3D Volume Reconstruction by Serially Acquired 2D Slices Using a Distance Transform-Based Global Cost Function

Stelios Krinidis, Christophoros Nikou, and Ioannis Pitas

Aristotle University of Thessaloniki
Department of Informatics
Box 451, 54006 Thessaloniki, Greece

Abstract. An accurate, computationally efficient and fully-automated algorithm for the alignment of 2D serially acquired sections forming a 3D volume is presented. The method accounts for the main shortcomings of 3D image alignment: corrupted data (cuts and tears), dissimilarities or discontinuities between slices, missing slices. The approach relies on the optimization of a global energy function, based on the object shape, measuring the similarity between a slice and its neighborhood in the 3D volume. Slice similarity is computed using the distance transform measure in both directions. No particular direction is privileged in the method avoiding global offsets, biases in the estimation and error propagation. The method was evaluated on real images (medical, biological and other CT scanned 3D data) and the experimental results demonstrated the method's accuracy as reconstruction errors are less than 1 degree in rotation and less than 1 pixel in translation.

1 Introduction

Three-dimensional reconstruction of medical images (tissue sections, CT and autoradiographic slices) is now an integral part of biomedical research. Reconstruction of such data sets into 3D volumes, via the registrations of 2D sections, has gained an increasing interest. The registration of multiple slices is of utmost importance for the correct 3D visualization and morphometric analysis (e.g. surface and volume representation) of the structures of interest. Several alignment algorithms have been proposed in that framework. A review of general medical image registration methods is presented in [1], [2], [3].

The principal 3D alignment (reconstruction from 2D images) methods may be classified in the following categories: fiducial marker-based methods [4], feature-based methods using contours, crest lines or characteristic points extracted from the images [5], [6], and gray level-based registration techniques using the intensities of the whole image [7], [8], [9], [10]. Most of the above mentioned techniques do not simultaneously consider the two major difficulties involved in medical and CT scanned data registration.

At first, consecutive slices may differ significantly due to distortions, discontinuities in anatomical structures, cuts and tears. These effects are more pronounced when distant slices are involved in the registration. From this point of view, a registration method must be robust to missing data or outliers [7], [10].

Besides, registering the slices sequentially (the second with respect to the first, the third with respect to the second, etc.) leads to different types of misregistration. If an error occurs in the registration of a slice with respect to the preceding slice, this error will propagate through the whole volume. Also, if the number of slices to be registered is large, a global offset of the volume may be observed, due to error accumulation [8].

In this paper, a solution to the above mentioned shortcomings is presented. A global energy function having as variables the rigid transformation parameters (2D translation and rotation) of a given slice with respect to a local symmetric neighborhood is proposed. Global energy functions are a powerful tool in computer vision applications but they have not yet been considered for the registration of serially acquired slices.

Our approach was inspired by the technique presented in [11], which consists in minimizing a global energy function with the Iterative Closest Point algorithm [12], to register multiple, partially overlapping views of a 3D structure. The global energy function implemented in our approach is associated with a pixel similarity metric based on the Euclidean distance transform [13].

The remainder of the paper is organized as follows. The global energy function formulation and the associated registration algorithm is presented in section 2, experimental results are presented in section 3 and conclusions are drawn in section 4.

2   A Global Energy Function Formulation

Before presenting the alignment method, the notations used in our formulation are introduced. A set of 2D serially acquired slices is represented by:

$$V = \{I_i | i = 1 \ldots N\} \tag{1}$$

where I_i is a slice (a 2D image) and N denotes the total number of slices. A pixel of a 2D slice is represented by: $p = (x, y)^T$, so that $I_i(p)$ corresponds to the gray level (intensity) of pixel p of slice i. N_x and N_y designate the number of pixels of each slice in the horizontal and vertical direction respectively.

Standard two-dimensional rigid alignment consists of estimating the rigid transformation parameters (translation t_x, t_y and rotation by angle θ) that have to be applied to the image to be aligned (floating image) in order to match a reference image.

In the approach proposed here, the alignment of the 2D sections, within the 3D volume, is considered globally by minimizing an energy function $E(\cdot)$, which expresses the similarity between the 2D sections:

$$E(\Theta) = \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{p=1}^{N_x \times N_y} f(I_i(T_{\Theta_i}(p)), I_j(T_{\Theta_j}(p))) \tag{2}$$

where $f(\cdot)$ is a similarity metric, I_k denotes slice k and T_{Θ_k} designates a rigid transformation with parameters $\Theta_k = \{t_x^k, t_y^k, \theta^k\}$.

Equation (2) indicates that for a given set of rigid transformation parameters T_{Θ_i}, applied to the slice to be aligned I_i, the similarity between the *transformed slice* $I_i(T_{\Theta_i}(p))$ and *all of the other already transformed slices* $I_j(T_{\Theta_j}(p))$ in the volume is accumulated in the energy function.

Assuming that function $f(\cdot)$ is symmetric:

$$f(I_i(T_{\Theta_i}(p)), I_j(T_{\Theta_j}(p))) = f(I_j(T_{\Theta_j}(p)), I_i(T_{\Theta_i}(p))) \tag{3}$$

which is the case for the pixel similarity functions considered here, yields the following global minimization problem:

$$\hat{\Theta} = arg \min_{\Theta} E(\Theta) = arg \min_{\Theta} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j<i}}^{N} \sum_{p=1}^{N_x \times N_y} f(I_i(T_{\Theta_i}(p)), I_j(T_{\Theta_j}(p))) \tag{4}$$

Without additional constrains, the optimization problem (4) has clearly an infinite number of solutions (if the set of rigid transformations $\{T_{\hat{\Theta}_1}, T_{\hat{\Theta}_2}, \ldots T_{\hat{\Theta}_N}\}$ is a solution, the same holds true for $\{T_{\hat{\Theta}_1} \circ T_\Delta, T_{\hat{\Theta}_2} \circ T_\Delta, \ldots T_{\hat{\Theta}_N} \circ T_\Delta\}$, where T_Δ is an arbitrary 2D rigid transformation). To remove this ambiguity, the transformation $T_{\hat{\Theta}_l}$ applied to an arbitrary chosen slice k is constrained to the identity transformation (we have chosen $k = 1$ in our implementation). As a result, there are $3(N-1)$ parameters to estimate.

It is common sense that distant slices present very little similarity due to anatomy and it would be more appropriate to measure the energy function only for slices presenting at least some similarities. Therefore, the support region of function $f(\cdot)$ has been limited to a neighborhood of radius R centered at each slice and set:

$$f(I_i(T_{\Theta_i}(p)), I_j(T_{\Theta_j}(p))) = 0, \quad \forall \ |i - j| > R \tag{5}$$

Thus, the following alignment algorithm is associated with the energy function (4):

– **do** until convergence.
  • declare all slices unvisited.
  • **do** until all slices are declared visited.
    ∗ randomly chose an unvisited slice $I_i \in V$.

∗ update the rigid transformation parameters T_{Θ_i} bringing into alignment slice I_i with the other slices in the neighborhood of i, by minimization of the following local energy function:

$$E_i(\Theta_i) \overset{\text{def}}{=} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ |i-j| \leq R}}^{N} \sum_{p=1}^{N_x \times N_y} f(I_i(T_{\Theta_i}(p)), I_j(T_{\Theta_j}(p))) \qquad (6)$$

∗ declare slice I_i visited.
- **end do**
− **end do**

The minimization of the local energy function (4) is conducted by a deterministic optimization algorithm, known as Iterated Conditional Modes (ICM) [14]. ICM is a discrete Gauss Seidel-like optimization technique, accepting only configurations decreasing the objective function. Let us notice that the parameter $\hat{\Theta}_i$ corresponding to the minimum value of the local energy function $E_i(\Theta_i)$ (Equ. 6) also corresponds to a local minimum value of the global energy function $E(\Theta)$ with respect to Θ_i, keeping all other parameters $\Theta_j, j \neq i$ fixed. It is thus easy to see that the described algorithm converges towards a local minimum of the initial energy function (2). This local minimum corresponds to a satisfactory registration, since the initial alignment of the 2D sections is generally close to the desired solution (if this is not the case, a good initialization may be obtained by a standard coarse alignment technique such as principal axes registration). It is thus not necessary to resort here to greedy global optimization procedures, such as simulated annealing or genetic algorithms.

Further improvement of the solution is obtained by a gradient decent technique. To speed the algorithm up a multigrid data processing is also implemented.

The pixel similarity metric associated with the above described global energy function is based on a distance transform ([13], [15]) (also known as chamfer matching technique [16]) and it is computed from the 3D object contours [17]. A distance transformation is an operation that converts a binary picture, consisting of feature and non-feature elements (contours), to a picture where each pixel has a value that approximates its distance to the nearest contour point.

Thus, using the distance transform $D(p)$ of the reference slice the method aligns the floating slice by minimizing the distance between the contours of the images. For further details of the chamfer matching method the reader may refer to [16].

Considering the slices per triplets, which is very common for standard reconstruction problems (i.e. setting $R=1$ in eq. 5), the estimation of the alignment parameters Θ involves the non-linear similarity metric:

$$f(T_{\Theta_i}(p)) = D_{i-1}(T_{\Theta_{i-1}}(p)) + D_{i+1}(T_{\Theta_{i+1}}(p)), \qquad I_i(T_{\Theta_i}(p)) \neq 0 \qquad (7)$$

where $I_i(T_{\Theta_i}(p)) \neq 0$ means that only the contour points of I_i are involved.

A large number of interpolations are involved in the alignment process. The accuracy of estimation of the rotation and translation parameters is directly related to the accuracy of the underlying interpolation model. Simple approaches such as the nearest neighbor interpolation are commonly used because they are fast and simple to implement, though they produce images with noticeable artifacts. Besides, as the translation and rotation parameters should compensate for accuracy by having subvoxel values, this type of interpolation would not be appropriate. More satisfactory results can be obtained by small-kernel cubic convolution techniques, bilinear, or convolution-based interpolation. According to sampling theory, optimal results are obtained using sinus cardinal interpolation, but at the expense of a high computational cost. As a compromise, a bilinear interpolation technique has been used in the optimization steps. At the end of the algorithm, the alignment parameters are refined using a sinus cardial interpolation that preserves the quality of the image to be aligned. This technique has proven to be fast and efficient.
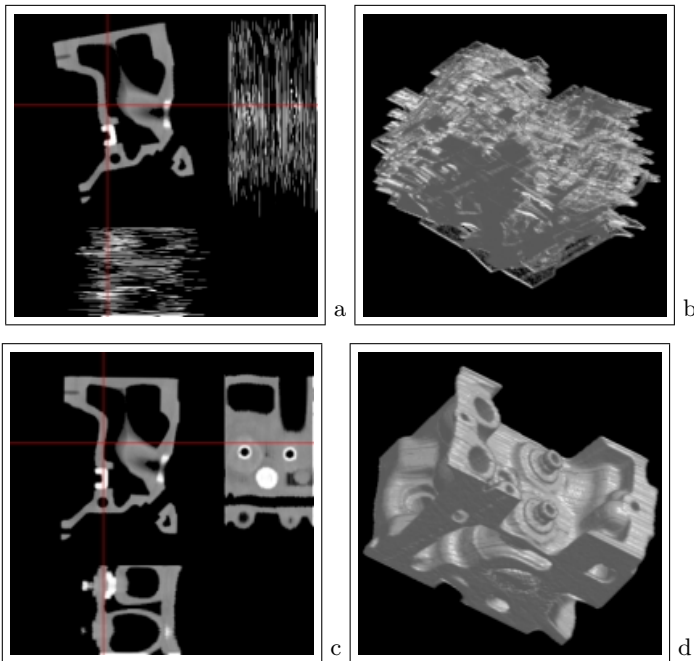

Fig. 1. Reconstruction of a 3D scanned mechanical part volume of 109 slices. **(a)** Multiplanar view of the volume before registration. **(b)** Three-dimensional view of the volume before registration. **(c)** Multiplanar view of the volume after registration. **(d)** Three-dimensional view of the volume after registration.

Table 1. A set of 109 slices of a 3D CT scanned mechanical part volume were artificially transformed using different rigid transformation parameters. Each slice was randomly transformed using translations varying from -10 to +10 pixels and rotations varying from -20 to +20 degrees. Statistics on the alignment errors for the rigid transformation parameters are presented. Translation errors are expressed in pixels and rotation error in degrees.

	Δt_x	Δt_y	$\Delta\theta$
median	0.33	0.31	0.06
maximum	1.07	0.93	0.25
mean $\pm$ s. dev	0.35 $\pm$ 0.25	0.38 $\pm$ 0.25	0.07 $\pm$ 0.06

3   Experimental Results

To evaluate our method, we applied the algorithm to the reconstruction of an artificially misaligned 3D CT scanned mechanical part (figure 1). The slices of the original $256 \times 256 \times 109$ CT volume were transformed using translations varying from -10 to +10 pixels and rotations varying from -20 to +20 degrees.

Table 2. A set of 100 slices of a 3D CT scanned mechanical part volume were artificially transformed using different rigid transformation parameters. Each slice was translated by 0.2 pixels in both directions and rotated by 0.4 degrees with respect to its preceding slice. Different statistics on the errors for the rigid transformation parameters are presented. Translation errors are expressed in pixels and rotation error in degrees.

	Δt_x	Δt_y	$\Delta\theta$
median	0.19	0.23	0.05
maximum	0.99	0.87	0.32
mean $\pm$ s. dev	0.29 $\pm$ 0.26	0.31 $\pm$ 0.26	0.07 $\pm$ 0.07

The transformations for each slice were random following a uniform distribution in order not to privilege any slice (figures 1(a) and 1(b)). Table 1 presents statistics on the alignment errors. The algorithm revealed robust in aligning this type of image providing small registration errors. Figures 1(c) and 1(d) present the reconstructed volume.

Moreover, we have uniformly transformed 100 slices of the same 3D volume (mechanical part of an engine) by applying to each slice I_i a translation of $t_x^i = t_x^{i-1} + 0.2$ pixels and $t_y^i = t_y^{i-1} + 0.2$ pixels and a rotation of $\theta^i = \theta^{i-1} + 0.4$ degrees. As the volume has 100 slices, the last slice is translated by 20 pixels in both directions and rotated by 40 degrees with respect to its initial position. Table 2 presents the registration errors of the method. It is illustrated that our approach has subvoxel mean, median and maximum errors.
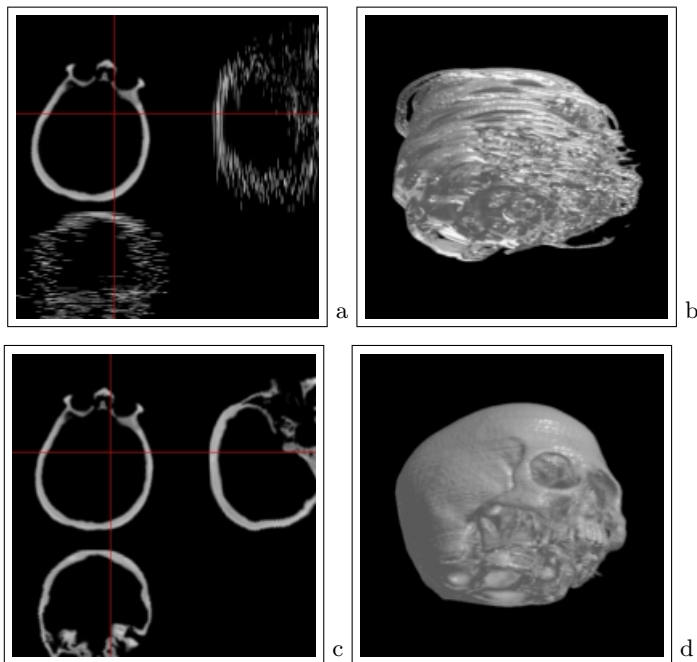
Fig. 2. Reconstruction of a 3D human skull volume of 140 slices. **(a)** Multiplanar view of the volume before registration. **(b)** Three-dimensional view of the volume before registration. **(c)** Multiplanar view of the volume after registration. **(d)** Three-dimensional view of the volume after registration.

Table 3. A set of 140 slices of a 3D CT human skull volume were artificially transformed using different rigid transformation parameters. Each slice was randomly transformed using translations varying from -10 to +10 pixels and rotations varying from -20 to +20 degrees. Different statistics on the errors for the rigid transformation parameters are presented. Translation errors are expressed in pixels and rotation error in degrees.

	Δt_x	Δt_y	$\Delta \theta$
median	2.10	0.33	0.07
maximum	1.45	2.02	2.42
mean ± s. dev	0.37 ± 0.28	0.38 ± 0.30	0.19 ± 0.35

The same evaluation procedure was performed on a 3D human skull volume with 140 slices (figure 2). The algorithm aligned the artificially (randomly and uniformly) misaligned slices of the volume and the errors are drawn in Tables 3 and 4. Human skull presents discontinuities, and consecutive slices may differ significantly due to anatomy but the global energy function is robust to these shortcomings. As it can be seen, median and mean translation and rotation errors are less than 1 pixel and 1 degree respectively. Also maximum errors are
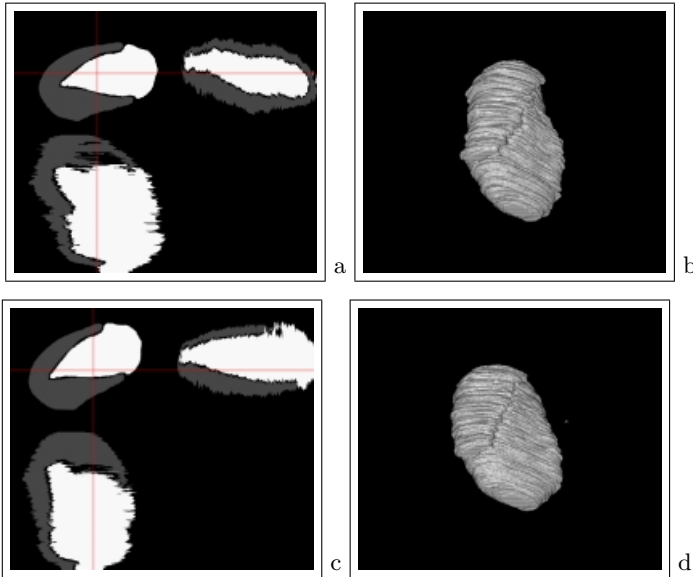
Fig. 3. Reconstruction of a 3D tooth volume of 265 slices. **(a)** Multiplanar view of the volume after alignment by an expert dentist. **(b)** Three-dimensional view of the volume after alignment by an expert dentist. **(c)** Multiplanar view of the volume after registration. **(d)** Three-dimensional view of the volume after registration.

Table 4. A set of 140 slices of a 3D CT human skull volume were artificially transformed using different rigid transformation parameters. Each slice was translated by 0.15 pixels in both directions and rotated by 0.3 degrees with respect to its preceding slice. Different statistics on the errors for the rigid transformation parameters are presented. Translation errors are expressed in pixels and rotation error in degrees.

<div align="center">

Alignment error statistics

	Δt_x	Δt_y	$\Delta \theta$
median	0.23	0.21	0.26
maximum	1.95	1.94	1.64
mean $\pm$ s. dev	0.33 ± 0.32	0.34 ± 0.33	0.25 ± 0.25

</div>

slightly superior to 1 pixel and 1 degree respectively showing the robustness of the technique.

Furthermore, the algorithm was applied to the reconstruction of volumes (tooth germs, biological tissues) with unknown ground truth. The method's performance was compared with the manual alignment accomplished by an expert physician. Figure 3 shows the reconstruction of a tooth germ by an expert dentist (fig. 3(a) and 3(b)) and by our method (fig. 3(c) and 3(d)). It is illustrated that human intervention fails to correctly align the slices, whilst our method is efficient and can achieve alignment with high accuracy.
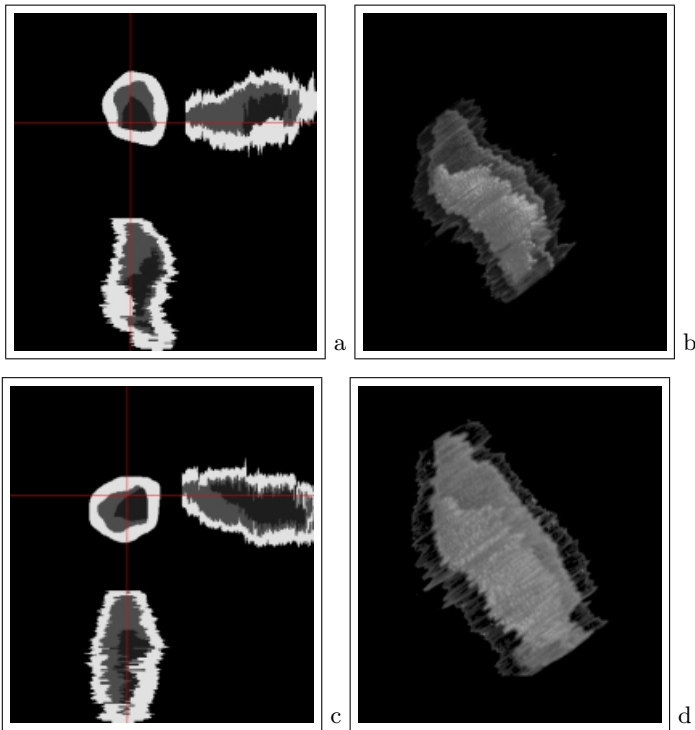
Fig. 4. Reconstruction of a 3D tooth volume of 194 slices. **(a)** Multiplanar view of the volume after alignment by an expert dentist. **(b)** Three-dimensional view of the volume after alignment by an expert dentist. **(c)** Multiplanar view of the volume after registration. **(d)** Three-dimensional view of the volume after registration.

The same stands for the example presented in figure 4 where another tooth reconstruction is presented.

Also, Figure 5 depicts a 3D tissue containing a large number of vessels. Figures 5(a) and 5(b) show the volume aligned by an expert biologist and Figures 5(c) and 5(d) the tissue after alignment by our method.

This volume presents cuts and discontinuities and the tissues had been stretched during the cut procedure. Despite these drawbacks, according to the expert biologist, the algorithm aligned correctly the slices.

Finally, let us notice that the algorithm has a computational complexity $O(N_x N_y N)$ and requires approximately 10 min. to reconstruct a $256 \times 256 \times 140$ volume on a Pentium III (800 MHz) workstation.

4   Conclusion

The alignment method described in this paper is akin to the global energy function formulation proposed in [11] to register multiple views of a 3D surface in
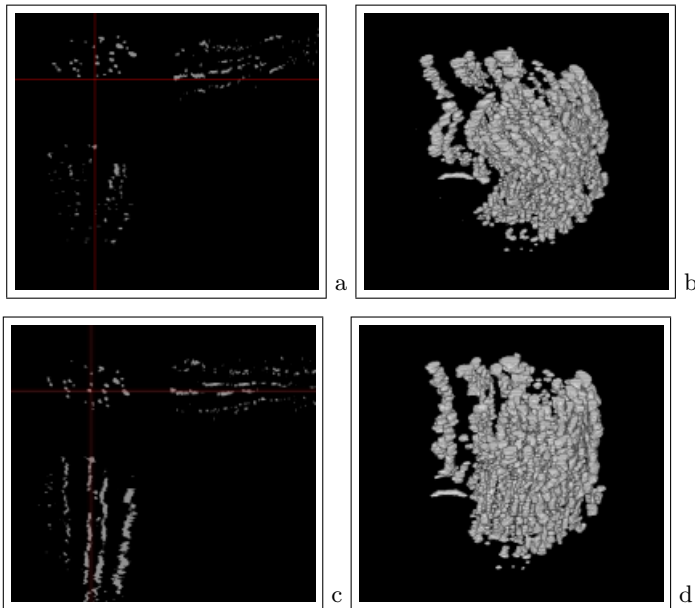
Fig. 5. Reconstruction of a 3D tissue volume of 237 slices. **(a)** Multiplanar view of the volume after alignment by an expert biologist. **(b)** Three-dimensional view of the volume after alignment by an expert biologist. **(c)** Multiplanar view of the volume after registration. **(d)** Three-dimensional view of the volume after registration.

computer vision applications. The main contribution of the approach is to consider the alignment problem globally on the 3D volume, by minimizing a global objective function expressing the similarity between neighboring slices. The approach does not privilege any particular direction in the registration process. By these means, the major problems set by the registration of serially acquired slices are addressed. With the global (isotropic) formulation of the registration problem (rather than a standard step by step, sequential formulation), no global offset nor error propagations are observed in the final alignment. The approach seems promising and its association to more sophisticated but time consuming pixel similarity metrics (mutual information [18], robust estimation-based measures [19]) may improve its accuracy and is a perspective of this work.

References

1. L. Gottesfeld-Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.
2. J. B. A. Maintz and M. A. Viergever. A survey of medical image registration techniques. *Medical Image Analysis*, 2(1):1–36, 1998.
3. P. Van den Elsen, E. J. D. Paul, and M. A. Viergever. Medical image matching - a review with classification. *IEEE engineering in Medicine and Biology*, 12(1):26–39, 1993.

4. A. F. Goldszal, O. J. Tretiak, P. J. Hand, S. Bhasin, and D. L. Mac Eachron. Three-dimensional reconstruction of activated columns from 2-[^{14}c] deoxy-d-glucose data. *NeuroImage*, 2:9–20, 1995.

5. L. Hibbard and R. Hawkins. Objective image alignment for three-dimensional reconstruction of digital autoradiographs. *Journal of Neuroscience Methods*, 26:55–75, 1988.

6. A. Rangarajan, H. Chui, E. Mjolsness, S. Pappu, L. Davachi, P. Goldman-Rakic, and J. Duncan. A robust point-matching algorithm for autoradiograph alignment. *Medical Image Analysis*, 1(4):379–398, 1997.

7. W. Zhao, T. Young, and M. Ginsberg. Registration and three-dimensional reconstruction of autorediographic images by the disparity analysis method. *IEEE Transactions on Medical Imaging*, 12(4):782–791, 1993.

8. A. Andreasen, A. M. Drewes, J.E. Assentoft, and N. E. Larsen. Computer-assisted alignment of standard serial sections without use of artificial landmarks. a practical approach to the utilization of incomplete information of 3d reconstruction of the hippocampal region. *Journal of Neuroscience Methods*, 45:199–207, 1992.

9. B. Kim, J. Boes, K. Frey, and C. Meyer. Mutual information for automated unwarping of rat brain autorediographs. *NeuroImage*, 5:31–40, 1997.

10. S. Ourselin, A. Roche, G. Subsol, X. Pennec, and C. Sattonnet. Automatic alignment of histological sections for 3d reconstruction and analysis. Sophia Anipolis, France, 1998.

11. R. Ben-Jemaa and F. Schmitt. A solution for the registration of multiple 3d points sets using unit quaternions. In *Notes in Computer science. Proceedings of the 5th European Conference on Computer Vision (ECCV'98)*, volume 2, pages 34–50, Freiburg, Germany, June 1998.

12. M. J. Besl and N. McKay. A method for the registration of 3d shapes. *IEEE transactions of Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

13. G. Borgefors. Distance transformations in arbitrary dimensions. *Computer Vision, Graphics, and Image Processing*, 27:321–345, 1984.

14. J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, 48(3):259–302, 1986.

15. Per-Erik Danielsson. Euclidean distance transform. *Computer Graphics and Image Processing*, 14:227–28, 1980.

16. G. Borgefors. Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE transactions of Pattern Analysis and Machine Intelligence*, 10:849–965, November 1988.

17. J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 679–698, 1986.

18. W. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multimodal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1):33–51, 1996.

19. C. Nikou, J. P. Armspach, F. Heitz, I. J. Namer, and D. Grucker. Mr/mr and mr/spect registration of brain images by fast stochastic optimization of robust voxel similarity measures. *NeuroImage*, 8(1):30–43, 1998.