.*, our approach is fully automated. It does not require *a priori* structures, landmarks, or transformation matrices to be specified, but will estimate them from a dense similarity map.

We detail our method in Section 2, and then discuss in Section 3 registration results on a variety of biomedical images, and 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 [18] to compute the correspondence field.

2.1.1 Block-matching algorithm

We associate with I_F and I_R two rectangular lattices $L_F = \{(i,j) \in [1,\ldots,w_F] \times [1,\ldots,h_F]\}$ and $L_R = \{(i',j') \in [1,\ldots,w_R] \times [1,\ldots,h_R]\}$ 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. 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.

The block-matching algorithm works as follows (see Figure 3): 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 $b_{I_F}^{i,j}$ 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 $sim\left(b_F^{i,j}, b_R^{k,l}\right)$ 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.

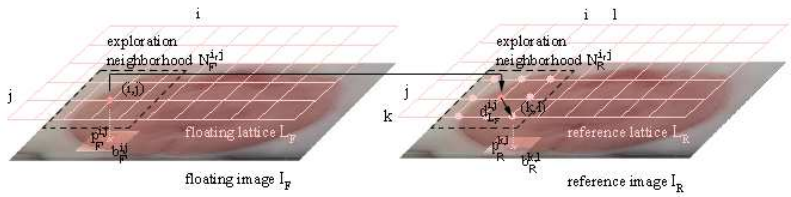

Figure 3: Block-matching algorithm

- 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 respectively (see [9] and [13] 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.

The Appendix Section provides an exhaustive list of the parameters of all the algorithms used in this study, along with comments on their impact on the overall registration quality and 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 [26], 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 Extended correlation coefficient

A ubiquitous choice for image registration [29], 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:

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

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 re-written (see [28] for instance):

$$1 - \mathrm{cor}\left(b_F^{i,j}, b_R^{k,l}\right)^2 = \frac{\min_{A,B} E\left[b_F^{i,j} - A.b_R^{k,l} - B\right]^2}{\mathrm{var}\left(b_F^{i,j}\right)} \tag{2}$$

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 stems from 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 [29, 26].

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:

$$\mathrm{ecor}\left(b_F^{i,j}, b_R^{k,l}\right)^2 = \frac{\sum_{u,v}\left(b_F^{i,j}(u,v) - M_F^{i,j}\right).\left(b_R^{k,l}(u,v) - M_R^{k,l}\right)}{\sum_{u,v}\left(b_F^{i,j}(u,v) - M_F^{i,j}\right)^2.\sum_{u,v}\left(b_R^{k,l}(u,v) - M_R^{k,l}\right)^2} \tag{3}$$

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

By estimating A and B on a larger neighborhood, we homogenize the computation of the similarity field (and of the displacement field) over the entire image. Note that 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 extended coefficient). It should however be carefully re-considered for multi-modal registration applications (we had to revert to the classic coefficient for the registration of autoradiographies to MRI data in the monkey brain, see Figure 8).

Furthermore, by homogenizing the similarity measures, we also justify the computation of distances between our similarity distributions (see Section 2.2.1).

Figure 4 displays the similarity distributions and the "arg max" displacement field for two consecutive histological sections of the brain (60 μm myelin stained coronal sections

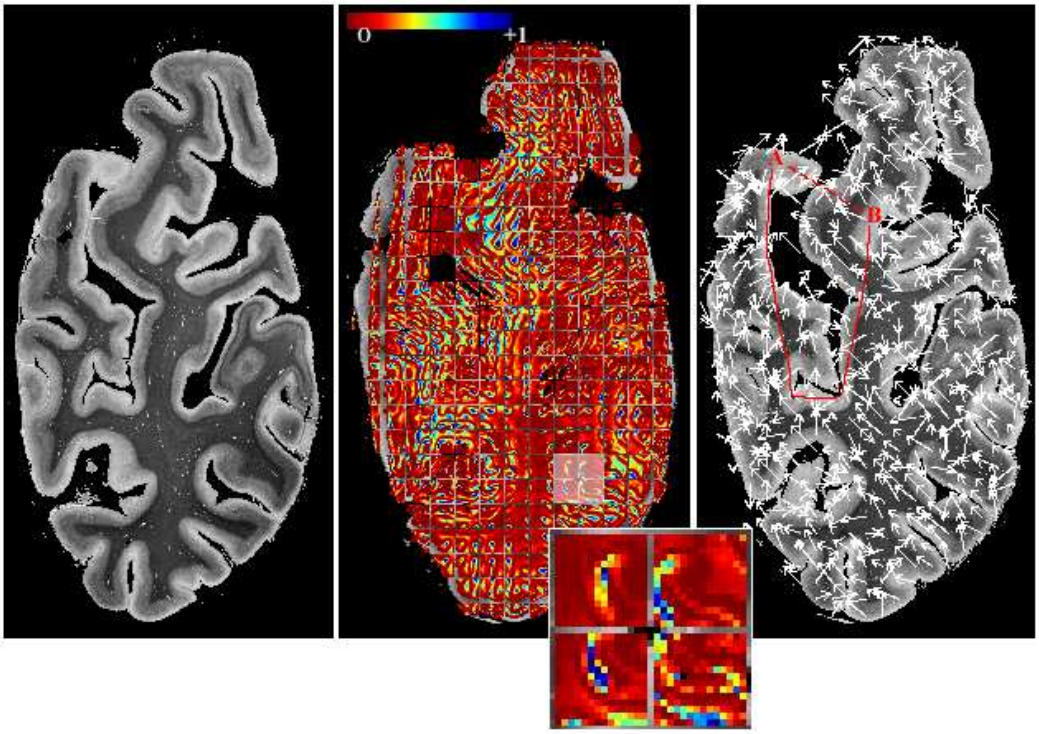

Figure 4: Dense correspondence field: input reference image (left column), input floating image with superimposed similarity distribution (middle column) and superimposed "arg max" displacement field (right column) for two consecutive myelin stained histological sections. The color bar in the middle column shows 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.

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 map.

Obviously, the similarity squares present clear patterns, 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.

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 [3]. 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: $L_F = \{s \, s.t. \, \exists!(i,j) \in L_F, s = (i,j)\}_{t=1}^{w_F.h_F}$ (Table 2 shows the main notations we use for our approach). Our clustering method is adapted from the standard agglomerative hierarchical clustering algorithm described in [19]:

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 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.

The number of clusters can either be specified by the user (our case here), or pre-indicators like the Davies-Bouldin index [12] or the cophenetic correlation coefficient [3] can assist this choice. The Appendix Section briefly 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:

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

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 (the continuous red line in Figure 4 is the geodesic path between A and B, the dotted line the Euclidean path). 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 [10] which achieves a good trade-off of precision over speed.

Given two sites t and u, their centroid distance is written:

$$D_{centroid}\,(t, u) = D_{geodesic}\left(p_F^t, p_F^u\right) \tag{5}$$

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 [32] 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 $\rho^t\left(p_R^u - p_F^t\right) = \frac{sim\left(b_F^t, b_R^u\right)}{\sum_{v \in N_t} sim\left(b_F^t, b_R^v\right)}$. Such distributions are depicted in Figure 4 (middle column).

As a distance between distributions, we chose the Earth mover's distance [31], a discrete solution to the discrete Monge-Kantorovich mass-transfer problem [16]. 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.* [31], 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:

$$D_{distribution}\left(t,u\right) = D_{EMD}\left(\rho^t, \rho^u\right) \tag{6}$$

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

$$D_{site}\left(t,u\right) = \alpha\, D_{geodesic}\left(p_F^t, p_F^u\right) + \left(1-\alpha\right) D_{EMD}\left(\rho^t, \rho^u\right) \tag{7}$$

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 [3] 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 Sum of Squares) estimator, see Section 2.3). 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 [3] 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:

$$D_{trsf}\left(T^a, T^b\right) = \begin{cases} \sum_{i,j}\left[T^a T^{b^{-1}} - Id\right]_{i,j}^2 + \sum_{i,j}\left[T^{a^{-1}} T^b - Id\right]_{i,j}^2 & \text{if both } T^{xx} \text{ are defined} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

(where i, j are matrix indices).

Finally, given two clusters of sites $C^a = \{a_1, \ldots, a_{n_a}\}$ and $C^b = \{b_1, \ldots, b_{n_b}\}$, the cluster distance between them is the longest distance from any site in C^a to any site in C^b

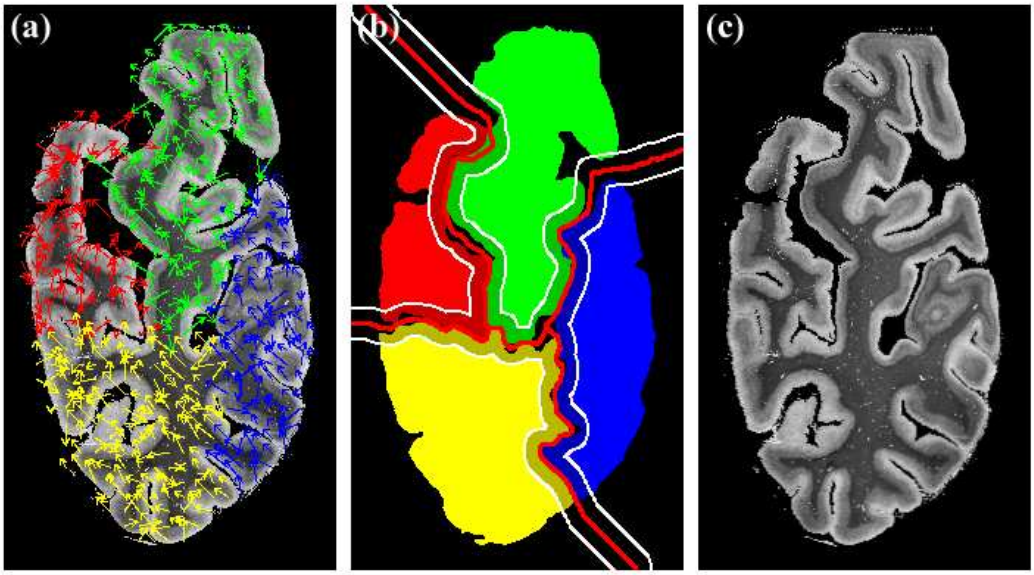

Figure 5: Composing the locally registered floating image: (a) input floating image with clustered "arg max" displacement field; (b) transformed binarized sub-images with darkened eroded pixels; (c) final composed locally registered floating image.

(complete-link) *plus* the "transformation distance" wherever it can be computed:

$$D_{cluster}(C^a, C^b) = \beta \max_{i,j} D_{site}(a_i, b_j) + (1 - \beta)\ D_{trsf}(T^a, T^b) \tag{9}$$

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

Figure 5.a shows the clustering of the correspondence field for the two consecutive histological sections of Figure 4 (with 4 clusters).

As an alternative, we could have used one of the numerous optical flow segmentation algorithms developed in the literature ([36] 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, $C = \{C^1, \ldots, C^{N_C}\}$ the cluster partition of L_F, and $\{c_1^i, \ldots, c_{n_i}^i\}$ the n_i sites of the ith 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_i^j (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:

$$I_F^i = \{(x, y) \in I_F \text{ such that } D((x, y), c_j^i) \leq cover_{radius} \text{ for some } c_j^i \in C^i\} \tag{10}$$

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

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

$$I_R^i = \{(x, y) \in I_R \text{ such that } D((x, y), c_j^i + d^{c_j^i}) \leq cover_{radius}, \text{ for some } c_j^i \in C^i\} \tag{11}$$

Again, we use the L_∞ distance here, with $cover_{radius} = b_size$ (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 Composing the registered floating image

Once we have extracted the reference and floating sub-images, we use the robust affine block-matching algorithm described in [26] 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 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:

$$(R^*, t^*) = \arg\min_{R,t} \sum_{i,j} \left\| \left(p_F^{i,j} + d^{i,j}\right) - R.p_F^{i,j} - t \right\|^2 \tag{12}$$

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

However, given the rather noisy appearance of the displacement field, an LTS estimator (Least Trimmed Sum of Squares, see [30] 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 12), 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. A robustness study on rat brains sections presented in [26] demonstrated its ability to recover rotations up to 28 degrees.

Then, for each pair of sub-images $\left\{I_R^l, I_F^l\right\}$, $l \in 1 \ldots N_C$, 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.

We then compose the final registered floating image using the Little *et al.* method [23]. Their approach applies user-provided affine transforms to user-defined structures and ensures a smooth interpolation in between them. 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. This erosion algorithm works as follows:

- We first apply the transformations to the floating sub-images ($\forall l \in 1 \ldots N_C$, $T^l\left(I_F^l\right)$ 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 $T^l\left(I_F^l\right)$ are disjoint.
  - A Euclidean distance map of the background of J is computed.
- A medial axis algorithm then extract 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: $\forall l \in 1 \ldots N_C$, $\overline{I_F^l} = I_F^l - T^{l\,-1}\left(T^l\left(I_F^l\right) \cap N\right)$.

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). We are currently investigating technique to link ν to the distance between the transformations associated to neighboring sub-images.

We choose as landmarks the corners of the original images, I_R and I_F (after the initial rigid registration), to further constraint the interpolation scheme, and use the modified Hardy multi-quadric recommended in [23] as a basis function for interpolation, as this agrees with an affine transform at infinity. Figure 5 exemplifies our composition process on the myelin stained histological section of Figures 1 and 4. We show in 5.b the transformed floating binarized sub-images in color, the skeleton of the background of J in red and the $2\nu = 20$ pixel wide band of eroded pixels in darkened colors with white borders.

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. We are currently exploring these aspects.

3 Results

We present here the various experiments we have conducted to assess the performances of our local registration approach. We first discuss the ability of the clustering algorithm to correctly classify phantom images. We then detail how the various components of our registration system influence the quality of the match for series of biomedical images.

3.1 Phantom Images

Phantom images enable us to evaluate the quality of the clustering algorithm under controlled textural conditions. Namely, pairs of synthetic images were created with artificial textures (grayscale cloud pattern, see Figure 6).

A first experiment was designed to illustrate the behavior of our approach on images consisting of several connected components. We considered 3 structures with 2 components each: in #1 a small square inside a larger one, in #2 and #3 the two vertical halves of a large square. To each component, a different affine transformation was applied:

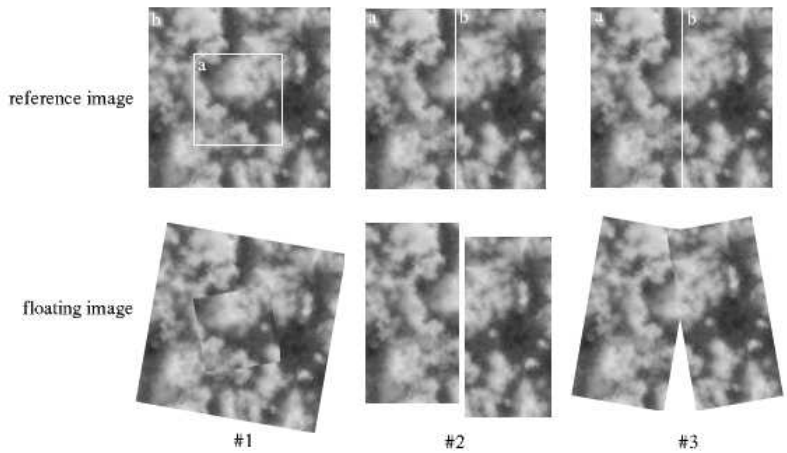

Figure 6: Clustering of three 2-component structures

- component 1a: T^{1a} = counter-clockwise $\frac{\pi}{8}$ rotation;
- component 1b: T^{1b} = clockwise $\frac{\pi}{8}$ rotation;
- component 2a: $T^{2a} = (-10, -20)$ translation;
- component 2b: $T^{2b} = (-5, +15)$ translation;
- component 3a: T^{3a} = counter-clockwise $\frac{\pi}{8}$ rotation;
- component 3b: T^{3b} = clockwise $\frac{\pi}{8}$ rotation.

Table 1 reports the classification results with and without the transformation based distance, and with and without distribution distance (EMD). We set $\alpha = 0.5$, $\beta = 0.5$, $\theta_{cluster} = 20$ and $N_C = 2$, and used the extended correlation coefficient for rigid registration. For each structure, we show, for each component, both the number of pixels of that component that were correctly classified and the number of vectors of the complementary component that were wrongly classified as this one, as a percentage of their respective number of pixels.

We observed better performances when the EMD distribution distance was used than when it was not. The similarity distribution distance actually helps the clustering algorithm to form, at early stages, sensible clusters that are then adequately agglomerated with the aid of the robust estimation of the associated transformations (even though, due to the noisy nature of the "arg max" displacement field, those may sometimes hinder the clustering process). Finally, use of the geodesic distance was particularly beneficial for components 2a/b and 3a/b.

	with transformation distance		without transformation distance	
components	with EMD	without EMD	with EMD	without EMD
# 1a	43 / 47	40.1 / 34	40 / 43.2	37.6 / 41
# 1b	53 / 57	66 / 59.9	56.8 / 60	59 / 72.4
# 2a	86.6 / 26.8	68.3 / 46.2	74.3 / 32.4	72 / 34.2
# 2b	73.2 / 13.4	53.8 / 31.7	67.6 / 25.7	65.8 / 28
# 3a	76.5 / 19	59.2 / 26.5	54.3 / 27.5	57.2 / 11.5
# 3b	81 / 23.5	73.5 / 40.8	72.5 / 45.7	88.5 / 42.8

Table 1: Performance of our local registration algorithm for a variety of phantom images.

Note that a *perfect* clustering is not a necessary condition for our method to perform well, (1) since the subsequently extracted sub-images will be larger than the obtained clusters and will overlap, and (2) since the registration algorithm we use to register these sub-images is robust.

3.2 Biomedical Images

3.2.1 Two detailed examples of mono- and multi-modal registration

Figures 7 displays the results of our local registration for the pair of myelin-stained histological sections introduced in Figure 1, and Figure 8 for an autoradiographic monkey brain section and its associated MRI (obtained by co-registration of a series of contiguous autoradiographic sections with an MR volume of the same monkey [25, 4]). In both cases, a gyrus (top left corner in Figure 7, and bottom left corner in Figure 8) was detached during the histological preparation and manually realigned in an unsatisfactory fashion.

In Figure 7, we show the reference image (a), the transformed binarized floating sub-images with darkened eroded pixels (b, refer to Section 2.3 for details), the locally registered floating image (c), the colored superposition of the reference image and the globally affinely registered floating image (d), the image of a regular grid convected with the associated hybrid affine/non-linear transformation with the superimposed transformed eroded floating sub-images (e), and the colored superposition of the reference and the locally registered floating image (f). In Figure 8, the transformed sub-images composition is replaced by the floating image with superimposed clustered "argx max" displacement field (b), the colored superpositions are replaced by the reference image on which we superimposed the edges of the globally registered floating image (d) and of the locally registered floating image (f).

These pairs of images were locally *rigidly* registered by our approach, with $\alpha = 0.5$, $\beta = 0.5$, $\theta_{cluster} = 20$ and standard parameters for the block matching algorithm (see Table 3 in Appendix) with an extended correlation coefficient and $N_C = 4$ for the myelin sections, and a classic coefficient with $N_C = 2$ for the autoradiography (tests with normalized mutual information as a similarity measure yielded less good results). Our clustering algorithm

adequately isolated in a separate sub-image the floating gyrus of the myelin section pair (red area in Figure 7.b) and the moving part in the monkey case (green area in Figure 8.b) which were subsequently correctly registered to their counterpart in the reference image. An affine transform would of course further decrease the discrepancy between the pairs of sub-images. 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.

3.2.2 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 two pairs of images presented above (Figures 7 and 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 or autoradiographic images. Previous work [26, 25, 4] 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 describe here the reconstruction of a 3-D histological volume from a series of 70 images. These were $50\mu m$ 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. Note that this process require 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)
```

We used here the same parameters as for the registration of the myelin-stained sections of Figure 7): $\alpha = 0.5$, $\beta = 0.5$, $\theta_{cluster} = 20$ and standard parameters for the block matching algorithm with an extended correlation coefficient and $N_C = 6$.

Figure 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 2.3). In both cases (rigid and piece-wise rigid) we show in (b) a 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, in (a) the 50th image (the immediately preceding section) with edges of the 51st one superimposed in red, and in (c) the 52nd image (the immediately following section) with edges of the 51st image superimposed.

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 all through the 3-D volume confirmed the enhanced continuity of the 3-D structures.

4 Conclusion and Perspectives

We have presented a fully automated local registration method, capable of dealing with a variety of 2-D images. It builds complex spatial transformations by elastically interpolating between rigid or affine transforms that are locally defined on pairs of sub-images. Consequently, we manage to minimize the overall number of degrees of freedom of the transformation, thereby agreeing with the guidelines of the parsimony principle (see [21] for a discussion of the problems inherent to high-dimensional transformations). These sub-images represent geometrically coherent components (in our biomedical applications, they are even anatomically coherent components). They are automatically extracted from an initial displacement field computed between the images to be registered. All user interaction is avoided, by contrast with other approaches [23].

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. Even though obtaining a *perfect* clustering may not actually be necessary, we are currently investigating stochastic approaches to improve this step.

Results on real data in both 2-D registration and 3-D reconstruction cases proved that the proposed method is adequate for several specific problems in biomedical imaging. Finally, even though the presented registration method works in 2-D, it could readily be extended to 3-D with a close-to-linear increase in processing time.

Acknowledgement

This work was partially funded by European project MAPAWAMO (ref. QLG3-CT-2000-30161; coordinator: Prof. Guy Orban, Lab. of Neuro- and Psychophysiology, Dept. of Neurosciences and Psychiatry, K.U. Leuven, Belgium). We thank W. Vanduffel, K. Nelissen, D. Fize and G. Orban for providing us with the monkey autoradiography and MRI. We also thank J. Annese from the LONI laboratory for providing us with the human brain sections. 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 between EPIDAURE and LONI: http://www-direction.inria.fr/international/EQUIPES_ASSOCIEES/index.html.

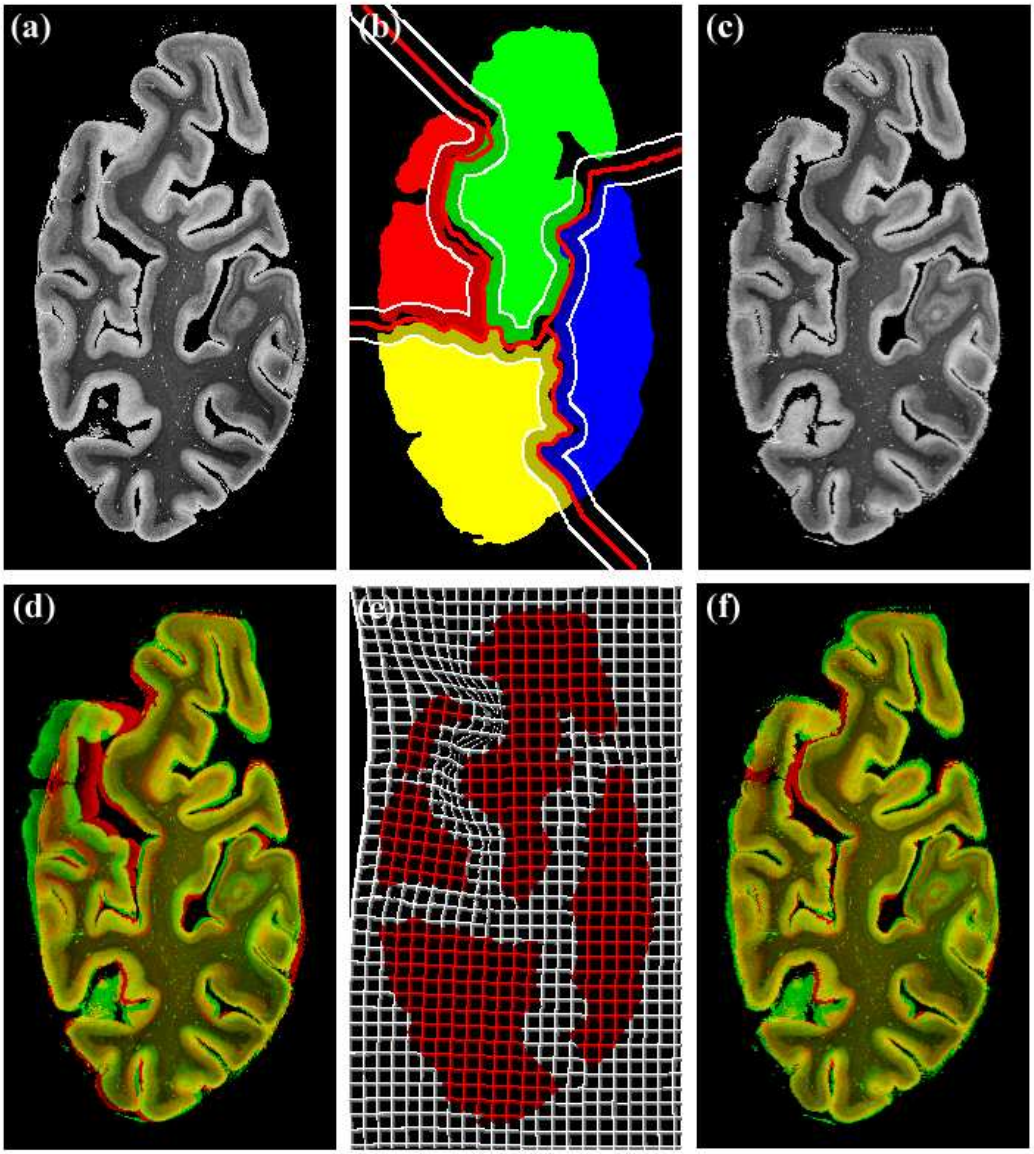

Figure 7: Registration of two consecutive myelin-stained histological sections of the human brain: (a) reference image, (b) transformed binarized floating sub-images with darkened eroded pixels, (c) locally registered floating image, (d) superposition of the reference image (red) and of the globally affinely registered floating image (green), (e) image of a regular grid convected by the associated hybrid affine/non-linear transformation with superimposed transformed eroded floating sub-images (in red), (f) superposition of the reference image (red) and of the locally registered floating image (green).

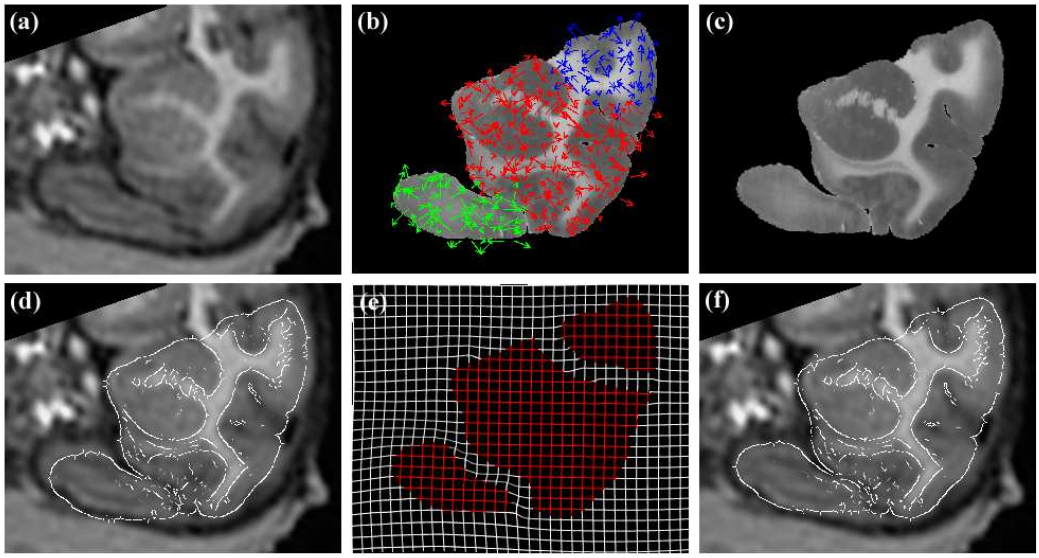

Figure 8: Registration of an autoradiographic monkey brain section and its associated MRI: (a) reference image, (b) floating image with clustered optimal "arg max" displacement field, (c) locally registered floating image, (d) reference image with superimposed edges of the globally affinely registered floating image, (e) image of a regular grid convected with the associated hybrid affine/non-linear transformation with superimposed transformed eroded floating sub-images (in red), (f) reference image with the superimposed edges of the locally registered floating image.

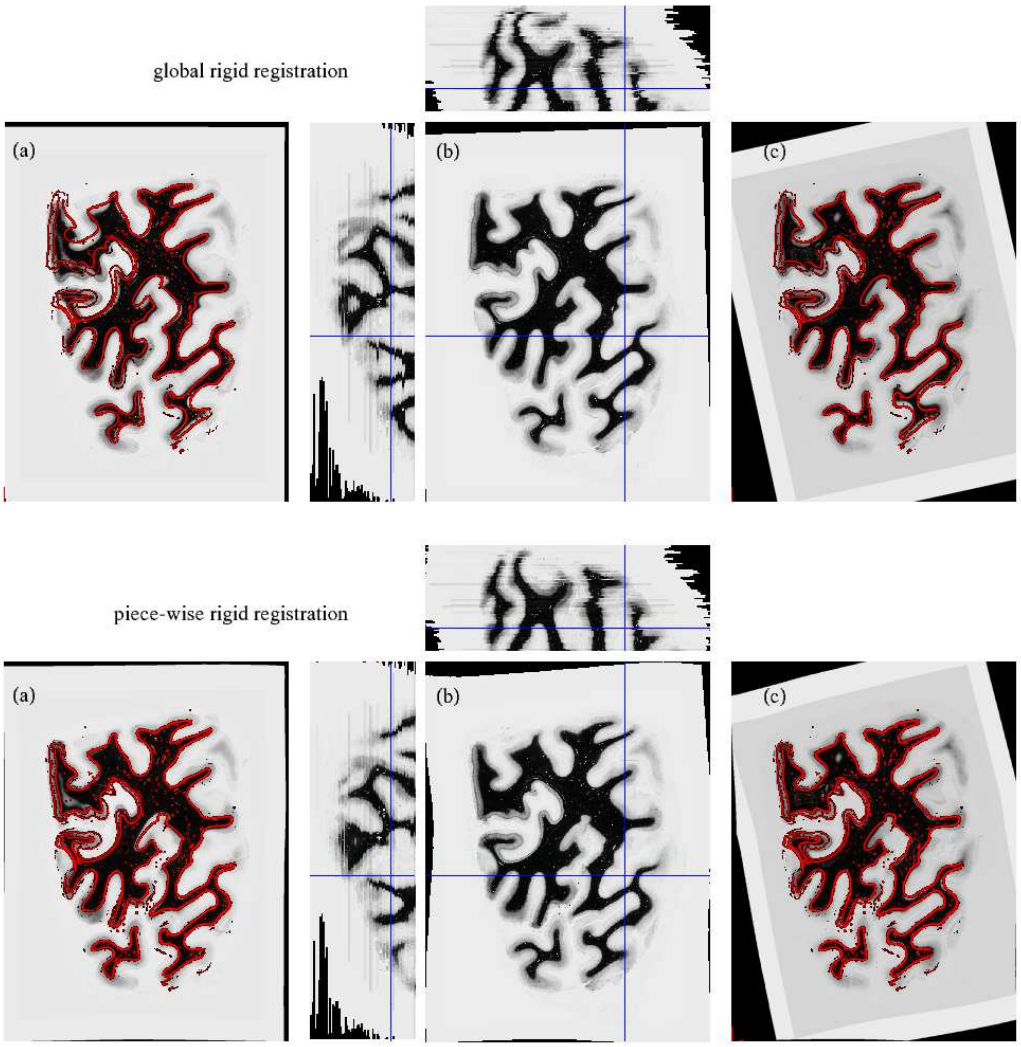

Figure 9: Reconstruction of a 3-D histological volume with a globally rigid (top) and our piece-wise rigid (bottom) registration algorithm: (b) coronal view (middle) of the 3-D reconstructed volume corresponding to the 51st image of the stack with the associated axial (top) and sagittal (left) views, (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

Appendix

Notations

Table 2 summarizes the notations that we used throughout this article.

Notation	Description	Comment
I_R	reference image (size $w_R \times h_R$)	fixed image
I_F	floating image (size $w_F \times h_F$)	moveable image
L_R, L_F	reference and floating lattice	list of sites
(i,j), t, u	site, site in L_F, site in L_R	
$N_R^{i,j}$, $N_F^{i,j}$	neighborhood of a site	
$b_R^{i,j}$, b_R^t	block of I_R associated with site (i,j) or t	group of pixels (usually a square)
p_F^t	centroid of the block of I_F associated with site t of L_F	2-D point
cor	classical correlation coefficient	computed on b_R^t and b_R^u
ecor	extended correlation coefficient	taking into account N_R^t and N_F^u
d^t	"arg max" displacement vector associated with site t	2-D vector
ρ^t	similarity distribution associated with site t	
C^a	cluster	unordered set of sites
$\{c_1^a, \ldots, c_{n_a}^a\}$	the n_a sites of cluster C^a	
T^a	estimated transform of C^a	rigid or affine
I_R^l, I_F^m	reference and floating sub-images	

Table 2: Notations

Standard values

Table 3 reports standard values for the various parameters of the algorithms we use in our approach.

Number of clusters

To evaluate the influence of the specified number of clusters on the final registration quality, we used the synthetic images of Figure 6. We checked 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 part of the transformation of the enclosing component (with minimal error, 2% on average).

Parameter	Algorithm	Name	Typical value	Comments
block size	block matching	b_{size}	6×6	values between 4 and 7 give similar results.
size of $N_R^{i,j}$	block matching	*<none>*	20×20	depends on maximal. distance between corresponding components.
lattices step size	block matching	*<none>*	5×5	little/no impact on the registration quality.
similarity measure	block matching	*sim*	*cor* or *ecor*	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	little/no impact on the registration quality.
space between eroded structures	sub-images erosion	ν	10	values between 5 and 20 give similar results.

Table 3: Standard values for algorithm parameters

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.

References

[1] V. Arsigny, X. Pennec, and N. Ayache. Polyrigid and Polyaffine Transformations: A New Class of Diffeomorphisms for Locally Rigid or Affine Registration. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI'03)*, 2003.

[2] J. Ashburner and K. J. Friston. Nonlinear spatial normalization using basis functions. *Human Brain Mapping*, 7(4):254–266, 1999.

[3] E. Backer. *Computer-assisted reasoning in cluster analysis.* Prentice Hall, 1995.

[4] E. Bardinet and G. Malandain. Fusion of autoradiographies with an mr volume using 2-d and 3-d linear transformations. Research report 4791, INRIA, 2003.

[5] E. Bardinet, S. Ourselin, D. Dormont, G. Malandain, D. Tandé, K. Parain, N. Ayache, and J. Yelnik. Co-registration of histological, optical and mr data of the human brain. In T. Dohi and R. Kikinis, editors, *Medical Image Computing and Computer-Assisted Intervention (MICCAI'02)*, volume 2488 of *LNCS*, pages 548–555, Tokyo, September 2002. Springer.

[6] G. E. Christensen. Consistent linear-elastic transformations for image matching. In A. Kuba, M. Samal, and A. E. Todd-Pokropek, editors, *Proc. Information Processing in Medical Imaging (IPMI'99)*, volume 1613 of *LNCS*, pages 224–237, Visegrád, Hungary, 1999. Springer.

[7] A. Collignon, F. Maes, D. Delaere, D. Vanderneulen, P. Suetens, and G. Marchal. Automated multimodality image registration using information theory. In Y. Bizais, C. Barillot, and R. Di Paola, editors, *Proc. Information Processing in Medical Imaging (IPMI'95)*, Computational imaging and vision, pages 263–274. Kluwer, 1995.

[8] D. L. Collins, A. C. Evans, C. Holmes, and T. M. Peters. Automatic 3D segmentation of neuro-anatomical structures from MRI. In Y. Bizais, C. Barillot, and R. Di Paola, editors, *Proc. Information Processing in Medical Imaging (IPMI'95)*, Computational imaging and vision, pages 139–152. Kluwer, 1995.

[9] D. L. Collins, A. P. Zijdenbos, T. Paus, and A. C. Evans. Use of registration for cohort studies. In J. Hajnal, D. Hawkes, and D. Hill, editors, *Medical Image Registration.* 2001.

[10] O. Cuisenaire. *Distance Transformations: Fast Algorithms and Applications to Medical Image Processing.* PhD Thesis, 1999.

[11] C. Davatzikos. Spatial transformation and registration of brain images using elastically deformable models. *Computer Vision and Image Understanging*, 66(2):207–222, 1997.

[12] D. L. Davies and D. W. Bouldin. A cluster separation measure. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1:224–227, 1979.

[13] J. Dengler. Estimation of discontinuous displacement vector fields with the minimum description length criterion. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'91)*, pages 276–282, Lahaina, Maui, Hawaii, 1991. IEEE Computer Society Press.

[14] J. Feldmar and N. Ayache. Rigid, affine and locally affine registration of free-form surfaces. *The International Journal of Computer Vision*, 18(2), May 1996.

[15] J. C. Gee, M. Reivich, and R. Bajcsy. Elastically deforming 3D atlas to match anatomical brain images. *Journal of Computer Assisted Tomography*, 17(2):225–236, March/April 1993.

[16] S. Haker, S. Angenent, and A. Tannenbaum. Minimizing flows for the Monge-Kantorovich problem. *SIAM Journal of Math Analysis*, 2003. [to appear].

[17] S. M. Haney, P. M. Thompson, T. F. Cloughesy, J. R. Alger, and A. W. Toga. Tracking tumor growth rates in patients with malignant gliomas: A test of two algorithms. *American Journal of Neuroradiology*, 22(1), Jan. 2001.

[18] A. K. Jain. Image data compression: a review. *Proceedings of the IEEE*, 69(3):349–389, Mar. 1981.

[19] S. C. Johnson. Hierarchical clustering schemes. *Psychometrika*, 2:241–254, 1967.

[20] F. A. Jolesz, A. Nabavi, and R. Kikinis. Integration of interventional MRI with computer-assisted surgery. *Journal of Magnetic Resonance Imaging*, 13(1):69–77, Jan. 2001.

[21] U. Kjems, L. K. Hansen, and C. T. Chen. A Non-linear 3D brain co-registration method. In P. C. Hansen, editor, *Proceedings of the Interdisciplinary Inversion Workshop 4*, Lyngby, Denmark, 1996. IMM, Technical University of Denmark.

[22] B. Likar and F. Pernus. Registration of serial transverse sections of muscle fibers. *Cytometry*, 37:93–106, 1999.

[23] J. A. Little, D. L. G. Hill, and D. J. Hawkes. Deformations incorporating rigid structures. *Computer Vision and Image Understanding*, 66(2):223–232, 1997.

[24] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, Mar. 1998.

[25] G. Malandain and E. Bardinet. Fusion of autoradiographies with an MR volume using 2-D and 3-D linear transformation. In *Proc. Information Processing in Medical Imaging (IPMI'03)*, 2003.

[26] S. Ourselin, A. Roche, G. Subsol, X. Pennec, and N. Ayache. Reconstructing a 3D Structure from Serial Histological Sections. *Image and Vision Computing*, 19(1-2):25–31, Jan. 2001.

[27] D. Rey, G. Subsol, H. Delingette, and N. Ayache. Automatic Detection and Segmentation of Evolving Processes in 3D Medical Images: Application to Multiple Sclerosis. *Medical Image Analysis*, 6(2):163–179, June 2002.

[28] A. Roche. *Recalage d'images médicales par inférence statistique.* Thèse de sciences, Université de Nice Sophia-Antipolis, Feb. 2001.

[29] A. Roche, G. Malandain, and N. Ayache. Unifying Maximum Likelihood Approaches in Medical Image Registration. *International Journal of Imaging Systems and Technology: Special Issue on 3D Imaging*, 11(1):71–80, 2000.

[30] P. Rousseeuw. Least median of squares regression. *Journal of the American Statistical Association*, 79:871–880, 1984.

[31] Y. Rubner, C. Tomasi, and L. Guibas. A metric for distributions with applications to image databases. In *Proc. of International Conference on Computer Vision (ICCV'98)*, Bombay, India, 1998. IEEE Computer Society, Narosa.

[32] A. Singh and P. Allen. Image-flow computation: an estimation-theoretic framework and a unified perspective. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 56(2):152–177, 1992.

[33] C. Studholme, D. L. G. Hill, and D. J. Hawkes. Incorporating connected region labelling into automated image registration using mutual information. In *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA'96)*, pages 23–31. IEEE Computer Society Press, 1996.

[34] G. Taubin. An improved algorithm for algebraic curve and surface fitting. In *Proc. of International Conference on Computer Vision (ICCV'93)*, pages 658–665, Berlin, Germany, May 1993. IEEE Computer Society Press.

[35] P. M. Thompson, R. P. Woods, M. S. Mega, and A. W. Toga. Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain. *Human Brain Mapping*, 9(2):81–92, 2000.

[36] J. Weber and J. Malik. Rigid Body Segmentation and Shape Description from Dense Optical Flow Under Weak Perspective. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:139–143, Feb 1997.

[37] W. M. Wells, P. Viola, H. Atsumi, and S. Nakajima. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1):35–51, 1996.

[38] R. P. Woods, S. T. Grafton, C. J. Holmes, S. R. Cherry, and J. C. Mazziotta. Automated image registration: I. general methods and intrasubject, intramodality validation. *Journal of Computer Assisted Tomography*, 22:141–154, 1998.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399