3D volume reconstruction of a mouse brain from histological sections using warp filtering

Tao Ju [e,*], Joe Warren [a], James Carson [f], Musodiq Bello [d], Ioannis Kakadiaris [d],
Wah Chiu [b], Christina Thaller [b], Gregor Eichele [c]

[a] *Rice University, Houston, TX, USA*
[b] *Baylor College of Medicine, Houston, TX, USA*
[c] *Max Planck Institute of Experimental Endocrinology, Hannover, Germany*
[d] *University of Houston, Houston, TX, USA*
[e] *Washington University, St. Louis, MO, USA*
[f] *Pacific Northwest National Laboratory, Richland, WA, USA*

Abstract

Sectioning tissues for optical microscopy often introduces upon the resulting sections distortions that make 3D reconstruction difficult. Here we present an automatic method for producing a smooth 3D volume from distorted 2D sections in the absence of any undistorted references. The method is based on pairwise elastic image warps between successive tissue sections, which can be computed by 2D image registration. Using a Gaussian filter, an average warp is computed for each section from the pairwise warps in a group of its neighboring sections. The average warps deform each section to match its neighboring sections, thus creating a smooth volume where corresponding features on successive sections lie close to each other. The proposed method can be used with any existing 2D image registration method for 3D reconstruction. In particular, we present a novel image warping algorithm based on dynamic programming that extends Dynamic Time Warping in 1D speech recognition to compute pairwise warps between high-resolution 2D images. The warping algorithm efficiently computes a restricted class of 2D local deformations that are characteristic between successive tissue sections. Finally, a validation framework is proposed and applied to evaluate the quality of reconstruction using both real sections and a synthetic volume.
© 2006 Elsevier B.V. All rights reserved.

Keywords: 3D reconstruction; Filtering; Image warping; Dynamic programming; Histology

1. Introduction

In medical imaging, volumetric data generated by 3D imaging methods such as MRI and CT have wide applications in the visualization and analysis of organs. 2D imaging methods, such as optical microscopy, typically generate serial sections with much higher resolution than MRI or CT scans. Reconstruction of these 2D sections in 3D has therefore become an important tool for understanding anatomical structures in 3D, and in particular, for building high-resolution templates (atlases) of organs (Timsari et al., 1999; Armstrong et al., 1995; Rosen et al., 2000; Cannestra et al., 1997; MacKenzie-Graham et al.,

2004; Sidman, 2005), or of whole animals (Brune et al., 1999). Our work is motivated by the need for automatic and efficient methods in reconstructing, from a stack of high-resolution Nissl-stained sections, a smooth representative volume for building a high-quality 3D atlas of a mouse brain. Applications of the resulting atlas include visualization, physical computing, as well as constituting a standard coordinates for mapping and comparing brain data from different individuals. The authors have previously constructed a 2D atlas of the mouse brain that has been successfully used to build an atlas-based database of gene expression patterns (Ju et al., 2003).

Unfortunately, direct 3D reconstruction of serial sections by stacking successive sections will not produce a smooth volume. While 3D imaging methods can be applied in vivo, planar imaging methods are applied ex vivo. The preparation steps required by 2D imaging methods may introduce undesirable tissue distortions specific to the preparation procedures, which

* Correspondence to: One Brookings Drive, Campus Box 1045, St. Louis, MO 63130, USA. Tel.: +1 314 935 6648.
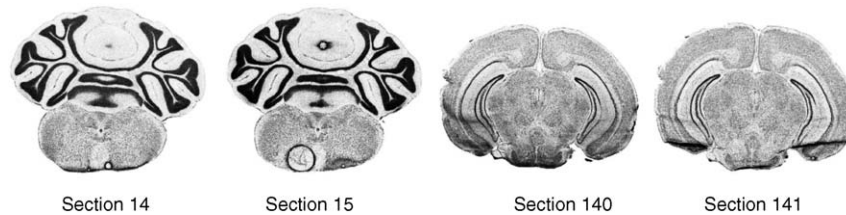*E-mail address:* taoju@cs.wustl.edu (T. Ju).

Fig. 1. Four coronal (i.e., vertically cut from front to back) sections from a stack of histological sections of a mouse brain acquired by cryo-sectioning. Besides rotational and translational differences, successive sections exhibit different nonlinear tissue distortions, such as compacting, stretching, and even folds and tears, as well as artifacts, such as air bubbles and dust.

have been closely examined by several authors (Berthold et al., 1982; Deverell et al., 1989; Schormann et al., 1995; Brey et al., 2002). The experimental data used in this paper were acquired by cryo-sectioning a frozen mouse brain into 25 µm thin slices followed by histology staining (Carson et al., 2002). Distortions on the resulted sections are mainly introduced in the sectioning stage, which include global translation and rotation, local stretching and compacting, and various image artifacts (see a more detailed discussion in Section 2.2.1). Since these deformations vary from section to section, as shown in the examples in Fig. 1, distortion correction is required to reconstruct a smooth volumetric representation.

1.1. Related works

There has been tremendous amount of work on 3D reconstruction from distorted serial sections. Typically, reconstruction is achieved by deforming individual 2D sections using image registration (warping) techniques. Here we briefly review related works on 2D image warping and 3D reconstruction.

1.1.1. Image warping

Image warping considers the task of finding the best deformation (warp) of a source image that matches a target image under a specific distance measure such as pixel intensities or mutual information (Maes et al., 1997). Image warping methods have been intensively studied in the field of medical imaging, and we refer interested readers to survey articles by Antoine Maintz and Viergever (1998), by Glasbey (1998) and by Lester and Arridge (1999), as well as books by Toga (1999), by Hajnal et al. (2001), and by Modersitzki (2004) for excellent reviews. Software packages implementing state-of-the-art warping methods are also available for medical applications, such as the Automated Image Registration (AIR) package from UCLA (Woods et al., 1998a) and the Insight Segmentation and Registration Toolkit (ITK) from NLM (Yoo et al., 2002).

For the purpose of 3D reconstruction, we are interested in the classification of image warping methods by the nature of the resulting image deformations. A number of methods compute deformations represented as global polynomials, such as rigid-body transformations (Woods et al., 1993; Maes et al., 1997), linear affine transformations (Rangarajan et al., 1997; Thevenaz et al., 1998), and higher-degree polynomial deformations (Woods et al., 1998b). Since these deformations are global in nature, they do not perform well in the case of local variations between images (Glasbey, 1998). To overcome this problem,

transformation frameworks that allow more localized deformations have been proposed, such as free-form B-spline based deformations defined on a regular grid of control points, which have been studied by numerous authors including Rueckert et al. (1999) and Studholme et al. (2000), and piecewise affine transformations considered by Pitiot et al. (2003). Local warping methods have more flexible mapping functions and are often regularized by some form of elastic energy, such as the thin-plate spline energy (Bookstein, 1989), to prevent excessive deformations. Common optimisation techniques for computing a regularized solution include solving partial differential equations (an excellent review of such methods can be found in the book by Modersitzki, 2004), finite-element methods (Guest et al., 2001), graph methods such as max-flow (Roy, 1999) and min-cut (Boykov et al., 1999), stochastic methods (Agazzi et al., 1993), and dynamic programming.

Since its first successful application in speech recognition by Sakoe and Chiba (1978), dynamic programming has been known in 1D as Dynamic Time Warping (DTW) and has been extended in various ways for warping 2D images (Levin and Pieraccini, 1992; Cox et al., 1996; Uchida and Sakoe, 1998; Ronee et al., 2001). Compared to other optimization methods in the continuous domain, dynamic programming finds discrete minimizers in a robust manner. However, due to the exponential time complexity of a fully 2D dynamic programming task, existing DTW-based image warping methods can only handle extremely low-resolution images.

1.1.2. 3D reconstruction

Previous methods on warp-based 3D reconstruction from serial sections can be appreciated from two perspectives: the types of image deformations performed, and the subjects from which the deformations are computed.

Over the past two decades, many authors have reported how linear transformations can be applied to automatically bring successive sections into alignment (Toga and Arnicar-Sulze, 1987; Lamers et al., 1989; Rydmark et al., 1992; Schormann et al., 1995; Hess et al., 1998; Streicher et al., 2000; Ourselin et al., 2001a,b; Brey et al., 2002; Nikou et al., 2003). The simplicity of linear transformations not only reduces the complexity of computation but also allows convenient manual adjustment through software interfaces (Rosen et al., 2000; Chen et al., 2003; Karen et al., 2003; Andrey and Maurin, 2005). However, as shown in several studies (Deverell et al., 1989; Jones et al., 1994; Boyle et al., 1997), the section distortions induced by the preparation process are local in nature. With the advance

of image registration techniques, reconstruction methods were proposed to correct localized section distortions by using local 2D deformations (Durr et al., 1989; Guest and Baldock, 1995; Kim et al., 1997; Timsari et al., 1999; Ali and Cohen, 1998) as well as elastic 3D surface deformations (Thompson and Toga, 1996; Mega et al., 1997; Gabrani, 1998; Gefen et al., 2003).

A large number of the above reconstruction methods compute the warp of each section so that the warped section matches a neighboring section. For example, Durr et al. (1989) computes elastic deformation between pairs of distorted images, while Karen et al. (2003) uses software tools for user-assisted alignment between every two consecutive sections by rigid-body transformations. However, unlike image registration where exact matching is preferred, the goal of reconstruction is to form a smooth volume allowing natural progression of features through successive images (Guest and Baldock, 1995). Recent methods, and in particular elastic 3D surface deformation methods, focus on warping distorted sections onto an undistorted reference, such as block-face photos (Mega et al., 1997; Kim et al., 1997; Gefen et al., 2003) or tissue markers (Streicher et al., 2000; Ourselin et al., 2001b), 2D sections of a 3D in vivo image (Thompson and Toga, 1996; Schormann et al., 1995; Ourselin et al., 2001a), or sections from an existing template (He et al., 1995; Ali and Cohen, 1998; Timsari et al., 1999). The problem with this approach is that in many applications an un-sectioned reference is not always available.

There have only been a few works so far (Montgomery and Ross, 1994; Guest and Baldock, 1995; Wirtz et al., 2004) that address the problem of smooth 3D reconstruction using elastic 2D image deformations in the absence of a reference volume (note that existing methods using elastic 3D surface deformations (Thompson and Toga, 1996; Mega et al., 1997; Gefen et al., 2003) cannot be applied due to the lack of a reference). Without the knowledge of the original object before sectioning, these works based their reconstruction on the following assumption: the shape of an anatomical structure varies slowly with respect to section thickness. In other words, corresponding points on adjacent sections are likely to be located close together. Both Guest and Baldock (1995) and Wirtz et al. (2004) consider all sections at the same time and aim at minimizing an overall energy functional consisting of distances between every two consecutive sections (e.g., spring forces between corresponding points (Guest and Baldock, 1995) or pixel-wise squared differences (Wirtz et al., 2004)) and the sum of elastic deformation potentials on every section. Although numerical solutions can be obtained by using finite element method (Guest and Baldock, 1995) or by approximating the differential equations using finite differences (Wirtz et al., 2004), solving the global minimization problem involving all sections is non-trivial due to its massive size.

A completely different approach was taken by Montgomery and Ross (1994). The idea is to reposition a point on each section by applying a local Laplacian smoothing operator on the position of that point and the positions of its corresponding points on the two adjacent sections. However, their method requires manual delineation of contours on each section to establish correspondence. Moreover, the Laplacian operator is only applied to points on the contours, hence the deformation is restricted to

the overall shape of a contoured region and has limited power in matching interior features with neighboring sections.

1.2. Contributions

In this paper, we introduce a new approach for reconstructing a smooth representative volume by elastically deforming serial sections, which requires no human intervention or the use of a reference volume. Since it is impossible to "undo" the tissue distortions without the knowledge of the original object before sectioning, we base our reconstruction on the same assumption from previous works (Guest and Baldock, 1995), that is, the section thickness (25 μm) is small comparing to variations in the shape of anatomical structures. Instead of considering all sections at the same time (Guest and Baldock, 1995; Wirtz et al., 2004) or only two neighbors for each section (Montgomery and Ross, 1994), we consider an extended neighborhood to compute the deformation of each section. Our goal is to produce a smooth volume where a point on a section lies close to its corresponding points within its neighborhood. To achieve this goal, we present:

(1) An automatic 3D reconstruction method, called *warp filtering*, based on 2D image warps between each pair of adjacent sections. During the reconstruction, each section is deformed using an average warp computed from the pairwise warps within a group of neighboring sections. The average warp results in a smooth volume by effectively repositioning each point on one section to the *weighted-average* location of the corresponding points on the neighboring sections. The algorithm works with any 2D image registration techniques for computing pairwise warps, and can be easily parallelized for speed. We performed quantitative and qualitative validation of the reconstruction method on both real and synthetic data, and the results revealed the effectiveness of our method in building a smooth volume from distorted sections.

(2) A new image warping algorithm based on dynamic programming for computing regularized warps between adjacent serial sections. Due to the nature of sectioning distortions, we consider a class of 2D warps that can be decomposed into 1D piecewise linear deformations with elastic constraints. The representation of such a decomposable 2D warp greatly facilitates the process of warp filtering. Moreover, the decomposition allows us to extend a well-known 1D discrete minimization method called Dynamic Time Warping (Sakoe and Chiba, 1978) to compute 2D elastic image warps even between images of high resolutions. Experimental results have shown that the proposed method achieves an efficiency comparable to state-of-the-art warping methods, while the resulting deformations often match successive tissue sections with improved accuracy.

The rest of the paper is organized as follows. We describe our image warping reconstruction algorithm in Section 2. In Section 3 we present our validation framework involving quantitative measures developed from the evaluation criteria suggested by Guest and Baldock (1995), and we report the experimental results on a MRI test volume and a stack of 350 histology sec-
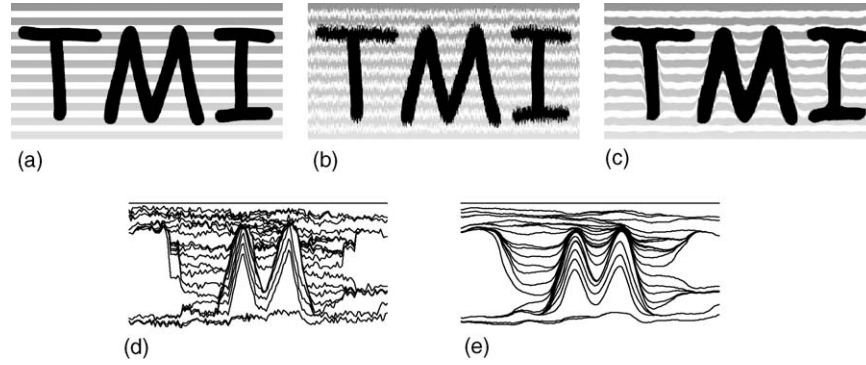
Fig. 2. Reconstruction of a smooth 2D image from deformed 1D columns using warp filtering. (a) The original grayscale image. (b) The same image with distorted columns. (c) The reconstructed image. (d and e) A subset of lines connecting corresponding points in successive columns before and after reconstruction.

tions. Finally, we discuss limitations of the method and future research in Section 4.

2. Methods

2.1. Warp filtering

To reconstruct a volume in which corresponding points on successive sections are located close to each other, warp filtering computes image warps that match each section to a *group* of neighboring sections based on pairwise warps between successive sections.

2.1.1. The algorithm

The input to the algorithm is a stack of N serial sections represented as 2D images $g_k: R^2 \to R$ ($k = 1, \ldots, N$) and pairwise warps represented as bivariate, bivalued functions $\phi_{k,k+1}: R^2 \to R^2$ ($k = 1, \ldots, N-1$) between every two successive images. The warps $\phi_{k,k+1}$ can be computed using any user-specified 2D image warping method so that the warped image $g_k \circ \phi_{k,k+1}$[1] best matches the neighboring image g_{k+1}.

The algorithm computes new warps $\Phi_k: R^2 \to R^2$ ($k = 1, \ldots, N$) that match each image g_k to a group of its neighboring images $\{g_{k-d}, \ldots, g_{k+d}\}$ ($d > 0$). Φ_k is represented as the following weighted average,

$$\Phi_k = \sum_{i=k-d}^{k+d} \gamma_i \phi_{k,i} \tag{1}$$

where γ_i ($i = k-d, \ldots, k+d$) are binomially distributed weights that approximate a Gaussian filter,

$$\gamma_i = \frac{\binom{2d}{i-k+d}}{2^{2d}},$$

and $\phi_{k,i}$ represent the warps from image g_k to each image g_i in the neighborhood $i \in [k-d, k+d]$. Given the pairwise warps $\{(\phi_{k-d,k-d+1}, \ldots, \phi_{k+d-1,k+d}\}$, $\phi_{k,i}$ can be constructed induc-

tively as[2]

$$\phi_{k,i} = \begin{cases} \phi_{k,i+1} \circ \phi_{i,i+1}^{-1}, & k-d \le i < k \\ \phi_{k,k}, & i = k \\ \phi_{k,i-1} \circ \phi_{i-1,i}, & k < i \le k+d \end{cases} \tag{2}$$

where $\phi_{k,k}$ denotes the identity.

In effect, under the warp Φ_k, each point on g_k is repositioned not to exactly match its corresponding points on a single neighboring image, but to match the weighted-average location of its corresponding points on a group of images $\{g_{k-d}, \ldots, g_{k+d}\}$. Consequently, in the warped stack of images $g_k \circ \Phi_k$ ($k = 1, \ldots, N$), high-frequency noise along lines of corresponding points through successive images, which are often induced by randomized sectioning distortions, are removed, and the distances between corresponding points on adjacent images are reduced.

We illustrate warp filtering in a 2D example in Fig. 2. Here we consider the simplified problem of reconstructing a coherent 2D image from a sequence of deformed 1D columns of pixels $g_k: R \to R$. Fig. 2a shows a grayscale image with letters "TMI", and Fig. 2b shows the same image but with each column g_k synthetically distorted using a randomized local function. To reconstruct a smooth 2D image, pairwise warps $\phi_{k,k+1}: R \to R$ that establish correspondences between points on column g_k to column g_{k+1} are computed between successive columns. In Fig. 2d, a subset of lines connecting corresponding points on successive columns is plotted based on these pairwise warps. Observe that the lines are jagged due to the local distortions that vary from column to column. Our algorithm computes the filtered warps Φ_k from pairwise warps $\phi_{k,k+1}$ and produces a stack of warped columns $g_k \circ \Phi_k$ that constitute a smooth 2D image shown in Fig. 2c. Observe in Fig. 2e that, after warping, corresponding points on successive columns lie close to each other and form smooth curves in space.

2.1.2. Comparison

The proposed algorithm is most closely related to the approach taken by Montgomery and Ross (1994). In comparison

---

[1] Operator '∘' denotes function composition such that $(f_1 \circ f_2)(x) = f_1(f_2(x))$.

[2] If pairwise warps $\phi_{k,k+1}$ are not invertible, reverse warps $\phi_{k+1,k}$ have also to be computed using the chosen image warping method between successive image pairs.

to their method, we present two major improvements. First, the correspondence between successive sections, which was established manually using contour lines by Montgomery and Ross (1994), is now computed automatically as image warps. Second, the simple two-neighbor Laplacian operator on contour lines is replaced by a more general de-noising Gaussian filter that operates on general warps between image pairs in an extended neighborhood.

Our method also offers several unique features when compared to global minimization methods (Guest and Baldock, 1995; Wirtz et al., 2004):

(1) *Simplicity*: Without the need to setup and solve a 3D minimization problem, our local approach involves only 2D image warping and simple weighted-averaging operations. In general, the algorithm can be used in conjunction with any existing 2D image registration techniques for smooth 3D reconstruction.
(2) *Stability*: In contrast with global minimization over all sections, computing pairwise warps between two individual sections is a minimization problem at a small scale and therefore less prone to errors due to variations in the input data. Such stability is greatly desired due to the systematic and random nature of section distortions.
(3) *Efficiency*: The decomposition of the reconstruction problem into warping and filtering tasks makes substantial performance increase possible through parallelization. In particular, since the pairwise warp between each image pair as well as the filtered warp for each image are computed independently of each other, linear reduction of computation time can easily be achieved in a distributed computing environment.

2.1.3. Warp representation

Although the warp filtering algorithm can be used with any image warping techniques, efficient computation using formula (1) and (2) requires appropriate representation of pairwise image warps $\phi_{k,k+1}$.

Regardless of how pairwise warps $\phi_{k,k+1}$ are represented, the filtered warp Φ_k can always be implicitly constructed by evaluating the righthand side of formula (1) and (2) at every point (x,y). This approach, however, needs to evaluate $\{\phi_{k-d,k-d+1}, \ldots, \phi_{k+d-1,k+d}\}$ for every point in the image. For efficiency, we can construct Φ_k as an explicit function using formula (1) and (2), which can then be used to directly evaluate $\Phi_k(x,y)$ at every point (x,y). The second approach requires the pairwise warps $\phi_{k,k+1}$ to be represented for the following operations on functions:

- Inversion: ϕ_{ij}^{-1}, representing the warp from image g_j to image g_i.
- Composition: $\phi_{i,j} \circ \phi_{j,k}$, representing the warp from image g_i to image g_k.
- Addition: $a\phi_{i,j} + b\phi_{i,k}$, representing a weighted average of the two warps.

Although linear functions are feasible, warps represented as higher-degree global polynomials or displacement fields in many local warping methods cannot be easily adapted to take advantage of these function operations. In the next section, we describe a new image warping method for computing a local, piecewise linear warping function that is readily represented for efficient warp filtering.

2.2. Computing image warps

It is well-known that the problem of computing the optimal warp is ill-posed (Modersitzki, 2004) and NP-complete (Keysers and Unger, 2003), hence regularization is required. For warping between successive serial cryo-sectioned images, we consider a specific type of regularized image warps such that

(1) Each 2D warp can be decomposed into a single 1D deformation in the horizontal direction and independent 1D deformations in the vertical direction for each column, and
(2) Each 1D deformation is represented as a monotonic piecewise linear function with elasticity constraints.

As we shall see in the following, such regularized warps are capable of representing local deformations that are characteristic between adjacent serial sections. In addition, the regularization allows fast warp filtering by supporting the three operations on warp functions (i.e., inversion, composition and addition) and efficient warp computation using dynamic programming.

2.2.1. Decomposing 2D warps

In our experiment (Carson et al., 2002), the frozen brain was cut coronally into 350 cryo-sections each 25 μm thick. Tissue sections were placed on slides and then a Nissl stain was applied. After staining, a coverslip was applied to each slide. Assuming that the sections have been rigidly aligned to correct the rotational and translational differences introduced when the tissue sections are placed on the slides (see Section 3.2 for details), localized differences between adjacent sections include (see examples in Fig. 1):

(1) Inherent anatomical variance between adjacent tissue sections.
(2) Regional tissue deformations in the form of vertical stretching and compacting induced both during cutting and when the section is placed onto the slide. Extreme deformations may lead to tearing and folding.
(3) Image artifacts, such as dust (dark artifact) and air bubbles (black ring-shaped artifact), introduced during coverslipping.

While the anatomical variance and image artifacts typically do not have specific orientation, regional distortions in the form of stretching and compacting mostly take place in the vertical (i.e., slicing) direction. To better represent the deformations caused by sectioning distortions while accommodating other possible local variances, we shall compute a restricted warp ϕ: $R^2 \rightarrow R^2$ of the following form:

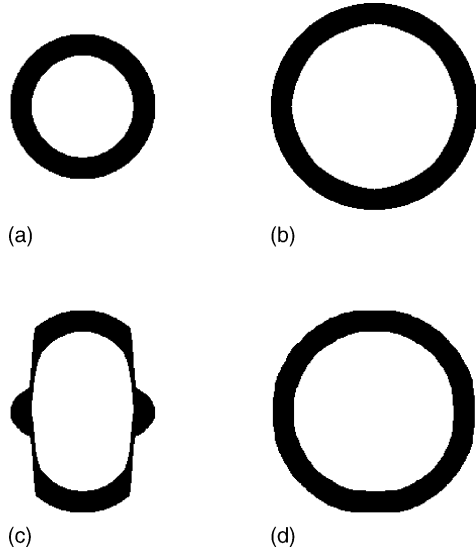$$\phi(x, y) = (\phi_X(x), \phi_Y(x, y))$$

Fig. 3. Warping from a source image (a) to a target image (b) by only computing the vertical deformation (c) and by computing deformations in both the horizontal and vertical direction (d).

where ϕ_X, ϕ_Y are single-valued functions. In other words, ϕ_X models an overall 1D deformation in the horizontal direction (i.e., shifting of image columns) and ϕ_Y models independent 1D deformations in the vertical direction for different values of x (i.e., shifting of pixels in each column). A similar decomposition has been considered previously by Agazzi et al. (1993) for text recognition. Such decomposition greatly simplifies the task of warp computation, since 1D (vector) warps can be obtained much more efficiently and robustly than 2D (image) warps.

Previous authors, such as Cox et al. (1996), have considered a more restricted form of 2D warps. In their methods, each 2D warp consists solely of independent 1D warps on each column (i.e., ϕ_X is the identity function). To illustrate the difference, Fig. 3 shows a simple example of warping between two images containing rings of different radii, a typical type of deformation between successive tissue sections. Fig. 3c shows the source image deformed using only 1D warps in the vertical direction, and Fig. 3d shows the same source image deformed using the combination of 1D warps in both vertical and horizontal directions. Observe that the horizontal warp ϕ_X is necessary to model the non-vertical variances between source and target images, which is required to handle the anatomical differences and image artifacts in tissue sections.

2.2.2. Piecewise linear representation

For convenient representations that will facilitate the function operations in warp filtering, we consider 1D warps $\{\phi_X, \phi_Y\}$ that are monotonic, piecewise linear functions. In particular, given two images s and t with $n + 1$ columns and $m + 1$ rows,

(1) The horizontal deformation ϕ_X is represented by a *pair* of piecewise linear functions $\sigma, \psi: [0, K \in Z^+] \rightarrow [0, n]$, which match column $\sigma(k)$ in image s to column $\psi(k)$ in image t for $k = 0, \ldots, K$.

(2) The vertical deformation ϕ_Y is represented by a sequence of piecewise linear function pairs $\sigma_k, \psi_k: [0, L \in Z^+] \rightarrow [0, m]$ for $k = 0, \ldots, K$, which match the point $(\sigma(k), \sigma_k(l))$ in image s to point $(\psi(k), \psi_k(l))$ in image t for $k = 0, \ldots, K$ and $l = 0, \ldots, L$.

We require functions σ, ψ and $\sigma_\Sigma, \psi_\Sigma$[3] to be *invertible*, so that the warp $\phi_{s,t}$ from image s to image t can be represented as:

$$\phi_{s,t}(x, y) = (\sigma(\hat{x}), \sigma_{\hat{x}}(\psi_{\hat{x}}^{-1}(y)))$$

where $\hat{x} = \psi^{-1}(x)$ for $x \in [0, n]$ and $y \in [0, m]$ (linear interpolation is used when subscript x assumes a non-integer value). The symmetry of σ and ψ allows the inverse warp $\phi_{t,s}$ from image t to image s to be represented in the same way with symbols σ and ψ exchanged. Such convenience will later become crucial in performing function operations during warp filtering in Section 2.2.4.

2.2.3. Elastic deformation

To prevent excessive image deformation, an image is typically modelled as an elastic material and image warps are constrained by some form of deformation energy. Due to the piecewise-linear nature of our warp representation, we consider a discrete form of deformation energy that consists of three terms:

(1) First and second order deformation energy in the X direction:

$$E_X(\sigma, \psi) = \alpha_X \sum_{k=0}^{K} \delta(k)^2 + \beta_X \sum_{k=0}^{K-1} (\delta(k+1) - \delta(k))^2$$

where $\delta(k) = \sigma(k) - \psi(k)$ and α_X, β_X are constant weights.

(2) Independent first and second order deformation energy in the Y direction for each vertical warp:

$$E_Y(\sigma_k, \psi_k) = \alpha_Y \sum_{l=0}^{L} \delta_k(l)^2 + \beta_Y \sum_{l=0}^{L-1} (\delta_k(l+1) - \delta_k(l))^2$$

where $\delta_k(l) = \sigma_k(l) - \psi_k(l)$, and α_Y, β_Y are constant weights.

(3) Coherence between neighboring vertical warp functions:

$$E_C(\sigma_\Sigma, \psi_\Sigma) = \gamma \sum_{k=0}^{K-1} \sum_{l=0}^{L} ((\sigma_{k+1}(l) - \sigma_k(l))^2 + (\psi_{k+1}(l) - \psi_k(l))^2).$$

where γ is a constant weight.

Note that the inclusion of the coherence term in (3) is necessary to ensure a smooth deformation field using a 2D warp that is decomposed into independent 1D warps.

---

[3] σ_Σ is the short hand for writing the whole sequence $\sigma_0, \ldots, \sigma_k$.
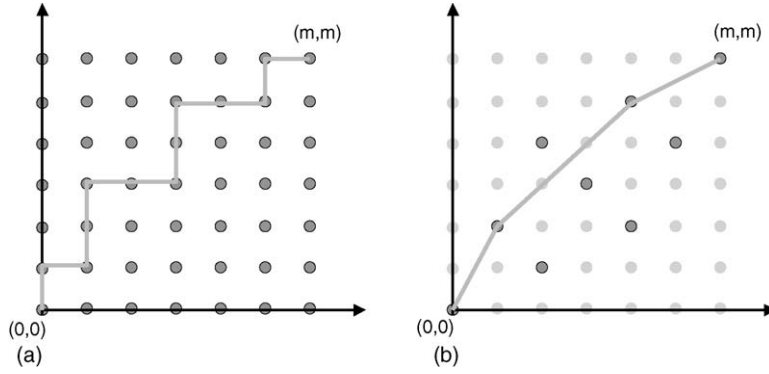
Fig. 4. Comparison of a Manhattan path (a) and a restricted path (b) whose slope is bounded between 1/2 and 2 and whose possible nodes (colored dark gray) lie in a skewed square along the diagonal.

Putting together, the warping problem that we consider becomes the minimization of the following error:

$$M(\sigma, \psi, \sigma_\Sigma, \psi_\Sigma) = \sum_{k=0}^{K} M_Y(\sigma(k), \psi(k), \sigma_k, \psi_k) + E_X(\sigma, \psi)$$
$$+ E_C(\sigma_\Sigma, \psi_\Sigma) \qquad (3)$$

where M_Y is the warping error for each independent vertical warp under the l^2-norm (e.g., sum of pixel-wise squared differences):

$$M_Y(\sigma(k), \psi(k), \sigma_k, \psi_k) = \sum_{l=0}^{L} (s(\sigma(k), \sigma_k(l)) - t(\psi(k), \psi_k(l)))^2$$
$$+ E_Y(\sigma_k, \psi_k) \qquad (4)$$

2.2.4. Dynamic programming

To compute a minimizer of (3), we first need an efficient method for minimizing the 1D warping error in (4). Dynamic Time Warping (DTW) (Sakoe and Chiba, 1978) is such a method that robustly computes a discrete minimizer in 1D matching problems. Here we describe how 1D DTW can be extended to warp 2D images by computing a discrete minimizer of (3).

2.2.4.1. 1D Warping. First we consider the 1D problem of computing the vertical warp pair (σ_k, ψ_k) between column $\sigma(k)$

$$e_{ij} = \begin{cases} \mathrm{Min} \begin{pmatrix} e_{i-2, j-1} + \mathrm{Left}(i, j) + \beta_Y \\ e_{i-1, j-2} + \mathrm{Down}(i, j) + \beta_Y \\ e_{i-3, j-3} + \mathrm{Diag}(i, j) \end{pmatrix}, & i > 0 \text{ or } j > 0 \\ \epsilon_{i, j}, & i = 0 \text{ and } j = 0 \\ \infty, & i < 0 \text{ or } j < 0 \end{cases} \qquad (5)$$

in image s and column $\psi(k)$ in image t, assuming the existence of σ and ψ. DTW discretizes $(\sigma_k(l), \psi_k(l))$ $(l = 0, \ldots, L)$ as *a path on the integer grid from* $(0, 0)$ *to* (m, m). The original DTW algorithm computes a Manhattan path from $(0, 0)$ to (m, m), as shown in Fig. 4a. However, the resulting piecewise linear functions (σ_k, ψ_k) would become non-invertible due to existence of
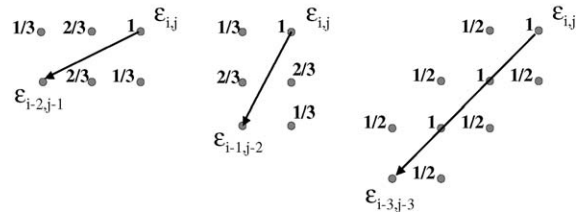

Fig. 5. Weights assigned to the entries in the error table ϵ for computing functions Left(i,j), Down(i,j) and Diag(i,j).

zero derivatives on the linear segments. To enforce invertibility, we adopt a similar technique to the method of Kovar and Gleicher (2003) that restricts the slope of every segment of the path to be 1/2, 1 or 2. An example path is shown in Fig. 4b. Using DTW, the path with bounded slope that minimizes the 1D warping error $M_Y(\sigma(k), \psi(k), \sigma_k, \psi_k)$ defined in (4) can be found in three steps:

(1) Construct a $m + 1$ by $m + 1$ error table ϵ, so that

$$\epsilon_{i, j} = (s(\sigma(k), i) - t(\psi(k), j))^2 + \alpha_Y (i - j)^2, \quad \forall i, j \in [0, m]$$

(2) Construct a $m + 1$ by $m + 1$ auxiliary table e, in which e_{ij} is the error of the minimal-error path with bounded slope from entry $(0, 0)$ to entry (i,j) in the error table ϵ. We can compute e_{ij} inductively as follows:

where each function Left, Down, and Diag computes a weighted sum of entries in ϵ according to the direction and length of the previous path segment, as shown in Fig. 5. Note that the sum of the weights is proportional to the length of the path segment in each direction.

(3) The minimal-error path that leads to $e_{m,m}$ encodes the piecewise linear functions (σ_k, ψ_k): $[0, L] \to [0, m]$ where L is
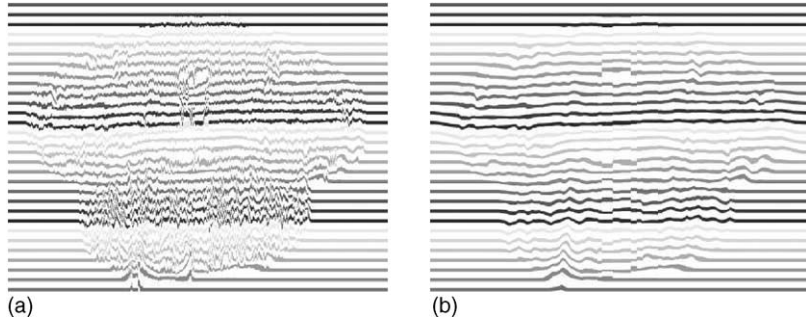
Fig. 6. A test pattern warped by applying a non-coherent warp (a) and a coherent warp (b) from sections 14 to 15 in Fig. 1.

the number of path segments and $e_{\sigma_k(l), \psi_k(l)}$ are entries along the path for $l = 0, \ldots, L$.

Both time and space complexity of the above algorithms is $O(m^2)$. However, we can dramatically improve the speed and the memory usage in practice. First, $\epsilon_{i,j}$ and $e_{i,j}$ only have to be computed for grid points i, j that lie within a skewed square of size $(m+2)/3$ along the diagonal, as shown in Fig. 4b, which effectively reduces the computation by a factor of 9. Furthermore, we can limit our computation to grid points $\{i,j\}$ within a narrow band along the diagonal, whose half-width is the maximum deformation $||(_k(l) - \psi_k(l)||$. When warping between adjacent tissue sections, we typically observe a maximum deformation of less than 10% in both horizontal and vertical directions, hence banding reduces the computation further by an approximate factor of 5. However, this percentage can be increased or decreased by the user, based on how well the sections are aligned with respect to each other, so that dynamic programming will be computed over a larger or smaller neighborhood.

2.2.4.2. 2D Warping. The 2D warping task $M(\sigma, \psi, \sigma_\Sigma, \psi_\Sigma)$ is achieved in two stages. First, observing a similar form on the right-hand side of the 2D minimization goal in (3) (ignoring the coherence term E_C) and the 1D minimization goal in (4), we consider functions (σ, ψ) such that $(\sigma(k), \psi(k))$ $(k = 0, \ldots, K)$ form *a path on the integer grid from* $(0, 0)$ *to* (n, n) with bounded slope. Similarly, using DTW, such a path that minimizes the sum of vertical warping error $\sum_{k=0}^{K} M_Y(\sigma(k), \psi(k), \sigma_k, \psi_k)$ and the horizontal elastic energy $E_X(\sigma, \psi)$ can be found in three steps:

(1) Construct a $n+1$ by $n+1$ error table ϵ, so that

$$\epsilon_{i,j} = M_Y(i, j, \bar{\sigma}_i, \bar{\psi}_i) + \alpha_X(i - j)^2, \quad \forall i, j \in [0, n]$$

where $(\bar{\sigma}_i, \bar{\psi}_j)$ is the minimal-error vertical warp between column i in s and column j in t, computed using the previous 1D algorithm.
(2) Construct a $n+1$ by $n+1$ auxiliary table e in the same way as in (5) except constant β_Y is replaced by β_X.
(3) The minimal-error path that leads to $e_{n,n}$ encodes the piecewise linear functions $(\sigma, \psi): [0, K] \to [0, n]$ where K is the number of path segments and $e_{\sigma(k), \psi(k)}$ are entries along the path for $k = 0, \ldots, K$.

Next, given (σ, ψ), the vertical warps $(\sigma_\Sigma, \psi_\Sigma)$ can be then computed using the previous 1D algorithm. However, these minimal-error vertical warps yield 2D warps that may vary significantly from column to column. For example, Fig. 6a shows the result of applying a 2D warp computed from sections 14 to 15 in Fig. 1, without considering the coherence among vertical warps, to a test pattern (a sequence of uniformly spaced horizontal bars). Notice the high-frequency noise in the warped test pattern due to the large disparity between vertical warps in neighboring columns. Typically, the true 2D deformations induced during the sectioning process are much smoother. To incorporate the coherence energy $E_C(\sigma_\Sigma, \psi_\Sigma)$, given (σ, ψ), instead of computing a single best vertical warp (σ_k, ψ_k) for every $k = 0, \ldots, K$, we compute a subset of the low-error warps $(\sigma_{k_\Sigma}, \psi_{k_\Sigma})$ in step 3 of the 1D algorithm. Finally, we add an extra dynamic programming pass to choose the best warp (σ_k, ψ_k) from each group $(\sigma_{k_\Sigma}, \psi_{k_\Sigma})$ that minimizes the sum of vertical warping error $\sum_{k=0}^{K} M_Y(\sigma(k), \psi(k), \sigma_k, \psi_k)$ and the coherence energy $E_C(\sigma_\Sigma, \psi_\Sigma)$ (each vertical warp function must be first re-parameterized to have the same domain $[0, L]$, see Section 2.2.5). Fig. 6b shows the coherent warp applied to the same test pattern, which exhibits a much smoother deformation.

The dominant operation in 2D warping is the construction of error table ϵ in step 1, which has time complexity $O(n^2 m^2)$. However, the computation can be efficiently implemented by restricting the calculation to a subset of entries and using appropriate banding, as described in 1D warping. Experiment results have shown that the speed of our algorithm is comparable to other state-of-the-art methods while producing higher quality warps.

2.2.4.3. Examples and comparison. Fig. 7 shows two examples of warping adjacent Nissl-stained cryo-sections shown in Fig. 1. In each example, the goal is to find the warp that deforms the source section s to match the target section t. Sections are represented as grayscale images of dimensions $n = 850$ and $m = 670$. We compare the result of applying our method[4] to the result of using two other popular warping methods: Automated Image Registration (AIR) (Woods et al., 1998a), and NLM Insight Seg-

---

[4] With elastic energy weights $\alpha_X = \alpha_Y = 0.00001$, $\beta_X = \beta_Y = 0.02$, $\gamma = 0.005$.
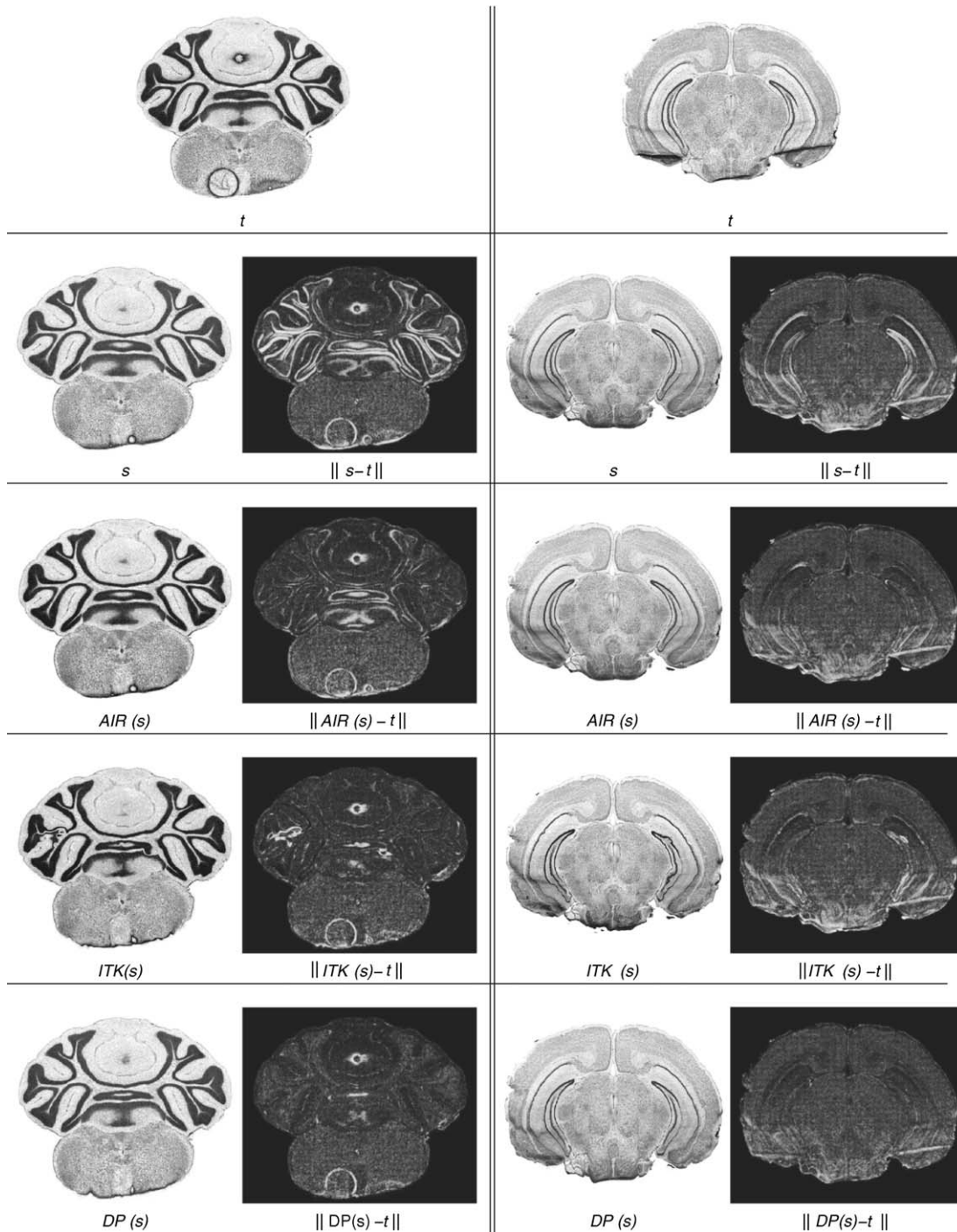
Fig. 7. Two examples of warping from a source image *s* to a target image *t* using global polynomial transformation in AIR package (AIR(*s*)), deformable registration implementation in ITK (ITK(*s*)), and the dynamic programming method presented in this paper (DP(*s*)).

mentation and Registration Toolkit (ITK) (Yoo et al., 2002). For AIR, we used the 182-parameter 12-degree global polynomial non-rigid transformation model, which provides the highest degree of deforming flexibility among all AIR models.[5] For ITK,

we used the FEM-based local deformable registration method.[6] In both AIR and ITK, we started from the standard parameter settings and modified them slightly to achieve the best deformation result on our testing images. The l^2-norm between the warped

___

[5] In our test, alignwarp program was used with model menu number (*m*) 32, threshold (*t*1, *t*2) 2, and convergence threshold (*c*) 0.00001.

[6] In our test, FEM Registration Filter was used with mean square metric, single-resolution, 8 pixels per element, 40 iterations with elasticity (*E*) and density (RhoC) set to 1e5.

Table 1
Performance comparison of three warping methods applied to the two examples in Fig. 7

	Example 1		Example 2	
	Time (s)	l^2-norm	Time (s)	l^2-norm
Source image (s)	–	11717.1	–	17069.7
Polynomial transformation (AIR(s))	36	5907.5	252	12419.5
Deformable registration (ITK(s))	1325	5005.9	1190	14314.6
Dynamic programming (DP(s))	165	3951.9	173	9207.4

The l^2-norm is computed between the source image (or warped source image) and the target image.

source image and the target image in each method as well as the computation time are compared in Table 1. Note that the dynamic programming method used in this paper achieves a comparable efficiency to a global warping method, while the quality of the resulting 2D warp is often better than a conventional local warping method when deforming between successive tissue sections.

Observe from Fig. 7 that the proposed warping method also handles extreme deformations (e.g., severe folding in the second example) and image artifacts (e.g., the big air bubble in the first example) in a reasonable manner. These random artifacts and distortions are examples of incompatible features between the source and target images, which imply that there does not exist a perfect warp that exactly matches one image to the other. Such incompatible features may also assume the form of tissue tears, or appearing (disappearing) anatomical features (e.g., the dense circular feature in upper middle of the target image in the first example) that are typical through successive tissue sections. The proposed warping method matches compatible features between the two images well while producing moderate modifications in places where incompatibility arises without triggering excessive deformations.

2.2.5. Warp operations

As we mentioned before, the piecewise linear 2D warp representation facilitates the three operations involved in warp filtering (i.e., inversion, addition and composition). First, we need to resolve a subtle problem that arises in our dynamic programming approach. That is, horizontal warps (σ, ψ) between different images may have different domain K, and even vertical warps (σ_Σ, $(_\Sigma)$) in a same 2D warp may have different domain L. For convenience, we re-parameterize every horizontal warp σ, $\psi: [0, K] \to [0, n]$ onto a fixed domain $[0, 2n]$ as follows:

$$\sigma^*(x) = \sigma(\omega^{-1}(x)), \ \psi^*(x) = \psi(\omega^{-1}(x)) \tag{6}$$

where $\omega(x): [0, K] \to [0, 2n]$ is a monotonic, invertible mapping defined as $\omega(x) = \sigma(x) + \psi(x)$. Intuitively, (6) re-parameterizes the two functions σ and ψ along the *diagonal* from $(0, 0)$ to (n, n), as shown in Fig. 8a. Every vertical warp (σ_k, ψ_k) can be diagonalized onto domain $[0, 2m]$ in a similar fashion.

2.2.5.1. 1D warp operations.

- Inversion: Given a pair of 1D warps $\{\sigma, \psi\}$, the inverted warp $\{\sigma, \psi\}^{-1}$ is easily obtained by exchanging the two functions as $\{\psi, \sigma\}$.

- Addition: Given two pairs of warps $\{\sigma, \psi\}$ and $\{\hat{\sigma}, \hat{\psi}\}$ with domain $[0, 2n]$, the addition $a\{\sigma, \psi\} + b\{\hat{\sigma}, \hat{\psi}\}$ is simply computed as the symmetric warp $\{a\sigma + b\hat{\sigma}, a\psi + b\hat{\psi}\}$. Note that the new warping functions preserve the monoticity, invertibility and the same domain. Fig. 8b illustrates this operation for $a = b = 1/2$.
- Composition: The composition of two pairs of warps $\{\sigma, \psi\}$ and $\{\hat{\sigma}, \hat{\psi}\}$ with domain $[0, 2n]$, represented as $\{\sigma, \psi\} \circ \{\hat{\sigma}, \hat{\psi}\}$, is computed as a new symmetric warp $\{\sigma(\psi^{-1}), \hat{\psi}(\hat{\sigma}^{-1})\}$. Note that the new warping functions are defined on the domain $[0, n]$, so diagonalization (6) must be performed to re-parameterize $\sigma(\psi^{-1})$ and $\hat{\psi}(\hat{\sigma}^{-1})$ onto domain $[0, 2n]$.

2.2.5.2. 2D warp operations. Let $\{\sigma, \psi, \sigma_\Sigma, \psi_\Sigma\}$ be a symmetric 2D warp between two images, the inverted warp is easily obtained by exchanging the symbols σ and ψ.

Given two 2D warps $\{\sigma, \psi, \sigma_\Sigma, \psi_\Sigma\}$ and $\{\hat{\sigma}, \hat{\psi}, \hat{\sigma}_\Sigma, \hat{\psi}_\Sigma\}$, the addition can be computed as simply $\{a\sigma + b\hat{\sigma}, a\psi + b\hat{\psi}, a\sigma_\Sigma + b\hat{\sigma}_\Sigma, a\psi_\Sigma + b\hat{\psi}_\Sigma\}$. Composition of the two warps can be performed as follows. First, the horizontal warp in the composite 2D warp, denoted by $\{\bar{\sigma}, \bar{\psi}\}$, is computed as the composition of horizontal warps $\{\sigma, \psi\} \circ \{\hat{\sigma}, \hat{\psi}\}$. Then, the vertical warps in the composite 2D warp, denoted by $\{\bar{\sigma}_\Sigma, \bar{\psi}_\Sigma\}$, are computed such that

$$\{\bar{\sigma}_k, \bar{\psi}_k\} = \{\sigma_{k_1}, \psi_{k_1}\} \circ \{\hat{\sigma}_{k_2}, \hat{\psi}_{k_2}\}$$

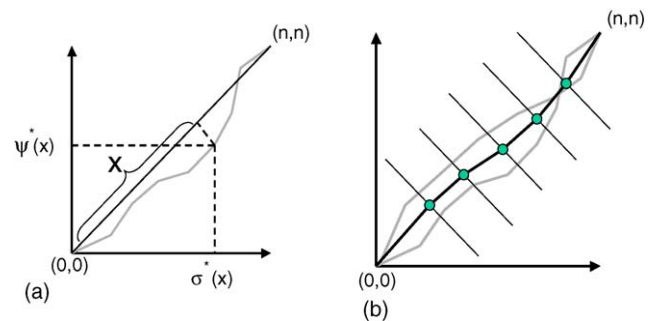where $k_1 = \sigma^{-1}(\bar{\sigma}(k))$ and $k_2 = \hat{\psi}^{-1}(\bar{\psi}(k))$.

Fig. 8. (a) Re-parameterized functions σ^* and ψ^* of the restricted path (colored in gray) along the diagonal from $(0, 0)$ to (n, n). (b) Symmetric addition of two diagonalized warps (colored in gray), the resulted warp is plotted in solid black.

3. Results

Here we report the validation methods and the corresponding results for evaluating the effectiveness of the proposed reconstruction method both qualitatively (i.e., by visual examination) and quantitatively (i.e., by computing distance and smoothness measures). These methods form a general framework for validating any warp-based 3D reconstruction algorithm. The evaluation is carried out on two sets of data: a synthetic volume with known distortions, and real serial sections with unknown distortions. All computations are performed on a commodity PC with 1.5 GHz AMD Athlon processor and 2.5 GB memory.

3.1. Using a synthetic volume

We first apply our reconstruction method to a stack of synthetically distorted cross-sections from an existing 3D volume. In our experiment, we use a MRI volume of the C57BL/6J mouse brain that has been generously provided by the LONI group at UCLA. The volume has dimension 256^3 with a uniform spacing of 56 μm. Although this dataset is not representative of the current state-of-the-art MRI technology, which may generate images with higher resolutions, it is the best synthetic volume that has been made available to us, and we are using the dataset for the sole purpose of validating our reconstruction method.

Fig. 9a shows a sagittal cross-section of the MRI volume. We take each of the 256 coronal cross-sections of the volume and apply a B-spline based local image distortion that varies randomly from section to section. Fig. 9b and c shows one of the coronal sections from the MRI volume before and after the synthetic distortion. Notice that deformation mainly takes place in the vertical direction so as to simulate the actual cutting distortions in real cryo-sections. The sagittal cross-sectional view of the distorted volume is shown in Fig. 9d.

The reconstruction proceeds by first computing the pairwise warps between successive sections followed by warp filtering with width $d = 5$. The computation took 742 s (12 min and 22 s) in total, with 661 s spent on warp computation and 81 s on warp filtering and final image deformation.

3.1.1. Visual validation

Fig. 9e shows the sagittal cross-section at the same location through the reconstructed stack. In comparison with Fig. 9d, our reconstruction method removes the majority of high frequency noises caused by synthetic distortions and is able to form smooth and easily recognizable anatomical structures.

3.1.2. Quantitative validation

We further compute the l^2-norm from each coronal cross-section in the original MRI volume to the same section after synthetic distortion and after reconstruction, which are plotted in Fig. 10. The sum of all section-wise l^2-norm are reduced by 46% after automatic reconstruction. The l^2-norm quantitatively measures how close the reconstructed volume is to the real volume. However, one should realize that such measure is imperfect in that exact restoration of the un-distorted volume is not the goal
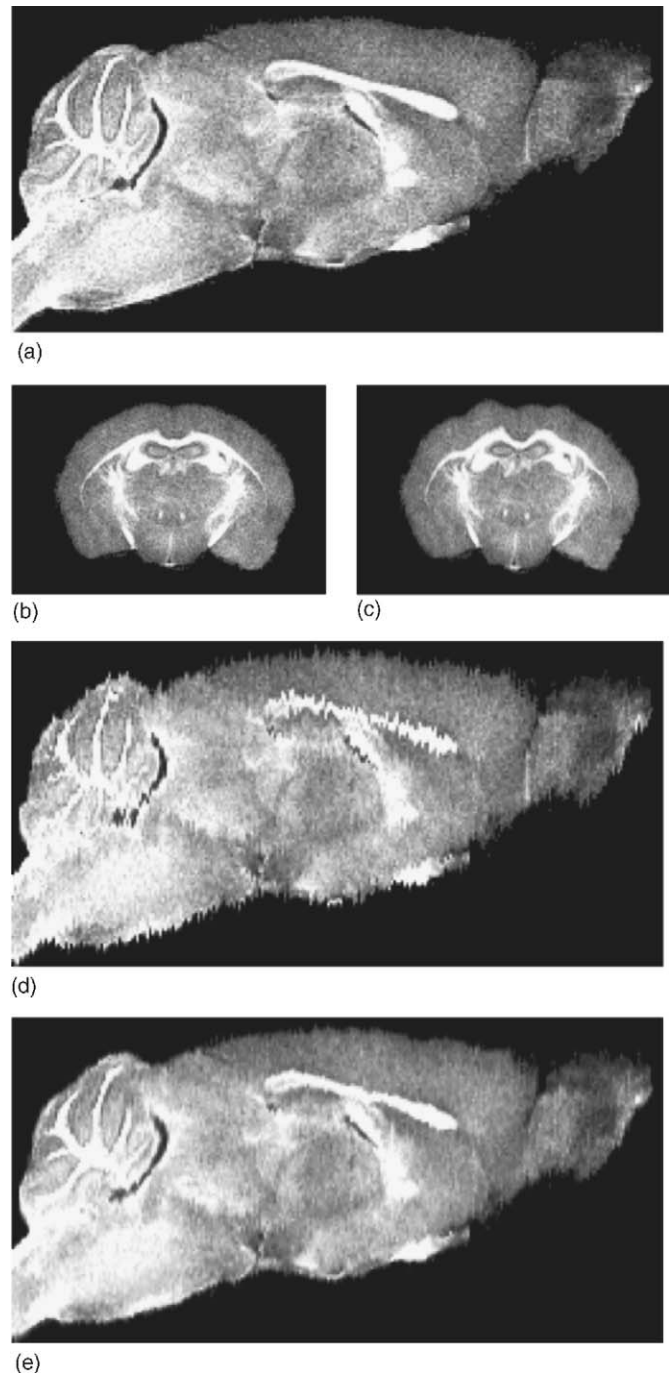

(a)

(b)   (c)

(d)

(e)

Fig. 9. Validation of warp filtering using an existing 3D volume. (a) A sagittal cross-section of a MRI volume of the mouse brain. (b) A coronal section from the MRI volume. (c) The same section in (b) after synthetic distortion. (d) Sagittal cross-section of the MRI volume after each coronal section has been randomly distorted. (e) Sagittal cross-section of the reconstructed volume from distorted sections using warp filtering.

of our reconstruction. Since there is no way to "undo" a distortion without the knowledge of the image before the distortion, the goal of our reconstruction is to instead produce a *smooth* volume representation. In the next experiment we present a better criteria that measures the *smoothness* of the reconstructed volume.
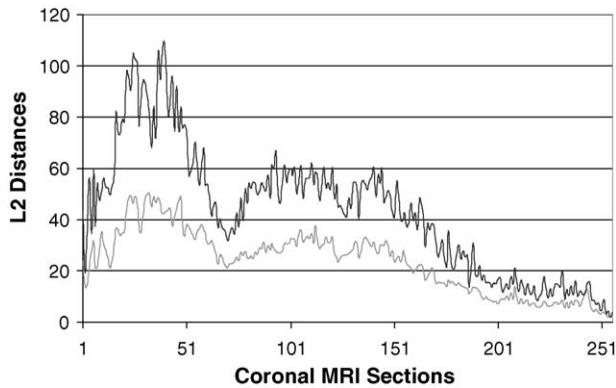
Fig. 10. The l^2-norm from each of the 256 coronal sections of a MRI volume to the same section after synthetic distortion (top curve) and to the reconstructed section after warp filtering (bottom curve).

3.2. Using serial sections

We next test our method in reconstructing a 3D volumetric representation of the cell-density for the mouse brain from serial sections. The input is a stack of 350 Nissl-stained images acquired by cyro-sectioning coronally a single frozen adult mouse brain (data preparation is detailed in Section 2.2.1). Each image is size 850×670 pixels at a resolution of 15 μm per pixel.

3.2.1. Pre-processing

The images were first registered using rigid-body transformations that align the symmetry axis of each coronal section to the vertical midline of the image and vertically adjust the sections using reference sagittal histology sections from Paxino's Atlas (Paxinos and Franklin, 2000). The axis alignment ensures the vertical direction of sectioning distortion, to which our warping method is tailored. The vertical adjustment avoids causing a straight cylinder to be reconstructed as a sloping cylinder by warp filtering. The resulting stack exhibits a natural 3D profile with minimal translational and rotational differences between adjacent sections. Fig. 12 shows synthetic cross-section cuts in the sagittal (a1) and horizontal (b1) direction through the rigid-body aligned stack.

Since the distance between adjacent sections (25 μm) far exceeds the size of a cell (∼10 μm), matching individual cells on two adjacent sections is not practical. To avoid aligning cellular details while matching macro features (e.g., the dark folds of the cerebellum) during warp computation, we first apply a smoothing filter on the tissue sections before computing the pairwise warps. Although a Gaussian filter could be used, an edge-preserving filter is more appropriate in that boundaries between anatomical structures are retained. Fig. 11 shows the result of applying a bilateral filter (Tomasi and Manduchi, 1998) that we used in our experiment (other filters may also be used). Note that the filtered image exhibits clearer boundaries between different anatomical features. After performing warp filtering on the bilaterally filtered images, the final warps would then be applied to the original images for accurate reconstruction.
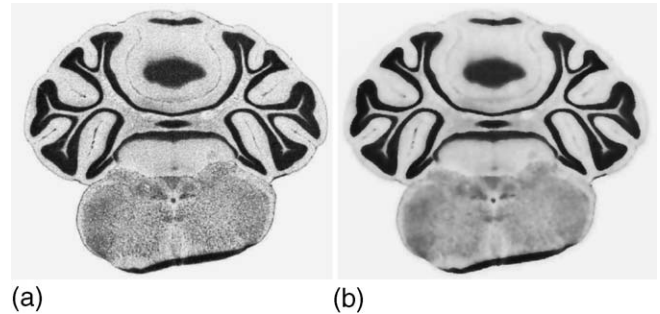

Fig. 11. A tissue section before (a) and after (b) applying the bilateral filter.

3.2.2. Performance

The computation of 349 pairwise warps between successive sections consumes 27 MB memory and a total of 976 min (16 h and 16 min), averaging 168 s for each single warp. The performance of the subsequent warp filtering stage at different filtering width d is summarized in Table 2. We want to point out that since every pairwise warp $\phi_{k,k+1}$ is computed independently of each other (and so is every filtered warp Φ_k), both warp computation and filtering can be greatly sped up by distributing the computation of different warps to different processors. Using this simple parallelization scheme on a cluster of 16 processors, for example, all pairwise warps would be computed within an hour while warp filtering would finish in approximately 10 min at $d = 20$.

3.2.3. Visual validation and comparison

Fig. 12 shows cross-section cuts in the sagittal (a2) and horizontal (b2) direction through the smooth volume reconstructed using filtering width $d = 20$ (while the filtered warps are computed from bilaterally filtered images, they are applied to the original images). Note that anatomical features become much more coherent in the reconstructed volume. The reconstruction also recovers the shape of some key structures, such as the folds in the cerebellum to the left and the hippocampus region in the middle, whose profiles are barely recognizable or completely illegible in the original stack.

For qualitative validation, we compare the reconstructed volume to real histology sections from Paxino's Atlas (Paxinos and Franklin, 2000) at similar sagittal (a4) and horizontal (b4) locations. We find that the shapes of the anatomical structures recovered in the reconstructed volume closely resemble the shapes of the corresponding structures in real tissue sections.

Table 2
Comparison of performance for different filtering width d and the corresponding mean and maximum smoothness measure S_k of the reconstructed volume

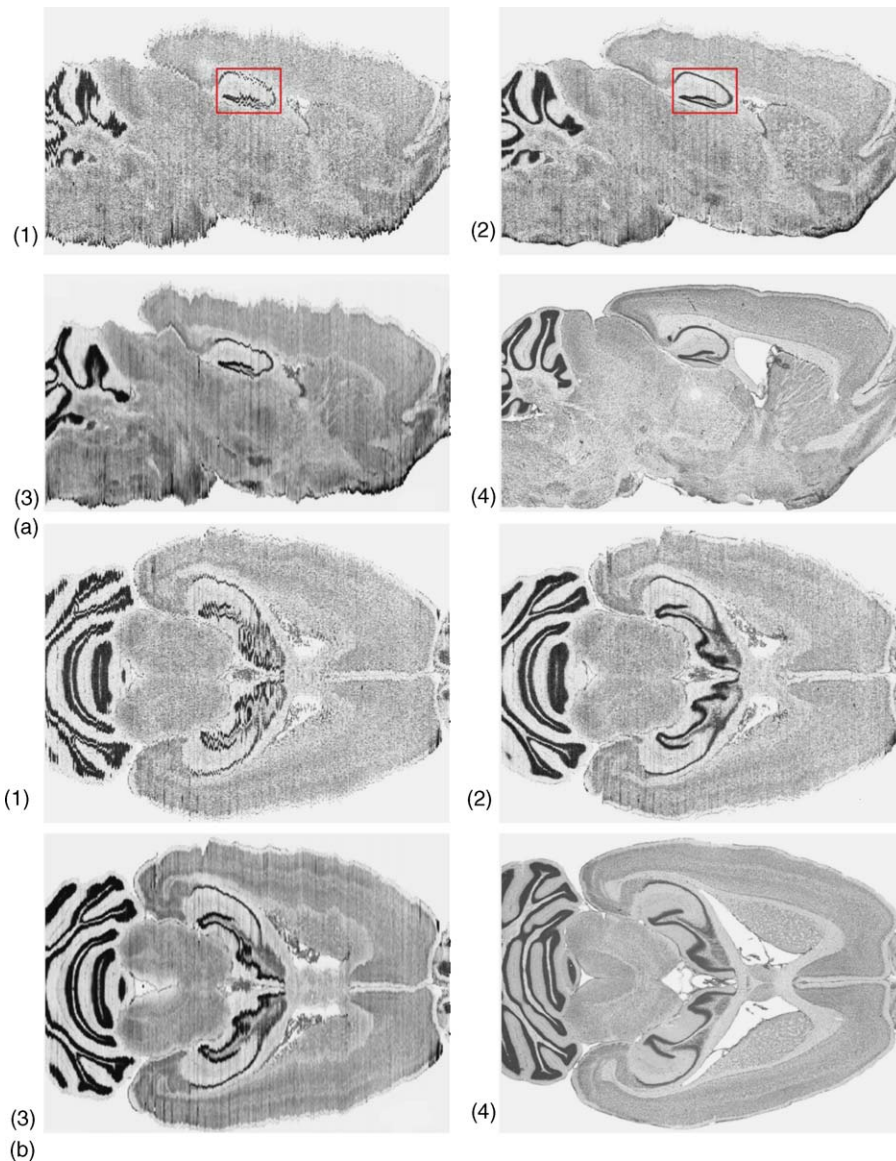| Filter width, d	Time (min)	Memory (MB)	Mean S_k	Maximum S_k
0[a]	–	–	85.43	1165.68
1	4.4	37	7.71	74.39
2	17.2	91	3.15	24.83
5	46.6	178	1.33	7.04
10	83.4	269	1.09	6.22
20	172.1	449	1.04	6.15

[a] Original stack.

Fig. 12. Comparison of sagittal (a1–a4) and horizontal (b1–b4) cross-sections: the original stack of 350 brain sections after rigid-body alignment (a1 and b1), the reconstructed volume using our method (a2 and b2) and reconstruction by the method of Guest and Baldock (1995) (a3 and b3) on filtered images, and histological sections from the Paxino's atlas (Paxinos and Franklin, 2000) at sagittal plate no. 110 (a4) and horizontal plate no. 157 (b4).

To further demonstrate of the effectiveness of our method, we performed visual comparison with the results by applying the FEM-based reconstruction method of Guest and Baldock (1995). The key idea behind their method is to model each section as a thin elastic plate, which is constrained so that points can only move within the plane. Similarly to our method, their method assumes that the sections have been coarsely registered, and computes point correspondences between every two successive sections using statistical tests of pixel populations in small local neighborhoods. Forces are then applied to each plate by matching corresponding points, and the system is released and allowed to reach equilibrium in a global minimization system involving all sections. In contrast, our method are mainly different in the use of the new warping method for computing correspondence between two sections and in the replacement of the whole-system minimization process by the local, direct

Gaussian averaging step. The local property of our method allows more direct control over the smoothness of the volume. For example, we compare in Fig. 12a2 and a3 and b2 and b3 the results of applying our method and the method of Guest and Baldock (1995). Please note that the input data that we provided Baldock's group were the bilaterally filtered coronal images, and the deformations computed by their method were applied to the same set of filtered images instead of the original images. Observe that in comparison, our method reconstructs smoother anatomical shapes and tissue boundaries by using direct and localized smoothing operations.

3.2.4. Quantitative validation

Objective validation with serial sections is difficult due to the lack of the knowledge of the original object before sectioning. One commonly used technique is to measure the correlation
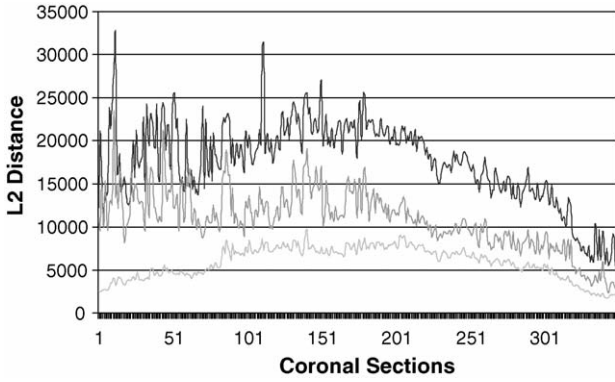
Fig. 13. The l^2-norm between successive images in the stack of 350 coronal sections before (top curve) and after (middle curve) rigid-body alignment. The elastic warping error between successive sections (bottom curve) is computed as the l^2-norm between the warped image $g_k \circ \Phi_k$ and image g_{k+1} for $k = 1, \ldots, 349$.

between adjacent sections before and after reconstruction, as done by Wirtz et al. (2004). However, as pointed out by Guest and Baldock (1995), such criteria are inappropriate since the purpose of reconstruction is not to exactly match one section to the next section. Inspired by the smoothness measure proposed by Guest and Baldock (1995), we propose the following two-step procedure to measure the smoothness of our reconstruction:

3.2.4.1. Correspondence evaluation. We first evaluate the quality of the pairwise warps $\phi_{k,k+1}$ that establish the correspondences between adjacent images g_k, g_{k+1} for $k = 1, \ldots, N$. This is accomplished by computing the l^2-norm between each warped image $g_k \circ \phi_{k,k+1}$ and target image g_{k+1}, as plotted in Fig. 13, which reveals the quality of matching. For comparison, we plot the l^2-norm between successive images in the original stack as well as in the rigidly registered stack. Observe that the pairwise warps reduce the pixel-wise differences between neighboring sections substantially by 60–90% in comparison to the original stack. Correspondence evaluation is a necessary step as the smoothness of corresponding features will make no sense if the correspondences are computed incorrectly.

3.2.4.2. Smoothness evaluation. We compute a smoothness measure S_k on each section k as follows:

$$S_k = \frac{\sum_{i=0}^{n} \sum_{j=0}^{m} (B(i, j) + C(i, j) - 2A(i, j))^2}{nm} \quad (7)$$

where $A(i, j) = \Phi_k^{-1}(i, j)$ denotes the location of (i, j) in the warped image $g_k \circ \Phi_k$, $B(i, j) = \Phi_{k-1}^{-1}(\phi_{k-1,k}(i, j))$ and $C(i, j) = \Phi_{k+1}^{-1}(\phi_{k,k+1}^{-1}(i, j))$ denote the locations of the corresponding points of (i, j) on the two neighboring sections $g_{k-1} \circ \Phi_{k-1}$ and $g_{k+1} \circ \Phi_{k+1}$ respectively after warping. As a reminder, ϕ refers to pairwise warps between adjacent sections, and Φ refers to the filtered warp. S_k is similar to the CAM measure proposed by Guest and Baldock (1995), which effectively measures how close a point lies to its corresponding points on the neighboring sections. For comparison, the smoothness mea-

sure of each section before reconstruction is computed by setting Φ_{k-1}, Φ_k and Φ_{k+1} to be identity functions.

Table 2 reports the mean and maximum smoothness measure S_k for all 350 sections in the reconstructed volume for different filtering width d. Observe that S_k decreases by an order of magnitude just by warp filtering using a single neighbor on each side of a section (i.e., $d = 1$). A reduction by two orders of magnitude is observed in both the mean and maximum S_k for $d \geq 5$. To interpret the numbers, we notice from (7) that S_k reflects the average distance from every point $A(i, j)$ to the mid-point of its two corresponding points $(B(i, j) + C(i, j))/2$. For example, at $d = 20$, $S_k = 1.04$, hence the average distance $||(B(i,j) + C(i, j))/2 - A(i,j)|| = 0.51$. In other words, in the reconstructed volume, each point deviates from the middle of its two corresponding points on the neighboring sections by only 0.51 pixels on average. These results appear to improve on the experimental results of a global reconstruction method reported by Guest and Baldock (1995) that uses a similar measurement of smoothness.

4. Discussion

4.1. Noise removal using majority filtering

Despite the use of elastic warping, image artifacts such as tissue folds and air bubbles remain after reconstruction. This is because the warping functions that we compute are one-to-one mapping functions, and such invertibility of the mapping is preserved during warp filtering. In other words, all image features in the original data will be preserved after reconstruction, and no extra features will be added. Yet in some applications, such as visualization, image artifacts are not desirable. An effective solution to clean up the artifacts that we found is to apply a *majority filter* to corresponding points through successive images based on the pairwise image warps. Specifically, for each pixel on an image, if the intensity of the pixel is the highest or the lowest among its corresponding pixels in a group of nearby images, we let it inherit the average intensity of its neighbors (see Fig. 14 top). The filter removes outliers along matched pixels through different images, which are often introduced by bubbles or tissue folding on a single image. The bottom of Fig. 14 shows a coronal section in the reconstructed volume before and after the majority filter is applied. Observe that the dark folds in the original image are much less apparent in the filtered image.

For efficient implementation, majority filtering of pixel intensities can be performed as the same time when the Gaussian filter is applied to the pairwise image warps. Fig. 15a and b show the cross-section cuts of the reconstructed volume after incorporating the majority filter, which is applied to three neighbors on each side of a given pixel. In comparison with Fig. 12a2 and b2, the majority filtering along lines of corresponding points through successive sections improve the smoothness of the appearance.

It should be noted that the purpose of majority filtering is to ensure a smooth appearance of the reconstructed volume. As we have stated in the introduction, the motivation of our work is to be able to construct a smooth anatomical atlas from the histological sections. Such atlas may not be completely accurate
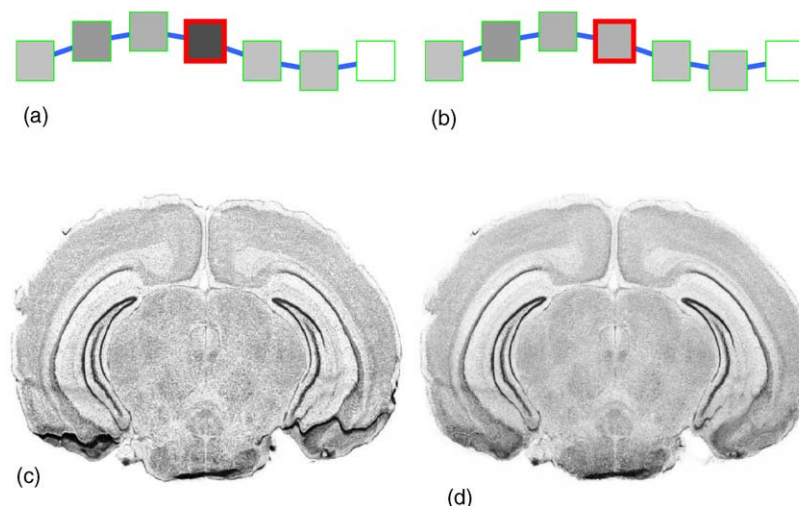
Fig. 14. (a) A dark pixel and its corresponding pixels on the neighboring sections. (b) Using majority filter, the dark pixel is replaced with the average pixel intensity of its neighbors. (c) A reconstructed coronal section. (d) The same coronal section after the majority filter is applied to the stack.

to the original brain subject due to unknown image distortions. However, it provides a good starting point for further examination of the 3D brain anatomy. More importantly, the atlas provides a common coordinate system onto which experimental sections collected from different individuals can be registered, and cross-individual comparisons can be performed. An example of mapping and comparison of brain data using a geometric atlas is provided in the work of Ju et al. (2003).

4.2. Smoothness of the volume

The smoothness of the reconstructed volume using our method is controlled by the filtering width d. Although d does not have direct physical meaning, it intuitively determines the number of consecutive sections that will be used for computing an average image for the middle section. Inappropriate choices of d, however, may cause the method to either flatten the highly curved or sharp features (when d is too big) or fail to remove the distortion errors (when d is too small). This problem of obtaining the desired amount of smoothness in the reconstructed volume arises from all smoothness-based reconstruction methods (Guest and Baldock, 1995; Wirtz et al., 2004).

A good choice of d depends on the input data. In most cases, the scale of randomized tissue distortions is small in compari-

son to the size of anatomical features, and the value of d can be determined via experiments and visual evaluation. For example, Fig. 16 shows the hippocampus region in a sagittal cross-section cut through the original stack of 350 images and the reconstructed stack with filter width $d = 5$, 10, 20. Note that although a choice of $d = 5$ already achieves a good local smoothness measure in Table 2, the dark stripe at the top of hippocampus still exhibits low-frequency noise inherent to the input data, which spans a larger neighborhood and requires a wider filtering width. At $d = 20$, the waviness is gone and the region becomes smooth. Still, the reconstructed volume needs to be examined to ensure that other key anatomical features have not been overly smoothed.

Sometimes a reasonable choice of d fails to exist, for example, when the scale of the largest distortion noise exceeds that of the smallest anatomical feature. To remove large distortions in one place while preserving small meaningful features in other locations, a more flexible method is to find independent filter widths d_k for each section g_k or even independent widths $d_k(i, j)$ for each individual pixel $g_k(i, j)$. These local filter widths will adapt to the regional noise-signal ratio based on the correspondence information between adjacent sections. Such local choices of filter widths can be easily incorporated into the warp filtering framework in (1) and (2). Moreover, edge-preserving filters,
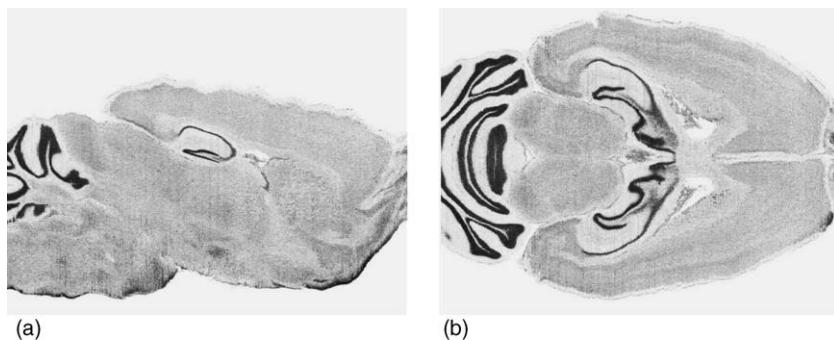

Fig. 15. Synthetic cross-section of the reconstructed brain volume after majority filtering in the sagittal (a) and horizontal (b) directions.
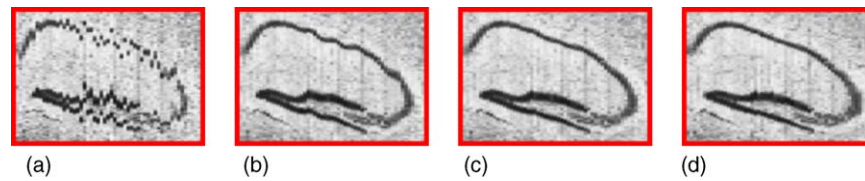
Fig. 16. Closeup looks at the hippocampus region in the synthetic sagittal cross-section of the rigid-body aligned stack (a) and the elastically aligned stack with filter width $d = 5$ (b), $d = 10$ (c) and $d = 20$ (d). The region is indicated as red boxes in Fig. 12a1 and a2.

such as a bilateral filter (Jiang et al., 2003), can be used in place of the Gaussian filter during warp filtering for preserving sharp features in the reconstructed volume that are often flattened out using the Gaussian filter.

5. Limitation of 2D warps

Any restricted image deformation, such as the regularized 2D warp described in Section 2.2, has its limitations. Specifically, our warp representation only models those 2D deformations that can be characterized by non-uniform deformations in the vertical direction and uniform deformation in the horizontal direction. Eligible deformations include translation, scaling, uniform growth/shrinkage of anatomical structures in the horizontal direction, and sectioning distortions in the vertical direction. The deformations that are not well modelled include rotation and non-uniform growth/shrinkage of anatomical structures in the horizontal direction. Although rotational differences between adjacent sections can be eliminated using rigid-body registration, we do observe, in a few cases, non-uniform variance of anatomical structures in the horizontal direction that are not modelled accurately by our method. Even for these rare cases, our method behaves in a reasonable manner, as observed from the uniform appearance of the graph of warping error in Fig. 13. To better handle a wider range of 2D deformations, we are currently investigating more flexible ways of decomposing 2D warps into 1D warps, such as the idea of allowing columns to bend and break during the horizontal deformation as discussed in (Levin and Pieraccini, 1992; Uchida and Sakoe, 1998; Ronee et al., 2001).

While we will continue our research on more accurate and efficient warping methods, we notice that comparing the quality of 2D warps generated by various methods is difficult due to the lack of standard benchmark data. In an effort to promote the development of benchmarks for 2D image warping methods as well as for 3D reconstruction methods, we have made the 350 histological sections in our experiment as well as our preliminary reconstruction result available for download at http://www.geneatlas.org/gene/data/histology.zip. We encourage those interested in this problem to apply their methods to this data.

References

Agazzi O, Kuo S, Levin E, Pieraccini R. Connected and degraded text recognition using planar hidden markov models. In: IEEE international conference on acoustics, speech, and signal processing (ICASSP-93), vol. 5, April; 1993. p. 113–6.

Ali W, Cohen F. Registering coronal histological 2-d sections of a rat brain with coronal sections of a 3-d brain atlas using geometric curve invariants and b-spline representation. IEEE Trans Med Imaging 1998;17(6):957–66.

Andrey P, Maurin Y. Free-d: an integrated environment for three-dimensional reconstruction from serial sections. J Neurosci Methods 2005;145:233–44.

Antoine Maintz JB, Viergever MA. A survey of medical image registration. Med Image Anal 1998;2(1):1–37.

Armstrong J, Kaiser K, Mller A, Fischbach K, Merchant N, Strausfeld N. Flybrain, an on-line atlas and database for the drosophila nervous system. Neuron 1995;15(1):17–20.

Berthold C, Rydmark M, Corneliuson O. Estimation of sectioning compression and thickness of ultrathin sections through vestopal-w-embedded cat spinal roots. J Ultrastruct Res 1982;80:42–52.

Bookstein F. Principal warps: thin-plate splines and the decomposition of deformations. IEEE Trans Pattern Anal Machine Intell 1989;11:567–85.

Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. In: Proceedings of International Conference on Computer Vision, vol. 1, September; 1999. p. 377–84.

Boyle J, Jenkins J, McKay I, McPhaden A, Lindop G. An assessment of the distortion of arteries due to sectioning in endomyocardial biopsies. J Pathol 1997;181(2):243–6.

Brey E, King T, Johnston C, McIntire L, Reece G, Patrick C. A technique for quantitative three-dimensional analysis of microvascular structure. Microvasc Res 2002;63:279–94.

Brune R, Bard J, Dubreuil C, Guest E, Hill W, Kaufman M, Stark M, Davidson D, Baldock R. A three-dimensional model of the mouse at embryonic day 9. Dev Biol 1999;216(2):457–68.

Cannestra A, Santori E, Holmes C, Toga A. A three-dimensional multimodality brain map of the nemestrina monkey. Brain Res Bull 1997;43(2):141–8.

Carson J, Thaller C, Eichele G. A transcriptome atlas of the mouse brain at cellular resolution. Curr Opin Neurobiol 2002;12(5):562–5.

Chen X, Ai Z, Rasmussen M, Bajcsy P, Auvil L, Welge M, Leach L, Vangveeravong S, Maniotis A, Folberg R. Three-dimensional reconstruction of extravascular matrix patterns and blood vessels in human uveal melanoma tissue: techniques and preliminary findings. Invest Ophthalmol Visual Sci 2003;44(7):2834–40.

Cox I, Hingorani S, Rao S, Maggs B. A maximum likelihood stereo algorithm. Comput Vis Image Und 1996;63(3):542–67.

Deverell M, Bailey N, Whimster W. Tissue distortion in three dimensional reconstruction of wax or plastic embedded microscopic structures. Pathol Res Pract 1989;185(5):598–601.

Durr R, Peterhans E, von der Heydt R. Correction of distorted image pairs with elastic models. Eur J Cell Biol 1989;48(25):85–8.

Gabrani M, 1998. Multidimensional spline theory and surface-based alignment of brains. PhD Dissertation, Elect. Comput. Eng., Drexel University, Philadelphia, PA.

Gefen S, Tretiak O, Nissanov J. Elastic 3-d alignment of rat brain histological images. IEEE Trans Med Imaging 2003;22(11):1480–9.

Glasbey C. A review of image warping methods. J Appl Stat 1998;25:155–71.

Guest E, Baldock R. Aautomatic reconstruction of serial sections using the finite element method. BioImaging 1995;3:154–67.

Guest E, Berry E, Baldock R, Fidrich M, Smith M. Robust point correspondence applied to two- and three- dimensional image registration. IEEE Trans Pattern Anal Machine Intell 2001;23:1–15.

Hajnal JV, Hill DLG, Hawkes DJ. Medical Image Registration. CRC Press; 2001. ISBN 0849300649.

He L, Sarrafizadeh R, Houk J. Three-dimensional reconstruction of the rubro-cerebellar premotor network of the turtle. Neuroimage 1995;2(1):21–33.

Hess A, Lohmann K, Gundelfinger ED, Scheich H. A new method for reliable and efficient reconstruction of 3-dimensional images from autoradiographs of brain sections. J Neurosci Methods 1998;84:77–86.

Jiang W, Baker M, Wu Q, Bajaj C, Chiu W. Applications of a bilateral denoising filter in biological electron microscopy. J Struct Biol 2003;144:114–22.

Jones A, Milthorpe B, Howlett C. Measurement of microtomy-induced section distortion and its correction for 3-dimensional histological reconstructions. Cytometry 1994;15(2):95–105.

Ju T, Warren J, Eichele G, Thaller C, Chiu W, Carson J. A geometry database for gene expression data. In: Eurographics Symposium on Geometry Processing; 2003.

Karen P, Jirkovska M, Tomori Z, Demjenova E, Janacek J, Kubinova L. Three-dimensional computer reconstruction of large tissue volumes based on composing series of high-resolution confocal images by gluemrc and linkmrc software. Microsc Res Tech 2003;62(5):415–22.

Keysers D, Unger W. Elastic image matching is np-complete. Pattern Recogn Lett 2003;24(1–3):445–53.

Kim B, Boes J, Frey K, Meyer C. Mutual information for automated unwarping of rat brain autoradiographs. Neuroimage 1997;5:31–40.

Kovar L, Gleicher M. Flexible automatic motion blending with registration curves. In: Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation. San Diego, California: Eurographics Association; 2003. p. 214–224. ISBN 1-58113-659-5.

Lamers W, Laan A, Huijsmans D, Smith J, Los J, Strackee J. Deformation-corrected computer-aided 3-d reconstruction of immunohistochemically stained sections of embryonic organs. Eur J Cell Biol 1989;48:103–6.

Lester H, Arridge SR. A survey of hierarchical non-linear medical image registration. Pattern Recognit 1999;32(1):129–49.

Levin E, Pieraccini R. Dynamic planar warping for optical character recognition. In: Proceedings of ICASSP, vol. III; 1992. p. 149–52.

MacKenzie-Graham A, Lee E, Dinov I, Bota M, Shattuck D, Ruffins S, Yuan H, Konstantinidis F, Pitiot A, Ding Y, Hu G, Jacobs R, Toga A. A multimodal, multidimensional atlas of the c57bl/6j mouse brain. J Anat 2004;204(2):93–102.

Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. IEEE Trans Med Imaging 1997;16:187–98.

Mega M, Chen S, Thompson P, Woods R, Karaca T, Tiwari A, Vinters H, Small G, Toga A. Mapping histology to metabolism: Coregistration of stained whole-brain sections to premortem pet in alzheimers disease. Neuroimage 1997;5:147–53.

Modersitzki J. Numerical methods for image registration. Oxford University Press; 2004.

Montgomery K, Ross M. Improvements in semiautomated serial-section reconstruction and visualization of neural tissue from tem images. Three-Dimensional Microsc (Proc SPIE) 1994;2184:264–7.

Nikou C, Heitz F, Nehlig A, Namer IJ, Armspach J-P. A robust statistics-based global energy function for the alignment of serially acquired autoradiographic sections. J Neurosci Methods 2003;124:93–102.

Ourselin S, Bardinet E, Dormont D, Malandain G, Roche A, Ayache N, Tande D, Parain K, Yelnik J. Fusion of histological sections and mr images: towards the construction of an atlas of the human basal ganglia. In: Proceedings of MICCAI 4th International Conference, Utrecht, The Netherlands; 2001a. p. 743–51.

Ourselin S, Roche A, Subsol G, Pennec X, Ayache N. Reconstructing a 3D structure from serial histological sections. Image Vision Comput 2001b;19(1–2):25–31.

Paxinos G, Franklin K. The mouse brain in stereotaxic coordinates. 2nd ed. Academic Press; 2000.

Pitiot A, Malandain G, Bardinet E, Thompson PM. Piecewise affine registration of biological images. In: International Workshop on Biomedical Image Registration, Philadelphia, USA; 2003.

Rangarajan A, Chui H, Bookstein FL. The softassign procrustes matching algorithm. Inform Proc Med Imaging 1997:29–42.

Ronee M, Uchida S, Sakoe H. Handwritten character recognition using piecewise linear two-dimensional warping. In: Proceedings of the 6th International Conference on Document Analysis and Recognition (ICDAR '01); 2001. p. 39–43.

Rosen G, Williams A, Capra J, Connolly M, Cruz B, Lu L, Airey D, Kulkarni K, Williams R, 2000. The mouse brain library @ www.mbl.org. In: International Mouse Genome Conference, vol. 14. p. 166.

Roy S. Stereo without epipolar lines: A maximum-flow formulation. Int J Comput Vision 1999;34(2):147–61.

Rueckert D, Sonoda L, Hayes C, Hill D, Leach M, Hawkes D. Non-rigid registration using free-form deformations: application to breast MR images. IEEE Trans Med Imaging 1999;18(1):712–21.

Rydmark M, Jansson T, Berthold C, Gustavsson T. Computer-assisted realignment of light micrograph images from consecutive section series of cat cerebral cortex. J Microsc 1992;165:29–47.

Sakoe H, Chiba S. Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans Acoustics Speech Signal Process 1978;26(1):43–9.

Schormann T, Dabringhaus A, Zilles K. Statistics of deformations in histology and application to improved alignment with MRI. IEEE Trans Med Imaging 1995;14(1):25–35.

Sidman R, 2005. High resolution brain atlas of c57bl/6j in 2D and 3D. http://www.hms.harvard.edu/research/brain/atlas.html.

Streicher J, Donat M, Strauss B, Sprle R, Schughart K, Mller G. Computer-based three-dimensional visualization of developmental gene expression. Nat Genet 2000;25(2):147–52.

Studholme C, Constable RT, Duncan J. Accurate alignment of functional EPI data to anatomical MRI using a physics-based distortion model. IEEE Trans Med Imaging 2000;19(11):1115–27.

Thevenaz P, Ruttimann U, Unser M. A pyramid approach to sub-pixel registration based on intensity. IEEE Trans Med Imaging 1998;7(1):27–41.

Thompson P, Toga A. A surface-based technique for warping three dimensional images of the brain. IEEE Trans Med Imaging 1996;15(4):402–17.

Timsari B, Tocco G, Bouteiller J, Baudry M, Leahy R. Accurate registration of autoradiographic images of rat brain using a 3-d atlas. In: International Conference on Imaging Science, Systems, and Technology; 1999.

Toga AW. Brain Warping. Academic Press; 1999. ISBN 0126925356.

Toga A, Arnicar-Sulze T. Digital image reconstruction for the study of brain structure and function. J Neurosci Methods 1987;20(1):7–21.

Tomasi C, Manduchi R. Bilateral filtering for gray and color images. In: Proceedings of the IEEE International Conference on Computer Vision; 1998. p. 59–66.

Uchida S, Sakoe H. A monotonic and continuous two-dimensional warping based on dynamic programming. In: Proceedings of the 15th International Conference on Pattern Recognition, vol. 1; 1998. p. 521–4.

Wirtz S, Fischer B, Modersitzki J, Schmitt O. Super-fast elastic registration of histologic images of a whole rat brain for three-dimensional reconstruction. In: Proceedings of the SPIE Medical Imaging 2004, vol. 5730, February; 2004.

Woods R, Grafton S, Holmes C, Cherry S, Mazziotta J. Automated image registration: i. General methods and intrasubject, intramodal-ity validation. J Comput Assist Tomogr 1998a;22:139–52.

Woods R, Grafton S, Watson J, Sicotte N, Mazziotta J. Automated image registration: ii. Intersubject validation of linear and nonlinear models. J Comput Assist Tomogr 1998b;22:153–65.

Woods R, Mazziotta J, Cherry S. Mri-pet registration with automated algorithm. J Comput Assist Tomogr 1993;17(4):536–46.

Yoo T, Ackerman M, Lorensen W, Schroeder W, Chalana V, Aylward S, Metaxes D, Whitaker R. Engineering and algorithm design for an image processing api: a technical report on ITK—the insight toolkit. In: Proceedings of medicine meets virtual reality. Amsterdam: IOS Press; 2002. p. 586–92.