Multimodal Image Registration using Statistically Constrained Deformable Multimodels

C. Nikou* † F. Heitz* J. P. Armspach†

*Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection
CNRS UPRES-A 7005 / Strasbourg I University
4 boulevard Sébastien Brant, 67400 Illkirch, France
nikou@mondrian.u-strasbg.fr, Fabrice.Heitz@ensps.u-strasbg.fr

†Institut de Physique Biologique, Faculté de Médecine
CNRS UPRES-A 7004 / Strasbourg I University
4 rue Kirschleger, 67085 Strasbourg, France
armspach@alsace.u-strasbg.fr

Abstract

The registration of multimodal images remains an intricate issue, especially when the multimodal image pair shows non overlapping structures, missing data, noise or outliers. In this paper, we present a deformable model-based technique for the rigid registration of 2D and 3D multimodal images. The deformable model embeds a priori knowledge of the spatial correspondence and statistical variability of the different (eventually non overlapping) image features which are used in the registration procedure. The method is applied to the intrasubject registration of medical MR/SPECT images of the brain by constructing a deformable model incorporating information on both MR (head) and SPECT (brain) contours. In an off-line training procedure, the spatial relations between head and brain as well as the anatomical variations of these two structures are learned. The registration is performed by minimizing an objective function involving the contours of the MR and SPECT images.

1 Introduction

The goal of image registration is to geometrically align two or more images so that pixels (or voxels) representing the same underlying structure may be superimposed. Image registration is an important preliminary step in many application fields involving, for instance, the detection of changes in temporal image sequences or the fusion of multimodal images [1, 2]. Medical imaging, with its wide variety of sensors (thermal, ultrasonic, X-Ray, MRI and nuclear) is probably one of the first application field [2], as are remote sensing, military imaging (visible, IR or radar), multisensor robot vision and multisource imaging used in the preservation of artistic patrimony.

Although a large variety of image registration methods have been proposed in the literature [1, 2], only a few techniques have attempted to address the registration of multimodal images showing gross dissimilarities, non overlapping structures, missing data, noise and outliers. The registration problem is indeed particularly difficult for multimodal images, showing an "overall" difference (due to differences in the characteristics of the scene observed by multiple sensors) and for which it is sometimes difficult to find image features that match exactly in the different modalities. This is the case when the multimodal pair is used for the complementary and non redundant information one image provides with respect to the other. When there are no overlapping structures in the image pair, the registration problem becomes involved. Image models, representing the relative location of the image features used in the different modalities, have to be introduced.

Since the seminal work of Kass *et al.* [3] on 2D shape models, deformable models have gained increasing popularity in computer vision to segment, match or track rigid and nonrigid objects [4, 5, 6, 7]. In medical image analysis, deformable models offer a unique and powerful approach to accommodate the significant variability of biological structures over time and across different individuals. A survey on deformable models as a promising computer-assisted medical image analysis technique may be found in [8]. A recent survey of medical image registration is presented in [2].

Among the different registration methods that can be identified, *segmentation-based registration methods* rely on rigid or deformable models of the structures (usually regions, surfaces or curves) used in the alignment procedure (see [2]). Statistical deformable models [9, 10, 7] in particular, are good candidate for representing the spatial relations and the statistical variability of the image features which are used as a support for the registration. In [11], for instance, a multislice 2D point distribution model (PDM) [9] is deformed to match various structures in single modal PET images. The approach proposed here relies on a 3D physical deformable model that embeds information on the structures that may be extracted from different image modalities. The different model parts are statistically constrained to represent the structures of interest and their spatial relations in the different image modalities. These constraints are learned from a representative population in an off-line training procedure.

The approach is applied to the registration of 2D and 3D Magnetic Resonance (MR) and Single Photon Emission Computed Tomography (SPECT) images by modeling the head contour in the MR images and the brain contour in the SPECT volumes. These structures are easily extracted from these image modalities (fig. 2) but they do not overlap in the registered image pair. The learning of the spatial relation between head and brain (using a representative training set) thus provides a robust solution to the registration problem, although the images in the pair show different structures (see fig. 2). This approach is also a first step towards the creation of an anatomical atlas that may be used both for segmentation, intrasubject and intersubject registration [2, 12].

2 A statistical deformable multimodel

To provide a training set, a representative collection of 3D MR/SPECT image pairs of different patients have first been rigidly registered to a reference MR image. To this end a supervised segmentation technique has been used, to remove non brain structures (skull and scalp), which are visible in the MR images, but have no counterpart in the SPECT images. The registration is performed on these pre-processed images, using an unsupervised robust multimodal registration technique developed by the authors [13, 14].

The head and the brain have then been segmented from the MR and the SPECT volumes respectively and have been parameterized by the amplitudes of the vibration modes of a deformable spherical mesh [15], by solving the system evolution equation:

$$\mathbf{M}\ddot{\mathbf{U}}(t) + \mathbf{C}\dot{\mathbf{U}}(t) + \mathbf{K}\mathbf{U}(t) = \mathbf{F}(t) \tag{1}$$

where $\mathbf{M}$, $\mathbf{C}$, $\mathbf{K}$ are the mass, dumping and stiffness matrices of the system, $\mathbf{U}$ stands for the nodal displacements and $\mathbf{F}$ is the image force [16]. Equation (1) is of order $3N$, where N is the total number of nodes of the spherical mesh. It is solved in the subspace of the vibration modes [15] using the following change of basis vectors:

$$\mathbf{U} = \mathbf{\Phi}^T\tilde{\mathbf{U}} = \sum_i \tilde{u}_i \phi_i, \tag{2}$$

where $\mathbf{\Phi}$ is a matrix and $\tilde{\mathbf{U}}$ is a vector, ϕ_i is the i^{th} row of $\mathbf{\Phi}$ and $\tilde{u}$ is a scalar. By choosing $\mathbf{\Phi}$ as the matrix whose columns are the eigenvectors of the eigenproblem:

$$\mathbf{K}\phi_i = \omega_i^2 \mathbf{M}\phi_i, \tag{3}$$

$\mathbf{M}$ and $\mathbf{K}$ are simultaneously diagonalized and the system (1) is simplified to $3N$ scalar equations [15]:

$$\ddot{\tilde{u}}_i(t) + \tilde{c}_i\dot{\tilde{u}}_i(t) + \omega_i^2\tilde{u}_i(t) = \tilde{f}_i(t). \tag{4}$$

In equation (4), ω_i designates the i^{th} eigenvalue of equation (3), $\tilde{u}_i$ is the amplitude of the corresponding vibration mode, $\tilde{c}_i$ are the nonzero (diagonal) elements of $\tilde{\mathbf{C}} = \mathbf{\Phi}^T\mathbf{C}\mathbf{\Phi}$ and $\tilde{f}_i(t) = \mathbf{\Phi}^T\mathbf{F}(t)$.

For each image pair i in the training set, a vector $\mathbf{a}_i$ containing the N' most important vibration modes describing the head in the MR image ($\mathbf{u}_i^{MR}$) and the N'' most important vibration modes describing the brain in the SPECT image ($\mathbf{u}_i^{SPECT}$) is then created:

$$\mathbf{a}_i = (\mathbf{u}_i^{MR}, \mathbf{u}_i^{SPECT})^T \tag{5}$$

where:

$$\begin{aligned} \mathbf{u}_i^{MR} &= (\tilde{u}_1^{MR}, \tilde{u}_2^{MR}, \ldots, \tilde{u}_{N'}^{MR})_i & (6)\\ \mathbf{u}_i^{SPECT} &= (\tilde{u}_1^{SPECT}, \tilde{u}_2^{SPECT}, \ldots, \tilde{u}_{N''}^{SPECT})_i & (7) \end{aligned}$$

with $3(N' + N'') < 6N$.

Random vector $\mathbf{a}_i$ is then statistically constrained by retaining the most significant variation modes in its Karhunen-Loeve (KL) transform [9, 7, 17]:

$$\mathbf{a}_i = \bar{\mathbf{a}} + \mathbf{P}\mathbf{b}_i \tag{8}$$

where $\bar{\mathbf{a}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{a}_i$ is the average vector of the training set, $\mathbf{P}$ is the matrix whose columns are the eigenvectors of the covariance matrix

$$\mathbf{\Gamma} = \mathbb{E}[(\mathbf{a}_i - \bar{\mathbf{a}})^T(\mathbf{a}_i - \bar{\mathbf{a}})] \tag{9}$$

and

$$\mathbf{b}_i = \mathbf{P}^T(\mathbf{a}_i - \bar{\mathbf{a}}) \quad (10)$$

are the coordinates of $(\mathbf{a}_i - \bar{\mathbf{a}})$ in the eigenvector basis.

The deformable model (corresponding to the head and brain contours here) is finally parameterized by the m most significant statistical deformation modes stacked in vector $\mathbf{b}$. By modifying $\mathbf{b}$, both the head and the brain are deformed (fig. 1), according to the anatomical variability observed in the training set. Given a double initial spherical mesh $\bar{\mathbf{X}}_{H,B}$, the deformable multimodel $\mathbf{X}_{H,B}(\mathbf{b})$ is thus represented by:

$$\mathbf{X}_{H,B}(\mathbf{b}) = \bar{\mathbf{X}}_{H,B} + \mathbf{\Phi}\bar{\mathbf{a}} + \mathbf{\Phi}\mathbf{P}\mathbf{b} \quad (11)$$

with a low dimension $m \ll 3(N' + N'')$ for $\mathbf{b}$. In our application, typical values are $N \simeq 20000$, $N' \simeq N'' \simeq \frac{N}{4} \simeq 5000$ and $m \simeq 10$. Thanks to the KL representation, only a few parameters are necessary to describe the variations of the deformable multimodel.

3 Deformable model-based registration

Registration of the multimodal image pair consists in estimating the rigid transformation parameters $\mathcal{S}_{rig}$ (3D rotation and translation parameters) that have to be applied to the image to be registered (here the SPECT image) in order to match the reference image (here the MRI). The registration relies on (noisy) head contours extracted from the MRI and (noisy) brain contours extracted from the SPECT image. These structures do not overlap but the deformable model represents the relative location of the head and brain contours and accounts for the anatomical variability observed among the training population. The deformable model is used as a probabilistic atlas that constrains the rigid registration of the image pair. The transformations between the deformable model and the image pair include the deformation parameter vector $\mathbf{b}$ (representing anatomical variability) and a rigid transformation $\mathcal{S}_{mod}$.

The rigid and deformable transformation parameters $\mathcal{S}_{rig}$, $\mathcal{S}_{mod}$ and $\mathbf{b}$ yielding the overall "best" match of the deformable multimodel with the head contour in the MR image and the registered brain contour in the SPECT image are estimated by minimizing a global energy function based on a distance between the model and the contours extracted from the image pair [16]:

$$(\mathcal{S}_{rig}, \mathcal{S}^*_{mod}, \mathbf{b}^*) = \arg \min_{\mathcal{S}_{rig}, \mathcal{S}_{mod}, \mathbf{b}} [\mathcal{E}(\mathcal{S}_{rig}, \mathcal{S}_{mod}, \mathbf{b})] \quad (12)$$

where:

$$\begin{aligned} \mathcal{E}(.) &= E_{MR}[\mathbf{X}_H(\mathcal{S}_{mod}, \mathbf{b})] \\ &+ E_{SPECT}[\mathbf{X}_B(\mathcal{S}_{rig}, \mathcal{S}_{mod}, \mathbf{b})]. \end{aligned} \quad (13)$$

E_{MR} is an energy function computed only for the points of $\mathbf{X}_{H,B}$ modeling the head and E_{SPECT} depends only on the points of $\mathbf{X}_{H,B}$ modeling the brain :

$$\begin{aligned} E_{MR}[\mathbf{X}_H(\mathcal{S}_{mod}, \mathbf{b})] &= \sum_{p \in \mathbf{X}_H} \Delta_{MR}(p) \\ E_{SPECT}[\mathbf{X}_B(\mathcal{S}_{rig}, \mathcal{S}_{mod}, \mathbf{b})] &= \sum_{p \in \mathbf{X}_B} \Delta_{SPECT}(p). \end{aligned}$$

In the above equation, $\Delta_{MR}(p)$ and $\Delta_{SPECT}(p)$ designate the chamfer distance between point p of the deformable model and the nearest contour point in the MR and SPECT image respectively [16] (contours are extracted using simple thresholding techniques [14]).

The global minimization (12) may be performed by using a stochastic method, alternately minimizing the objective function $\mathcal{E}$ with respect to the different transformation parameters ($\mathcal{S}_{rig}$,$\mathcal{S}_{mod}$,$\mathbf{b}$) [6]. In practice, for the application considered here, we have resorted to the following suboptimal (but fast) estimation technique :

- Initialize the template by the average model $\mathbf{X}_{H,B}(\mathbf{0})$.
- Do until convergence:
  - Estimate $\mathcal{S}_{mod}$ using a principal axes registration technique, bringing into alignment the deformable model and the MRI.
  - Deform the average model by iteratively computing the components of $\mathbf{b}$ minimizing (13).
- Estimate $\mathcal{S}_{rig}$ using a principal axes registration technique, by bringing into alignment the SPECT image with the deformable model.

Figure 2 illustrates an example of a 2D (single slice) MRI/SPECT registration using the proposed technique (experiments of the method on 3D images are in progress). As can be seen, although the MRI and SPECT head and brain contours do not overlap, the two images have been correctly registered. It was also not necessary to remove non brain structures (skull and scalp) before registration.

4 Conclusion

We have presented a statistical deformable model-based technique for the registration of non overlapping multimodal data. Thanks to the statistical constraints embedded in the deformable model, the registration of multimodal images showing gross dissimilarities, non overlapping structures, missing data, noise and outliers becomes possible.

The technique is applied to the registration of MR and SPECT images, showing partially non redundant (anatomical and functional) information. The above method is not limited to MR/SPECT modalities or medical images but may be adapted to other image modalities used in remote sensing, military imaging, multisensor robot vision or industrial applications.

The described approach is also a first step towards the creation of an anatomical atlas that may be used for brain segmentation, intrasubject and intersubject registration. A first application concerning the segmentation of the brain in MR images is described in [12].

Acknowledgments

This study has been supported by the Commission of the European Communities, DG XII, in the framework of the TMR program (Training and Mobility of Researchers), contract No ERBFMIBCT960701 and by the "Groupement d'Intérêt Scientifique-Sciences de la Cognition" (CNRS, CEA, INRIA, MENESR).

References

[1] L. G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, 1992.

[2] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration techniques," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[3] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *International Journal of Computer Vision*, vol. 4, no. 1, pp. 321–331, 1988.

[4] R. Bajcsy and S. Kovačič, "Multiresolution elastic matching," *Computer Vision, Graphics and Image Processing*, vol. 46, pp. 1–21, 1989.

[5] C. Davatzikos, "Spatial transformation and registration of brain images using elastically deformable models," *Computer Vision and Image Understanding*, vol. 66, no. 2, pp. 207–222, 1997.

[6] C. Kervrann and F. Heitz, "Statistical model-based segmentation of deformable motion," in *Proceedings of the 3rd IEEE International Conference on Image Processing (ICIP '96)*, Lausanne, Switzerland, 1996, pp. 937–940.

[7] C. Kervrann and F. Heitz, "A hierarchical Markov modeling approach for the segmentation and tracking of deformable shapes," *Graphical Models and Image Processing*, accepted for publication, Feb. 1998.

[8] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: a survey," *Medical Image Analysis*, vol. 2, no. 1, pp. 91–108, 1996.

[9] T. F. Cootes, C. J. Taylor, and J. Graham, "Active shape models - their training and application," *Computer Vision and Image Understanding*, vol. 1, no. 1, pp. 38–59, 1995.

[10] U. Grenander and M. I. Miller, "Representation of knowledge in complex systems," *Journal of the Royal Statistical Society*, vol. 56, no. 3, pp. 549–603, 1994.

[11] C. L. Huang, W. T. Chang, L. C. Wu, and J. K. Wang, "Three-dimensional PET emission scan registration and transmission scan synthesis," *IEEE Transactions on Medical Imaging*, vol. 16, no. 5, pp. 542–561, 1997.

[12] C. Nikou, F. Heitz, and J. P. Armspach, "Brain segmentation from 3D MRI using statistically learned physics-based deformable models," *Submitted to IEEE Medical Imaging Conference*, 1998, 8-14 November 1998, Toronto, Canada.

[13] C. Nikou, F. Heitz, and J. P. Armspach, "Robust registration of dissimilar single and multimodal images," in *Lecture Notes in Computer Science. Proceedings of the 5^{th} European Conference on Computer Vision (ECCV'98)*, Springer-Verlag, Ed., Freiburg, Germany, 2-6 June 1998.

[14] 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*, 1998, In press.

[15] C. Nastar and N. Ayache, "Frequency-based nonrigid motion analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 11, pp. 1069–1079, 1996.

[16] G. Borgefors, "On digital distance transforms in three dimensions," *Computer Vision and Image Understanding*, vol. 64, no. 3, pp. 368–376, 1996.

[17] C. Nastar, B. Moghaddam, and A. Pentland, "Flexible images: matching and recognition using learned deformations," *Computer Vision and Image Understanding*, vol. 65, no. 2, pp. 179–191, 1997.

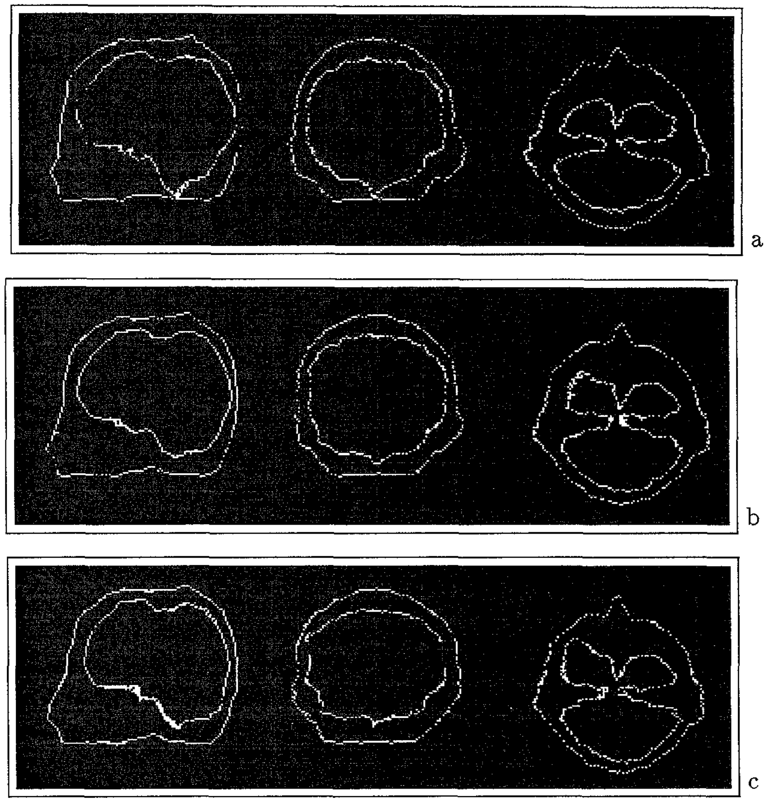

Figure 1: *Deformations of a 3D multimodel by varying the first statistical mode in vector* $\mathbf{b}$ *between* $-\sqrt{\lambda_1}$ *and* $\sqrt{\lambda_1}$. λ_1 *designates the first eigenvalue of the covariance matrix* $\mathbf{\Gamma}$. **(a)** *Configuration for* $\mathbf{b}[1] = -\sqrt{\lambda_1}$. **(b)** *Configuration for* $\mathbf{b}[1] = 0$, *i.e. the average model.* **(c)** *Configuration for* $\mathbf{b}[1] = 3\sqrt{\lambda_1}$. *Each row shows a multiplanar (sagital, coronal, transversal) view of the 3D model.*

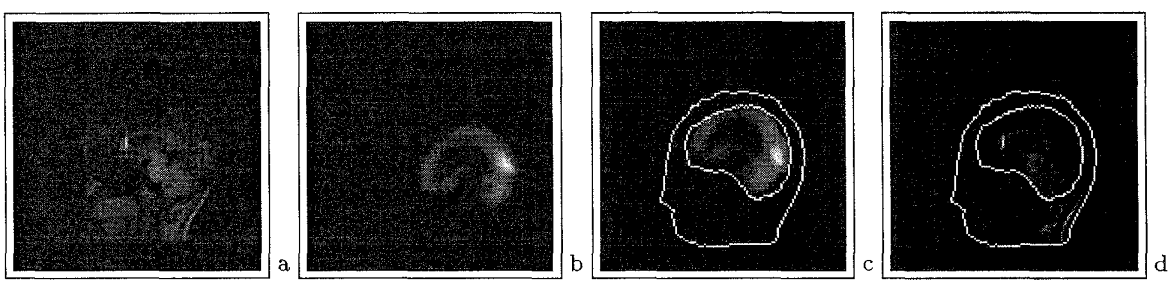

Figure 2: *2D MR/SPECT registration using the deformable multimodel.* **(a)** *MR image.* **(b)** *SPECT image.* **(c)** *Registered SPECT image with the deformable multimodel superimposed.* **(d)** *The MR image in (a) with the resulting deformable multimodel superimposed.*