A Fast Piece-wise Deformable Method for Multi-Modality Image Registration

Girish Gopalakrishnan S V Bharath Kumar Ajay Narayanan Rakesh Mullick

Imaging Technologies Lab, GE Global Research Center, Bangalore-560066, India.

{*girish.gopalakrishnan, bharath.sv, ajay.narayanan, rakesh.mullick*}*@ge.com*

Abstract

Medical image fusion is becoming increasingly popular for enhancing diagnostic accuracy by intelligently 'fusing' information obtained from two different images. These images may be obtained from the same modality at different time instances or from multiple modalities recording complementary information. Due to the nature of the human body and also due to patient motion and breathing, there is a need for deformable registration algorithms in medical imaging. Typical non-parametric (deformable) registration algorithms such as the fluid-based, demons and curvature-based techniques are computationally intensive and have been demonstrated for mono-modality registrations only. We propose a fast and deformable algorithm using a 2*-tiered strategy wherein a global MI-based affine registration is followed by a local piece-wise refinement. We have successfully tested this method on CT and PET images and validated the same using clinical experts.*

1. Introduction

Image registration (fusion) is an important and a classical problem in medical image analysis and understanding for patient care. The key step of image registration is to find a spatial transformation such that a chosen similarity metric between two or more images of the same scene achieves its maximum.

The accurate and reliable registration of multi-modality images is an important tool for assessing temporal and structural changes between images. Different imaging modalities provide information about different properties of the underlying tissues such as the X-ray attenuation coefficient from X-ray Computed Tomography (CT), and proton density or proton relaxation times from Magnetic Resonance (MR) imaging. These images allow clinicians to gather information about the size, shape and spatial relationship between anatomical structures and any pathology, if present. Other imaging modalities provide functional information such as the blood flow form ultrasound doppler or glucose metabolism from Positron Emission Tomography (PET) or Single-Photon Emission Tomography (SPECT), and permit clinicians to study the relationship between anatomy and physiology.

The bulk of registration algorithms in medical imaging [11, 13, 2] can be classified as being either frame-based, point landmark-based, surface-based or voxel-based. Recently, the voxel-based similarity approaches to image registration have attracted significant attention since these full-volume-based registration algorithms do not rely upon data reduction, require no segmentation, and involve little or no user interactions. More importantly, they can be fully automated and quantitative assessment becomes possible. Maintz *et al.* [10] give a good overview while Studholme *et al.* [4] and Penny *et al.* [8] provide a detailed comparative assessment. In particular, the voxel-based similarity measures based on joint entropy [3], mutual information (MI) [4, 1] and normalized mutual information [5, 6] have been shown to align images acquired with different imaging modalities robustly. The utility of these information-theoretic measures arises from the fact that they make no assumptions about the actual intensity values in the images, but instead measure statistical relationships between the two images [11].

Generally, global registration is performed to capture the gross misalignment between images. In the case of medical data, especially with thoracic imaging, deformable techniques are required to capture breathing artifacts, which calls for local registration so as to cope with local geometric differences. This means that the final mapping has to comprise, for example, a global rigid/affine transformation as well as a locally adaptive transformation so as to cope with gross and subtle misalignments between the images. The major class of locally adaptive non-parametric transformations in medical image analysis include computationally intensive methods such as the fluid, diffusion, and curvature-based techniques. Though these methods capture local vari-

ations, they would require landmarks to be defined before registration, making them semi-automatic (requiring the involvement of clinician). Furthermore, these methods are proven to be suitable especially for mono-modality, inter-subject and time-series registration [10]. On the other hand, in a multi-modality setting, rigid/affine transformation is widely used, which fails to recover motion artifacts such as breathing in the case of thoracic imaging.

In order to address the aforementioned issues, in this paper, we propose a generic framework for fast and accurate MI-based deformable registration of multi-modality images. Instead of performing elastic transformations which are very computationally intensive [10], we compute affine transformations both at the global and local levels (on pre-computed Volumes of Interest (VOIs) called *Blobs*), essentially resulting in piece-wise deformation model. These VOIs can be considered as analytical landmarks for other registration schemes, which can be both parametric and non-parametric. The proposed method involves a global registration followed by local registrations in the pre-computed VOIs. The initial global match is performed to correct for gross misalignments. A pyramid-based multi-resolution decomposition [12] at this stage helps to improve the robustness and timing of the global registration. The latter step involves automatic selection of a VOI, followed by local matching (typically affine/rigid or deformable) and finally interpolating the individual transformations obtained across the entire volume. Also, even if deformable methods are used locally, the overall registration is fast considering the data size in the VOIs as compared to the original data size.

The paper is organized as follows. In Section 2, we provide a broad overview on the rigid/affine registration method along with the details on metric and optimizer. We provide motivation for piece-wise affine transformation and build blob framework for fast deformable registration. We present the results of our proposed blob-based registration method on various modalities by comparing it with affine registration in Section 3. In Section 4, we conclude our paper by showing that our proposed method provides a framework for fast deformable registration of volumetric data from CT-PET images.

2 Blob-Based Deformable Registration

As mentioned in Section 1, registration of medical images requires computation of transformations at both global and local levels to achieve good accuracy. We propose a piece-wise affine registration which is an approximation to deformable method. The proposed method is motivated from the fact that the local subtle variations can be captured with a set of piecewise affine transformations within the pre-computed analytical landmarks called *blobs* and is computationally less intensive compared to that of elastic transformation. These analytical landmarks are computed automatically thus removing the need for clinician's intervention in marking the landmarks.

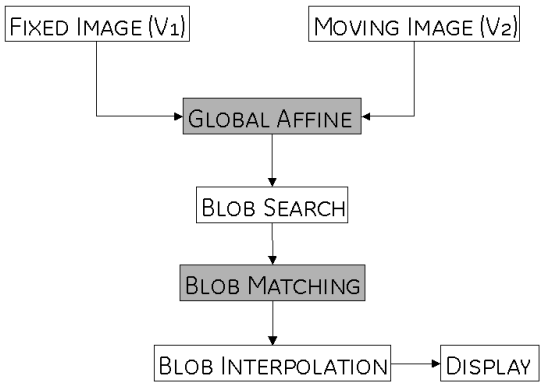


Figure 1. Deformable registration using a 2-tiered strategy

Figure 1 shows the steps involved in the 2-tiered strategy of our proposed blob-based deformable method. Blob-based deformable registration involves defining and identifying blobs/VOIs in the fixed volume. We then perform a blob-based local matching or registration (typically affine/rigid, can be deformable also), followed by blob-interpolation, i.e. estimate the deformation at every point in the moving volume based on the transformation matrices obtained at the blobs.

We automate the finding of this small set of 'landmark-like' points in the moving image, defined as the centers of 'blobs' with high levels of spatially organized differentiation in their intensity values, and low levels of symmetry. Blob search constitutes the finding of these 'landmark-like' points. We use standard methods to find an approximate transformation quickly by maximizing mutual information (MI) over affine candidates. For each blob in the moving image, this matching gives an initial guess for the blob's transformation to a particular position in the fixed image. We perform a local search for improvements on this guess and optionally retain the blob only if the match found exceeds a pre-assigned quality threshold. The search can also use a multi-resolution approach (seeking a coarse fit in a reduced resolution version of the images, then iteratively refining), both for computational efficiency and to avoid inappropriate local optima. This search process is called blob matching.

The blob centers and the points to which they correspond provide a matched landmark list, to which any standard interpolation or fitting method can be applied [9].

2.1 Affine Registration by Maximizing Mutual Information

The affine registration of two volumetric images, V_1 (fixed image) and V_2 (moving Image) involves the registration of V_2 to V_1 by determining the best affine transformation T^*, which maximizes a given metric, say $\phi(.)$.

$$T^* = \arg\max_T \phi(V_1, T(V_2)) \tag{1}$$

where

$$T = \begin{bmatrix} a_{11} & a_{12} & a_{13} & x' \\ a_{21} & a_{22} & a_{23} & y' \\ a_{31} & a_{32} & a_{33} & z' \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

is an affine transformation. In T, the sub-matrix

$$S = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{3}$$

can be decomposed into shear, scale and rotation and the vector $\begin{bmatrix} x' & y' & z' \end{bmatrix}^T$ contains the translations along the 3-dimensions. The volume $T(V_2)$ is the transformed image of V_2 using the affine transformation T. It can be seen that each of shear, scale, rotation and translation are represented with 3 parameters affecting the 3-dimensions.

The mutual information between two discrete random variables U and V corresponding to V_1 and V_2 is defined as

$$I(U;V) = -\sum_{u\in U', v\in V'} P(u,v)\log\frac{P(u)P(v)}{P(u,v)} \tag{4}$$

$$= -\sum_{i=1}^{N}\sum_{j=1}^{M} p_{ij}\log\frac{p_i q_j}{p_{ij}} \tag{5}$$

where the random variables U and V take values in the set $U' = \{u_1, u_2, \cdots, u_N\}$ and $V' = \{v_1, v_2, \cdots, v_M\}$ having probabilities $\{p_1, p_2, \cdots, p_N\}$ and $\{q_1, q_2, \cdots, q_M\}$ such that $P(U = u_i) = p_i$, $P(V = v_i) = q_i$, $p_i > 0$, $q_i > 0$ and $\sum_{u\in U'} P(u) = 1$, $\sum_{v\in V'} P(v) = 1$. $P(U,V)$ is the joint probability of the random variables U and V. MI represents the amount of information that one random variable, here V contains about the second random variable, here U and vice-versa. $I(U;V)$ is the measure of shared information or dependence between U and V. It is to be noted that $I(U;V) \geq 0$ with equality if, and only if, U and V are independent. MI measures the dependence of the images by determining the distance of their joint distribution p_{ij} to the joint distribution in case of complete independence, $p_i q_j$. Extending from (1), the best affine transformation T^*_{MI}, which maximizes MI defined in (4) is given as

$$T^*_{MI} = \arg\max_T I(V_1; T(V_2)) \tag{6}$$

2.2 Blob Search

A candidate VOI qualifies as a blob if it algorithmically resembles a region of radius R containing what a human would recognize as a landmark. The size of each blob is chosen such that when the spatial samples are chosen from a blob, it contains statistically sufficient voxels for further calculations. Blob search involves finding the centers of a small set of radius R regions in the moving image depending on the information content present in it. Blob search involves 3 steps.

1. Blob partitioning
2. Blob score evaluation
3. Blob ranking

2.2.1 Blob partitioning

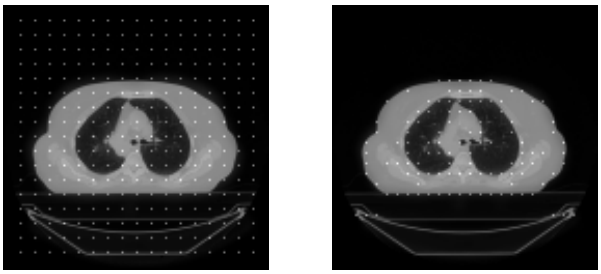
To start with, the moving image is partitioned into smaller VOIs/blobs. Each of these blobs would then be locally registered. The shape and placement of the blobs govern the accuracy of the deformation field estimation. The blobs can be cubes or spheres and the initial arrangement could be a Hexagonal Close Packing (HC), Simple Cubic Packing (SC) or Body Centered Cubic Packing (BCC) configuration, with the inspiration from molecular chemistry.

2.2.2 Blob score evaluation

Based on the information content in each blob, a statistical metric (normalized to the blob volume) is calculated. This metric has to be correlated with the predicted deformation within each blob. The metric can be the extent of surface mismatches between blobs, the MI value inside the blob, the volume overlap of the two images in the blob etc. The exact selection of the metric is application dependent.

2.2.3 Blob ranking

This step ranks the blobs based on their expected level of deformation in the descending order. The advantage of this step would be the tradeoff for computation time without considerable loss in accuracy by retaining the critically important blobs (see Figure 2). This step is analogous to a manual process wherein with increasing availability of time, the clinical expert would deposit more landmarks (while a skilled operator can define landmark pairs in as much as $6-10$ seconds, the program generated up to 200 landmarks in less than 1 second for a volume of size $128\times128\times109$). This actively reduces the number of landmarks to actually use in the registration step based on a cutoff percentage of the metric value or the number of landmarks to process can



(a) (b)

Figure 2. (a) Blob centers selected based on HC arrangement. (b) Blobs retained for local refinement after the 'blob ranking' step.

be specified. Let us consider two volumes V_1 and V_2. The metric based on the information content in the blobs between these volumes is computed as follows.

$$M = \alpha v_o(V_1, V_2) + \beta g(V_2) \tag{7}$$

where M is the metric, $v_o(V_1, V_2)$ is the volume overlap between V_1 and V_2 given as

$$v_o(V_1, V_2) = 0.5 - (DSC(V_1, V_2)\ mod\ 0.5) \tag{8}$$

and

$$g(V_2) = \frac{1}{n} \sum_i \frac{V_2^i}{\max(V_2)}\ if\ V_2^i > \gamma \tag{9}$$

where V_2^i is the i^{th} voxel's intensity of V_2, $\max(V_2)$ is the maximum voxel intensity of V_2 and γ is some threshold. α and β are the weighting factors such that $\alpha + \beta = 1$ so that $0 \leq M \leq 1$. $DSC(V_1, V_2)$ is the dice similarity coefficient between V_1 and V_2 defined as

$$DSC(V_1, V_2) = \frac{2(V_1 \cap V_2)}{(V_1 \cup V_2)} \tag{10}$$

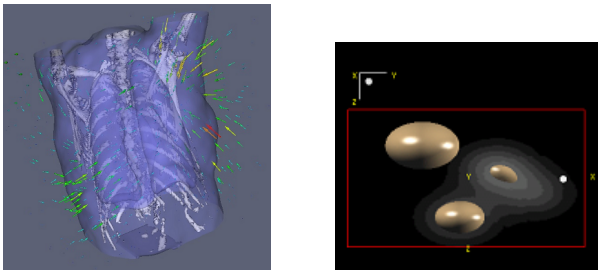
2.3 Blob Matching

With the selection of blobs, we perform a standard search for a 12-parameter affine fit between the images, for which many schemes exist in the literature. We prefer the use of an affine match to the search for a rigid match because the collection of affine transformations is a flat 12-parameter space within which search is easier than when confined to its curved 6-parameter subset of rigid motions. The affine search is thus by most methods faster, as well as having extra six degrees of freedom that allow a better match, unless there are strong a priori reasons to expect a rigid fit to be possible. We could use any other fitting method like rigid, affine or curvilinear, provided that it is fast and reliably gives a fair match: high accuracy is not necessary in this initial step of whole-image match.

Having removed the gross differences at global level by computing an affine transformation relating the moving and fixed images, we next proceed to remove the local subtle variations by performing the matching of precomputed blobs. An affine process like the one used in the whole-image match uses the initial guess from the global output to start a search process and thus converges on a nearby better fit. Because this region matching involves multiple pixels or voxels in the blob, it can often locate the best fit with an error substantially less than the spacing of neighboring grid points in the fixed image. Since the blob is small compared to the entire moving image, for most search methods convergence to the best fit is many times faster. To reduce the chances of a mismatch, the ratio $blobSizeMoving/blobSizeFixed$ is chosen to be > 1.

2.4 Blob Interpolation

The blob matching results in an affine transformation around each blob center in the moving image, which is linear when considered as mapping a space of vectors based at the said center to a space of vectors based at the corresponding point in the fixed image, which should approximate the global transformation around the said center. The transformations obtained for each blob, influences a definite region around it. Hence the corresponding transformation for each voxel will be proportional to its distance from the blob center of the nearest n blobs. This distance from the blob centers form a weighting factor to provide appropriate influence effects. An exponential decay around the center of each blob is used to define its zone of influence (see Figure 3). This ensures that points outside this region remain unaffected. This step is necessary to avoid forcing a transformation to an already well-transformed point. A well-transformed point is defined as that point which has not been identified as interesting and furthermore has obtained a convincing transformation after the global affine registration and thus does not require any further refinement. This step of blob interpolation handles the 3D translation component. One could also use a quaternion-based interpolation similar to Shekhar *et al.* [14] to represent 3D rotations. Once the displacement vectors are obtained for the entire volume, the corresponding pixel intensities are obtained from the moving image by tri-linear interpolation. This provides sub-pixel accuracy and avoids artifacts in the final image.



(a) (b)

Figure 3. (a) The deformation field overlaid on a lung image (b) The zone of influence of each blob

Patient	GA	BR	BI	Total	sec/slice
1	2.614	0.701	2.294	5.609	0.1002
2	2.784	1.172	3.034	6.990	0.0852
3	2.584	1.101	3.029	6.714	0.0819
4	2.805	1.032	2.423	6.260	0.1118
5	2.551	1.007	2.990	6.584	0.0803
6	2.735	0.981	2.795	6.511	0.0794
7	2.731	0.979	2.792	6.502	0.0793
8	2.641	0.719	2.342	5.702	0.1018

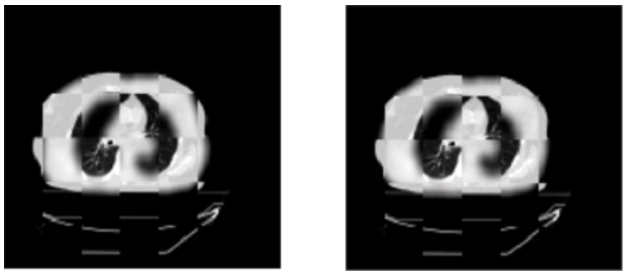
Table 1. Timing results (in seconds) for 8 pairs of data. GA=Global Registration, BR=Blob Registration, BI=Blob Interpolation

3 Results and Validation

The proposed approach was tested on implicit function-based 3-D phantom images [7] and on eight pairs of CT-PET thoracic datasets [see Figure 4]. The CT datasets were acquired using GE LightSpeed QXi/Plus multi-detector CT scanner (Matrix Size: 512×512, DFOV: $48 - 63$ cm, Slice Thickness: $2.5 - 5.0$ mm) and PET (Transmission [*Tr*] and Emission [*Em*]) datasets (Matrix Size: 128×128, DFOV: $55 - 66$ cm, Voxel Size: $4.25 - 5.14$ mm) from multiple vendors. In our experiments, 12-parameter affine transform with 2 levels of multi-resolution decomposition is used while maximizing MI as the similarity metric. While performing blob search, we used 'HC' as the initial packing arrangement for blob partitioning along with 'surface mismatching' as the metric for blob score evaluation. The registered outputs from these images were validated using multiple clinical experts. These readers were only shown the CT and *Em* data for each case and empowered with a visualization tool offering a common cursor, and allowing selection of the orthogonal views, window-level and zoom. Point-pairs were recorded in a sequential manner alternating between CT and PET (*Em*). The mis-registration between the recorded points was calculated on the basis of a mean-squared measure and categorized based on reader and registration methodology. The mean registration error across all cases for rigid registration was 7.86 mm ($\sigma = 2.93$ mm, $p = 0.552$) and that for non-rigid registration was 7.13 mm ($\sigma = 2.55$ mm, $p = 0.521$). Inter-observer variability (3 readers) was computed to be 2.66 mm. The timing results obtained on an Intel 2 GHz processor with 2 GB RAM are included in Table 1.

4 Conclusions and Future Direction

The current methodology used to down-sample the number of blobs in an image is dependant on the modality of the image and also the appropriate threshold that is selected.



(a) (b)

Figure 4. Fixed and moving images tiled using checkerboard (a) after global affine registration. (b) after blob-based registration

We plan on making this step more modality independent by applying image specific tests (homogeneity test etc). We are also experimenting on other techniques to extrapolate the local transforms.

References

[1] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multimodality image registration using information theory. In *Information Processing in Medical Imaging: Proc. 14th International Conference (IPMI '95)*, pages 263–274, 1995.

[2] C. R. Maurer and J. M. Fitzpatrick. A review of medical image registration. In *Interactive Image-Guided Neurosurgery*, pages 17–44, 1993.

[3] C. Studholme, D. L. G. Hill, and D. J. Hawkes. Multiresolution voxel similarity measures for mr-pet registration. In *Information Processing in Medical Imaging: Proc. 14th International Conference (IPMI '95)*, pages 287–298, 1995.

[4] C. Studholme, D. L. G. Hill, and D. J. Hawkes. Automatic three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Med. Phys.*, 24(1):25–35, January 1997.

[5] C. Studholme, D. L. G. Hill, and D. J. Hawkes. An overlap invariant entropy measure of 3d medical image alignment. *Pattern Recognition*, 32(1):71–86, 1998.

[6] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imag.*, 16(2):187–198, 1997.

[7] G. Gopalakrishnan, T. Poston, N. Nagaraj, R. Mullick, and J. Knoplioch. Implicit function-based phantoms for evaluation of registration algorithms. In *Proc. SPIE Int. Soc. Opt. Eng. Medical Imaging 2005: Image Processing*, volume 5747, pages 1310–1316, 2005.

[8] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes. A comparison of similarity measures for use in 2d3d medical image registration. *IEEE Trans. Med. Imag.*, 17(4):586–595, 1998.

[9] J. A. Little, D. L. G. Hill, and D. J. Hawkes. A review of medical image registration. *Computer Vision Image Understanding*, 66(2):223–232, 1997.

[10] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Med. Image Anal.*, 2(1):1–36, 1998.

[11] L. G. Brown. A survery of image registration techniques. *ACM Computing Surverys*, 24(4):325–376, December 1992.

[12] L. Ibanez, W. Schroeder, L. Ng, and J. Cates. *The ITK Software Guide: The Insight Segmentation and Registration Toolkit (version1.4)*. Kitware, 2003.

[13] P. A. Van Den Elsen, E-J. D. Pol, and M. A. Viergever. Medical image matching – A review with classification. *IEEE Eng. Med. Biol.*, pages 26–38, March 1993.

[14] R. Shekhar, V. Walimbe, S. Raja, and V. Zagrodsky. Automated 3-dimensional elastic registration of whole-body PET and CT from separate or combined scanners. *The Journal of Nuclear Medicine*, 46 (9):1488–1496, 2005.