

A Novel Approach for Global Lung Registration Using 3D Markov-Gibbs Appearance Model

Ayman El-Baz¹, Fahmi Khalifa¹, Ahmed Elnakib¹, Matthew Nitzken¹, Ahmed Soliman¹, Patrick McClure¹, Mohamed Abou El-Ghar², and Georgy Gimel'farb³

¹BioImaging Laboratory, Bioengineering Department, University of Louisville, Louisville, KY, USA.

²Radiology Department, Urology and Nephrology Center, University of Mansoura, Mansoura 35516, Egypt.

³Department of Computer Science, University of Auckland, Auckland, New Zealand.

Abstract. A new approach to align 3D CT data of a segmented lung object with a given prototype (reference lung object) using an affine transformation is proposed. Visual appearance of the lung from CT images, after equalizing their signals, is modeled with a new 3D Markov-Gibbs random field (MGRF) with pairwise interaction model. Similarity to the prototype is measured by a Gibbs energy of signal co-occurrences in a characteristic subset of voxel pairs derived automatically from the prototype. An object is aligned by an affine transformation maximizing the similarity by using an automatic initialization followed by a gradient search. Experiments confirm that our approach aligns complex objects better than popular conventional algorithms.

1 Introduction

Image registration aligns two or more images of similar objects taken at different times, from different viewpoints, and/or by different sensors. The images are geometrically transformed to ensure their close similarity. Registration is a crucial step in many applied image analysis tasks, e.g., early diagnosis of detected lung nodules. One of the most compelling motivations for identifying a potentially malignant nodule is to assess its growth rate. To quantify the growth rate of a nodule, one must be able to measure the volume of nodules and identify corresponding nodules in a follow-up scans. The principal difficulty in estimating the nodule growth rate is automatic identification and registration (alignment) of the corresponding nodules in follow-up scans. Registration of the lung tissues is a challenging task due to large displacements between them in successive CT scans, which may be caused by variation in respiratory volumes and patient positioning. For these reasons, the registration of the successive CT lung data taken at different times is the main goal of this paper.

Most of the known registration methods fall into two main categories: feature-based and area-based techniques [1]. Feature-based techniques use sparse geometric features such as points, curves, and/or surface patches, and their correspondences to compute an optimal transformation. Area-based methods, such

as the classical least square correlation, directly match image signals to avoid feature extraction [2]. More powerful mutual information (MI) based image registration [3] exploits a probabilistic similarity measure that allows for more general types of signal deviations than correlation. The statistical dependency between two data sets is measured by comparing a joint empirical distribution of the corresponding signals in the two images to the joint distribution of the independent signals (e.g., see [4, 5] for more details about the existing approaches for medical images registration).

In this paper, we consider a more general case of registering 3D segmented lung data to a prototype (the reference data, usually the CT collected at the first visit of the patient) with similar visual appearance under their relative 3D affine transformations and monotone variations of signal correspondences. To suppress the variations between the images, all test and training images were equalized to the same signal ranges, thus having the same dimensions in their co-occurrence matrices. Generally, any equalization scheme can be used. The co-registered equalized images are described with a characteristic subset of signal co-occurrence statistics. The description implicitly “homogenizes” the images, i.e., considers them as spatially homogeneous patterns with the same statistics. Our approach differs from the feature-based registration in that the statistics characterize the whole object, and from conventional area-based techniques in that the similarities between the statistics rather than pixel-to-pixel correspondences are measured.

2 MGRF Based Image Registration

2.1 Basic notation.

- $\mathcal{Q} = \{0, \dots, Q - 1\}$ – a finite set of scalar image signals (e.g., gray levels).
- $\mathbf{R} = [(x, y, z) : x = 0, \dots, X - 1; y = 0, \dots, Y - 1; z = 0, \dots, Z - 1]$ – a 3D arithmetic lattice supporting digital low dose CT (LDCT) image data $g : \mathbf{R} \rightarrow \mathcal{Q}$.
- $\mathbf{R}_p \subset \mathbf{R}$ – an arbitrary-shaped part of the lattice occupied by a prototype.
- $\mathcal{N} = \{(\xi_1, \eta_1, \zeta_1), \dots, (\xi_n, \eta_n, \zeta_n)\}$ – a finite set of (x, y, z) -coordinate offsets defining neighboring voxels, or neighbors $\{((x + \xi, y + \eta, z + \zeta), (x - \xi, y - \eta, z - \zeta)) : (\xi, \eta, \zeta) \in \mathcal{N}\} \wedge \mathbf{R}_p$ interacting with each voxel $(x, y, z) \in \mathbf{R}_p$.
- \mathbf{T} – an indicator of vector or matrix transposition.

The set \mathcal{N} yields a 3D neighborhood graph on \mathbf{R}_p describing translation invariant pairwise interactions between the voxels with $|\mathcal{N}|$ families $\mathcal{C}_{\xi, \eta, \zeta}$ of the 2nd-order cliques $c_{\xi, \eta, \zeta}(x, y, z) = ((x, y, z), (x + \xi, y + \eta, z + \zeta))$ shown in Fig. 1.

Quantitative interaction strengths for the clique families are given by a vector $\mathbf{V}^T = [\mathbf{V}_{\xi, \eta, \zeta}^T : (\xi, \eta, \zeta) \in \mathcal{N}]$ of potentials $\mathbf{V}_{\xi, \eta, \zeta}^T = [V_{\xi, \eta, \zeta}(q, q') : (q, q') \in \mathcal{Q}^2]$ being functions of signal co-occurrences in the cliques.

Data normalization: To account for possible monotone (order-preserving) changes of signals, e.g., due to different sensor characteristics, every LDCT data set is equalized using the cumulative probability distribution of its signals.

Markov-Gibbs random field-based appearance model: The main idea of learning the appearance model using an MGRF is to find the neighborhood system (i.e., the subsets of voxels which have strong relations with each current voxel) and to estimate the interaction between each two voxels in this neighborhood system (see Fig. 1).

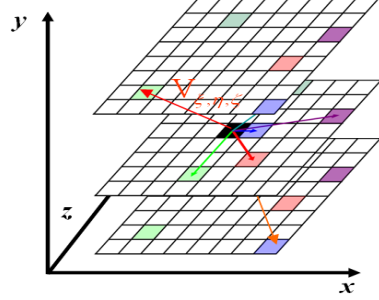


Fig. 1. Pairwise voxel interaction system.

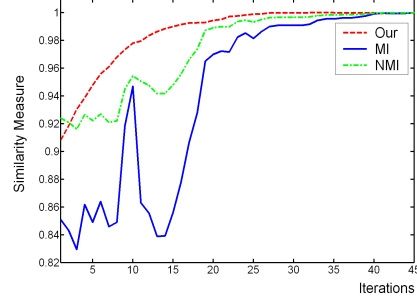


Fig. 2. Values of the Gibbs energy, MI, and NMI at the successive steps of the gradient ascent based search.

In a generic MGRF with multiple pairwise interaction in Fig. 1 [6], the Gibbs probability $P(g) \propto \exp(E(g))$ of an object g aligned with the prototype g° on \mathbf{R}_p is specified with the Gibbs energy $E(g) = |\mathbf{R}_p| \mathbf{V}^\top \mathbf{F}(g)$. Here,

- $\mathbf{F}^\top(g)$ is the vector of scaled empirical probability distributions of signal co-occurrences over each clique family: $\mathbf{F}^\top(g) = [\rho_{\xi,\eta,\zeta} \mathbf{F}_{\xi,\eta,\zeta}^\top(g) : (\xi, \eta, \zeta) \in \mathcal{N}]$;
- $\rho_{\xi,\eta,\zeta} = \frac{|\mathcal{C}_{\xi,\eta,\zeta}|}{|\mathbf{R}_p|}$ is the relative size of the family $\mathcal{C}_{\xi,\eta,\zeta}$, and
- $\mathbf{F}_{\xi,\eta,\zeta}(g)$ is the vector of empirical probabilities for this family: $\mathbf{F}_{\xi,\eta,\zeta}(g) = [f_{\xi,\eta,\zeta}(q, q'|g) : (q, q') \in \mathcal{Q}^2]^\top$ where
 - $f_{\xi,\eta,\zeta}(q, q'|g) = \frac{|\mathcal{C}_{\xi,\eta,\zeta;q,q'}(g)|}{|\mathcal{C}_{\xi,\eta,\zeta}|}$ are empirical signal co-occurrence probabilities, and
 - $\mathcal{C}_{\xi,\eta,\zeta;q,q'}(g) \subseteq \mathcal{C}_{\xi,\eta,\zeta}$ is a subfamily of the cliques $\mathcal{C}_{\xi,\eta,\zeta}(x, y, z)$ supporting the co-occurrence ($g_{x,y,z} = q, g_{x+\xi,y+\eta,z+\zeta} = q'$) in g .

The co-occurrence distributions and the Gibbs energy for the object are determined over \mathbf{R}_p , i.e., within the prototype boundary after an object is affinely aligned with the prototype. To account for the affine transformation, the initial image is resampled to the back-projected \mathbf{R}_p by bilinear interpolation.

The appearance model consists of the neighborhood \mathcal{N} and the potential \mathbf{V} to be learned from the prototype. Below we will show how to estimate \mathcal{N} and \mathbf{V} for the lung tissues from the LDCT images.

Learning the potentials: In the first approximation, the maximum likelihood estimate (MLE) of \mathbf{V} is proportional to the scaled and centered empirical co-

occurrence distributions for the prototype¹:

$$\mathbf{V}_{\xi,\eta,\zeta} = \lambda \rho_{\xi,\eta,\zeta} \left(\mathbf{F}_{\xi,\eta,\zeta}(g^\circ) - \frac{1}{Q^2} \mathbf{U} \right); (\xi, \eta, \zeta) \in \mathcal{N}$$

where \mathbf{U} is the vector with unit components. The common scaling factor λ is also computed analytically; it is approximately equal to Q^2 if $Q \gg 1$ and $\rho_{\xi,\eta,\zeta} \approx 1$ for all $(\xi, \eta, \zeta) \in \mathcal{N}$. In our case it can be set to $\lambda = 1$ because the registration uses only relative potential values and energies.

Learning the characteristic neighbors: To find the characteristic neighborhood set \mathcal{N} , the relative energies $E_{\xi,\eta,\zeta}(g^\circ) = \rho_{\xi,\eta,\zeta} \mathbf{V}_{\xi,\eta,\zeta}^\top \mathbf{F}_{\xi,\eta,\zeta}(g^\circ)$, i.e., the scaled variances of the corresponding empirical co-occurrence distributions for the clique families, are compared for a large number of possible candidates.

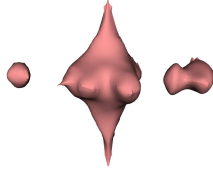


Fig. 3. The 3D neighborhood system estimated for the lung tissues.

Table 1. Alignment errors (in mm) for the expert-identified landmarks over the 100 data sets, and the execution times comparing the proposed approach against MI- and NMI-based methods. SD stands for standard deviation.

	Our	MI	NMI
Mean	1.9	5.1	4.8
SD	1.1	2.3	1.9
Time (min)	12	9.0	9.0

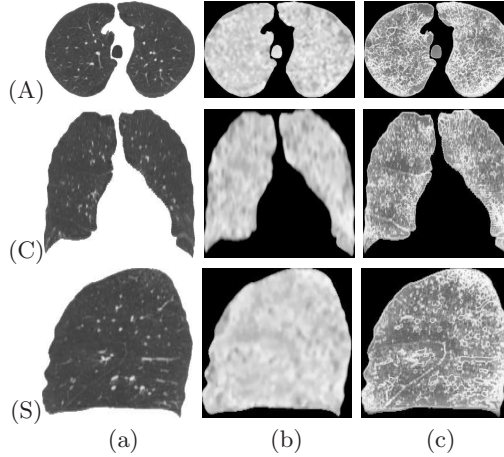


Fig. 4. 3D voxel-wise Gibbs energies projected onto the 2D axial (A), coronal (C), and sagittal (S) planes for visualization: 2D slices of the original LDCT images (a) and the voxel-wise Gibbs energies for $|\mathcal{N}| = 275$ which is estimated using LCDG model (b) and $|\mathcal{N}| = 3927$ which is estimated using tradition EM algorithm (c).

To automatically select the characteristic neighbors, we consider an empirical probability distribution of the energies as a mixture of a large “non-characteristic” low-energy component and a considerably smaller characteristic high-energy component: $P(E) = \pi P_o(E) + (1 - \pi) P_{hi}(E)$. The components

¹ For complete proof, please see: <https://louisville.edu/speed/bioengineering/faculty/bioengineering-full/dr-ayman-el-baz/supplemental-materials>

$P_{lo}(E)$ and $P_{hi}(E)$ are of arbitrary shape and thus are approximated with linear combinations of positive and negative Gaussians (LCDG) using efficient Expectation-Maximization-based algorithms introduced in [7] (the latter estimate both the components and the prior π). The intersection of the estimated mixture components using the LCDG model gives an energy threshold θ for selecting the characteristic neighbors: $\mathcal{N} = \{(\xi, \eta) : E_{\xi, \eta}(g^\circ) \geq \theta\}$ where $P_{hi}(\theta) \geq P_{lo}(\theta)\pi/(1 - \pi)$. The above example results in the threshold $\theta = 28$ producing 275 characteristic neighbors shown in Fig. 3. Figure 4 presents the relative 3D voxel-wise Gibbs energies $e_{x,y,z}(g^\circ)$ for this system:

$$e_{x,y,z}(g^\circ) = \sum_{(\xi, \eta, \zeta) \in \mathcal{N}} V_{\xi, \eta, \zeta}(g_{x,y,z}^\circ, g_{x+\xi, y+\eta, z+\zeta}^\circ)$$

Appearance-based registration The desired affine transformation of an object g corresponds to a local maximum of its relative energy $E(g_{\mathbf{a}}) = \mathbf{V}^\top \mathbf{F}(g_{\mathbf{a}})$ under the learned appearance model $[\mathcal{N}, \mathbf{V}]$. Here, $g_{\mathbf{a}}$ is the part of the object image reduced to \mathbf{R}_p by a 3D affine transformation $\mathbf{a} = [a_{11}, \dots, a_{34}]$: $x' = a_{11}x + a_{12}y + a_{13}z + a_{14}$; $y' = a_{21}x + a_{22}y + a_{23}z + a_{24}$; $z' = a_{31}x + a_{32}y + a_{33}z + a_{34}$. Its initial step is a pure translation ($a_{11} = a_{22} = a_{33} = 1$; $a_{12} = a_{13} = a_{21} = a_{23} = a_{31} = a_{32} = 0$) ensuring the most “energetic” overlap between the object and prototype. In other words, the chosen initial position ($a_{14}^*, a_{24}^*, a_{34}^*$) in Fig. 5 maximizes the energy. Then, the gradient ascent based search for the local energy maximum closest to the initialization selects all the 12 parameters \mathbf{a} . Note that this gradient-based optimization was used for all tested registration methods to estimate the goal parameters.

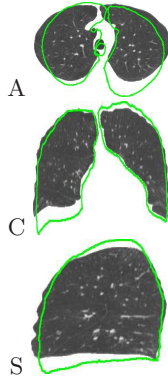


Fig. 5. Initialization of the proposed global registration algorithm.

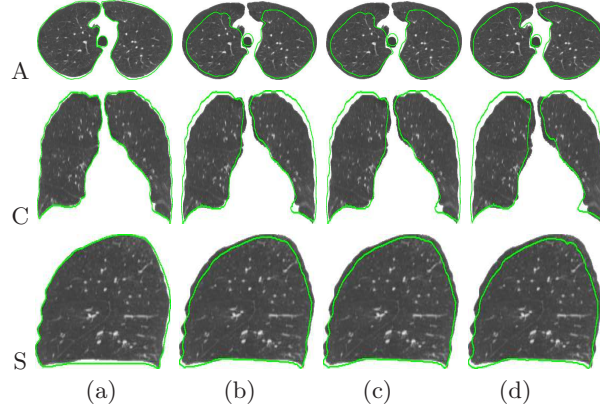


Fig. 6. Global registration results: our (a), MI-based (b), NMI-based (c), and SIFT-based (d) algorithms. These registration results were obtained from the segmented lung data without any pre-processing steps.

3 Experimental results

The proposed registration approach has been tested on the clinical data sets that have been collected from 100 patients at the end-expiration breath hold (EE-BH). Each patient has two LDCT scans, with the three-month period between each two successive scans. This clinical database was collected by the LDCT scan protocol using a multidetector GE Light Speed Plus scanner (General Electric, Milwaukee, USA) with the following scanning parameters: the slice thickness of 2.5 mm reconstructed every 1.5 mm; scanning pitch 1.5 mm; 140 KV; 100 MA; and the field-of-view 36 cm.

Results of the proposed global alignment of two lungs are shown in Figs. 6 and 7. It is clear from Fig. 6 (a) and Fig. 7 (c,d) that there are small misalignment at the lung surface, this is due to local deformation of lung tissues which come from breathing and heart beats. This can be handle by applying any local deformation transformation model (e.g., 3D cubic splines as a local transformation model [8]) as shown in Fig. 7 (f).

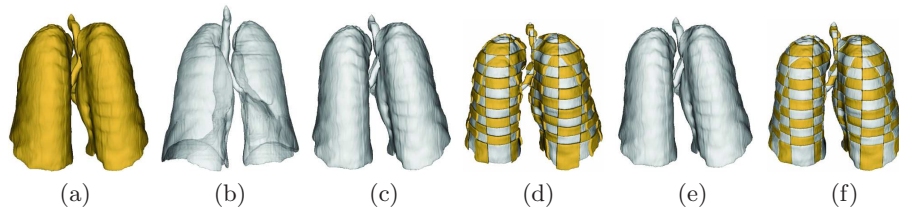


Fig. 7. 3D global and local registration: (a) the reference data, (b) the target data, (c) the target data after a 3D affine transformation, (d) the checkerboard visualization to show the motion of lung tissues, (e) non-rigid registration based using 3D cubic splines as a local transformation model [8], and (f) the checkerboard visualization to show the quality of the proposed local deformation model.

To highlight the advantages of the proposed registration approach, we compared, on segmented lung data, our global alignment to three popular conventional techniques, namely, to the area-based registration by mutual information (MI) [3] or normalized MI (NMI) [9] and to the feature-based registration that establishes correspondences between the images with 3D scale-invariant feature transform (SIFT) [10].

To clarify why the MI- or NMI-based alignment is less accurate, Fig. 2 compares the MI/NMI and Gibbs energy values for a sequence of affine parameters that appear at successive steps of the gradient ascent-based search for the maximum similarity in terms of mutual information or energy. Both the MI and NMI have many outstanding local maxima that potentially hinder the search, whereas the energy is much smoother and practically unimodal in these experiments. The 3D SIFT-based alignment fails because it cannot establish accurate correspondences between the similar lung areas.

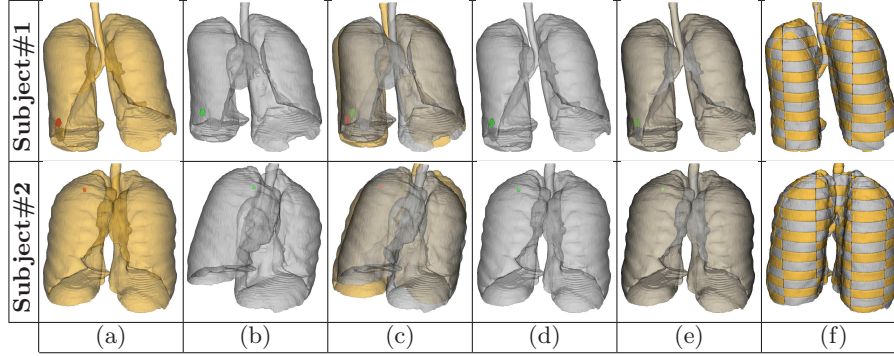


Fig. 8. Registration results of the proposed validation control study: (a) scan #1 before the movement of the patient; (b) scan #2 after patients rotation and translation; (c) the two scans superposed; (d) the registration results of the proposed approach; (e) superposition of the registered scan #2 with the reference scan #1; and (f) checkerboard visualization to show the quality of the proposed registration.

Validation: To validate the proposed approach, we have performed a control study using two subjects, each with one solid nodule. The CT data was collected with the same scanning protocol described above. In this new control study, we acquired two scans from the same patient at the EE-BH. The only difference between the two scans was that the patient was asked to make a global rotation and translation after the first scan. The purpose of this control study was to get the data from the same patient with minimal local deformation by breathing in order to test the proposed 3D global registration approach. Figure 8 illustrates the registration results of the proposed approach. We used scan #1 before the patients movement as a reference image to register scan #2 after the patient movement. The superposition of the registered and reference data and the checkerboard visualization in Fig. 8(e,f) demonstrates the high quality of the proposed registration. The average registration error is 1.4 mm with standard deviation ± 0.7 mm. This estimated error is based on calculating the Euclidean distance between 250 landmark points manually selected by a radiologist on the reference and the registered target data.

Additionally, our registration approach was validated on the 100 data sets based on using anatomical landmark correspondences selected by a radiologist. After the target data is aligned to the reference data, 10 anatomical landmarks, between both the reference and registered target data, are identified by the radiologist. The registration accuracy is quantitatively assessed by calculating the Euclidian distance for each expert-identified landmark on the registered data and its correspondence on the reference data. The error statistics for this validation experiment as well as the average execution time for our approach and the MI- and NMI-based methods are summarized in Table 1.

4 Conclusions

In this paper we introduced a new approach to align 3D CT data of a lung object with a given prototype whose appearance is modeled with a new 3D Markov-Gibbs random field with pairwise interaction model. Experimental results confirm that lung registration based on our new Markov-Gibbs appearance model is more robust and accurate than popular conventional algorithms. Moreover, it is worth mentioning that the proposed registration approach is not limited only to lung objects, but it is also suitable for registering any 3D texture medical objects. Furthermore, the proposed approach can be integrated with any nonrigid registration algorithm (e.g., cubic B-Splines based techniques). In our future work will use the evaluation framework proposed by van de Kraats et al. [11] to evaluate the performance of our registration approach against the MI- and NMI-based approaches as well as higher-order MI-based techniques (e.g., [12]).

Acknowledgment: This research work has been supported by Research Scholar Grants, 120556-RSG-11-266-01-CCE, from the American Cancer Society.

References

1. Zitova, B., Flusser, J.: Image Registration Methods: A Survey. *Image Vis. Comput.* 21, 977–1000 (2003).
2. Pope, P., Theiler, J.: Automated Image Registration (AIR) of MTI imagery. In: *Proc. SPIE 5093*. 27, 294–300 (2003).
3. Viola P.: Alignment by maximization of mutual information. Ph.D. Dissertation, MIT, Cambridge, MA (1995).
4. Pluim, J., Maintz, J., Viergever, M.: Mutual-information Based Registration of Medical Images: A survey. *IEEE Trans. Medical Imaging* 22(8), 986–1004 (2003).
5. Khalifa, F., Beache, G.M., Gimel'farb, G., Suri, J., El-Baz, A.: State-of-the-Art Medical Images Registration Methodologies: A Survey. In: El-Baz, A., Acharya, R., Mirmehdi, M., Suri, J. (eds), *Handbook of Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*. vol. 1, ch.9, pp. 235–277, Springer-Verlag (2011).
6. Farag, A., El-Baz, A., Gimel'farb, G.: Precise Segmentation of Multimodal Images. *IEEE Trans. Image Process.* 15(4), 952–968 (2006).
7. El-Baz, A., Elnakib, A., Khalifa, F., Abou El-Ghar, M., McClure, P., Soliman, A., Gimel'farb, G.: Precise Segmentation of 3D Magnetic Resonance Angiography. *IEEE Trans. Biomed. Eng.* (in press), DOI:10.1109/TBME.2012.2196434 (2012).
8. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J.: Non-rigid Registration Using Free-Form Deformations: Application to Breast MR Images. *IEEE Trans. Medical Imaging* 18(8), 712–721 (1999).
9. Studholme, C., Hill, D.L.G., Hawkes, D.J.: An Overlap Invariant Entropy Measure of 3D Medical Image Alignment. *Pattern Recognition* 32(1), 71–86 (1999).
10. Aly, A.: Local Features Invariance Beyond 2D Gray Spaces. Ph.D. Dissertation, University of Louisville, Louisville, KY (2007).
11. Van de Kraats, E.B., Pennney, G.P., Tomazevic, D., Van Walsum, T., Niessen W.J.: Standardized Evaluation Methodology for 2-D-3-D Registration. *IEEE Trans. Medical Imaging* 24(9), 1177–1189 (2005).
12. Rueckert, D., Clarkson, M.J., Hill, D.L.G., Hawkes, D.J.: Non-rigid Registration using Higher-Order Mutual Information. In: *Proc. SPIE 3979*. 27, 439–447 (2000).