Acta Cybernetica **16** (2003) 327-336.

Evaluation of a Fully Automatic Medical Image Registration Algorithm Based on Mutual Information*

Attila Tanács† and Attila Kuba‡

Abstract

Registration is a fundamental task in image processing. Its purpose is to find a geometrical transformation that relates the points of an image to their corresponding points of another image. Many registration algorithms have been proposed in the past decade. We present a fast, fully automatic algorithm that is capable of solving rigid-body registration of 3D images of the human brain where the images are taken by different imaging devices. We joined the Retrospective Registration Evaluation Project conducted by Vanderbilt University, USA. The evaluations of our results show that our method has the potential to produce satisfactory results, but visual inspection is necessary to guard against large errors.

Keywords: registration problem; automatic multimodal registration; registration accuracy;

1 Introduction

There is an increasing number of applications that require accurate aligning of one image with another taken from different viewpoints, by different imaging devices, or at different times. The geometrical transformation is to be found that maps a *floating image data set* in precise spatial correspondence with a *reference image data set*. This process of alignment is known as *registration*, although other words, such as *co-registration*, *matching*, and *fusion*, are also used. Examples of systems where image registration is a significant component include aligning medical images from different medical modalities for diagnosis, matching a target with a real-time image of a scene for target recognition, monitoring global land usage using satellite images, and matching stereo images to recover shape for autonomous navigation [1, 8].

*This work was supported by OTKA T023804 and FKFP 0908/1999 Grants.

†Department of Foundations of Computer Science, University of Szeged, 6720 Szeged, Árpád tér 2, Hungary, e-mail: tanacs@inf.u-szeged.hu

‡Department of Applied Informatics, University of Szeged, 6720 Szeged, Árpád tér 2, Hungary, e-mail: kuba@inf.u-szeged.hu

In this paper we focus on medical image registration which has a wide range of applications including combining information from multiple imaging modalities e.g., when relating functional information from nuclear medicine images to anatomy delineated in high-resolution MR or CT images, monitoring changes in size, shape, or image intensity over time intervals ranging from few seconds to even months or years, relating preoperative images and surgical plans to the physical reality of the patient in the operating room during image-guided surgery or during radiotherapy, and relating an individual's anatomy to a standardized atlas.

The registration technique for a given task depends on the knowledge about the characteristics of the type of variations. Registration methods can be viewed as different combinations of choices for the following four components [1]:

- *Search space* is determined by the type of transformation we have to consider, i.e., what is the class of transformations that is capable of aligning the images. Some widely used types are *rigid-body*, when only translations and rotations are allowed, *affine*, which maps parallel lines to parallel lines, and *nonlinear*, which can transform straight lines to curves.

- *Feature data set* describes what kind of image properties are used in matching. Features can be geometrical, e.g., automatically or manually selected landmark points, lines, and/or surfaces or the image intensity values can be used directly.

- *Similarity measure* is a function of the transformation parameters which shows how well the floating and the reference image fit. The task of registration is to optimize this function.

  In case of geometrical features this is usually a distance measure. When image intensity values are used, correlation, functions based on image intensity differences, or intensity similarity measures can be applied.

- *Search strategy* determines what kind of optimization method to use. Except for geometric features, where a direct solution of the problem might exist, an iterative approach is necessary.

In this paper we propose a fully automatic, iterative registration method that is capable of finding rigid-body transformations to align images from the same or different modalities (i.e., taken by the same or different imaging devices). Intensity similarity measures based on mutual information are used.

2 Methods

We follow the notations of [7]. Let X denote the object to be imaged, and let A and B be 3D images of X taken by the same or different imaging devices. The images usually have different fields of view, thus the domains Ω_A and Ω_B will be different:

$$A : x_A \in \Omega_A \mapsto A(x_A),$$
$$B : x_B \in \Omega_B \mapsto B(x_B).$$

$A(x_A)$ and $B(x_B)$ are referred to as the intensity values at spatial positions x_A and x_B, respectively. Intensity values represent some kind of measurement of the material in spatial positions of X, such as attenuation of X-ray beams in case of Computed Tomography (CT), changes in states of protons under changing the magnetic field properties in Magnetic Resonance Imaging (MRI), or distribution of nuclear tracers in case of Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT).

As the images A and B represent the same object X, there is a relation between the spatial locations in A and B. Position $x \in X$ is mapped to x_A in image A, and to x_B in image B. The registration process involves recovering the spatial transformation T which maps x_B to x_A over the entire domain of interest, which is the overlapping portion of the domains. This overlapping portion depends on the images A and B and on the spatial transformation T:

$$\Omega_{A,B}^T = \{x_A \in \Omega_A | T^{-1}(x_A) \in \Omega_B\}.$$

The medical images are discrete, they sample the object at a finite number of points. Taking this into account, we can define the domain Ω in the following way:

$$\Omega := \tilde{\Omega} \cap \Gamma_\zeta$$

where $\tilde{\Omega}$ is a bounded continuous set defining the volume of the patient imaged, and Γ is an infinite discrete sampling grid, which is characterized by the anisotropic sample spacing $\zeta = (\zeta^x, \zeta^y, \zeta^z)$. Sample spacing can be different for different images. These grid positions and the corresponding sample values together are referred to as voxels. For any given T, the intersection of discrete domains Ω_A and Ω_B might be the empty set, when no sample points will exactly overlap. To overcome this, we have to resample image intensities of image B in Ω_A. The simplest resampling method is to select the intensity value of the closest grid position of Ω_B. Linear or more complex interpolation methods can also be used. Let $\mathcal{T}$ denote the transformation that maps both the position and the associated intensity value at that position, and $B^{\mathcal{T}}$ the resampled image B.

The selection of the similarity measure is probably the most crucial part of a registration algorithm. We need a function which optimally has one global optimum at perfect alignment, has no local optimums, and is "smooth enough" to find this optimum fast. Practically it is very hard, or even impossible to find such a similarity measure, especially when the images are taken by different imaging devices. Many similarity measures were proposed in the past decade. We chose the measures based on the mutual information of the images proposed by Collignon et al. [4] and Wells et al. [12], and on the normalized mutual information of the images proposed by Studholme et al [11].

Both measures utilize the entropy of image A,

$$H(A) = -\sum_a p_A^T(a) \cdot \log p_A^T(a),$$

the entropy of image B,

$$H(B) = -\sum_b p_B^T(b) \cdot \log p_B^T(b),$$

and the joint entropy of images A and B,

$$H(A,B) = -\sum_a \sum_b p_{AB}^T(a,b) \cdot \log p_{AB}^T(a,b),$$

where p_A and p_B are the histograms, and p_{AB} is the co-occurrence matrix of the intensity values of images A and B. Mutual information is computed as

$$MI(A,B) = H(A) + H(B) - H(A,B),$$

and the normalized mutual information as

$$NMI(A,B) = \frac{H(A) + H(B)}{H(A,B)}.$$

We found that when mutual information is calculated over the overlapping domain $\Omega_{A,B}^T$, the failure rate is high [11]. We decided to use the whole Ω_A instead, in case of this measure, which solved the problem.

To speed up the registration process and to avoid falling into a local optimum, we use the Laplacian multiresolution pyramid representation of the images [2]. The search starts at the coarsest level. When an optimum is found, the result is propagated to the next, finer level. For the registration task of this project, we generate two new coarser pyramid levels.

We use Powell's direction set, iterative, nonlinear optimization algorithm to find the optimum of the similarity measure [10]. This method requires evaluating the similarity measure value for given transformation parameters only, no gradient or other information is necessary. The most time consuming part of the method is the evaluation itself, so it is crucial to avoid any unnecessary computations.

When resampling, we can take advantage of the fact that the transformation we are looking for is a linear one, which means that parallel lines, e.g., rows and columns remain parallel lines after applying the transformation. Using a general 3D line drawing algorithm [6], the resampling can be done using additions only, no multiplications are necessary. We use no interpolation of intensity values, we select the value of the nearest neighbor.

When the image sizes are no larger than 256 voxels, we can represent floating point numbers as 32-bit integers. Thus we have 1 sign bit, 9 bits for the integer part, and 22 bits for the fraction part. The precision of this representation is worse than that of the built-in floating point types, but is still good enough. We performed

numerical simulations to check the inaccuracy. Rigid-body transformations were generated randomly and applied to the points of a grid of size $256 \times 256 \times 25$, with grid spacing of $1.25, 1.25, 4.00$, respectively. Both real floating point and integer representations of reslicing methods were used and the maximum distance of the transformed points was calculated. The comparison showed that the maximum difference between spatial locations was about 0.02 voxels. For this price we get dramatic speed boost.

During resampling, we calculate probabilities p_A^T, p_B^T, and p_{AB}^T for each intensity value. The calculation of $MI(A,B)$ can be made faster as follows. By definition,

$$\begin{aligned} MI(A,B) &= -\sum_a p_A^T(a) \cdot \log p_A^T(a) - \sum_b p_B^T \cdot \log p_B^T(b) + \\ &\quad \sum_a \sum_b p_{AB}^T(a,b) \cdot \log p_{AB}^T(a,b) \\ &= \sum_a \sum_b (p_{AB}^T(a,b) \cdot \log p_{AB}^T(a,b) - p_A^T(a) \cdot \log p_A^T(a) - \\ &\quad p_B^T \cdot \log p_B^T(b)). \end{aligned}$$

Since the marginal probability distributions can be calculated from the joint probability distribution,

$$\begin{aligned} p_A^T(a) &= \sum_b p_{AB}^T(a,b), \\ p_B^T(b) &= \sum_a p_{AB}^T(a,b), \end{aligned}$$

mutual information can be calculated as

$$MI(A,B) = \sum_a \sum_b p_{AB}^T(a,b) \cdot (\log p_{AB}^T(a,b) - \log p_A^T(a) - \log p_B^T(b)).$$

The probabilities can have a value between 0 and 1, thus instead of calculating logarithmic values, we can use a precalculated lookup table, say the size of 10000 elements.

Real medical images can usually have intensity values ranging from -1000 to 4000. It means that the joint probability distribution table should have $5000 \cdot 5000 = 25000000$ elements, which is not feasible. That is why we scale intensity values so as to be in the $[0, 63]$, $[0, 127]$, or $[0, 255]$ ranges before registration.

Algorithm 1 summarizes the main steps of the method we applied.

3 Evaluation of the registration method

It is necessary to measure the degree of alignment in order to determine whether a given registration technique is adequate for a given problem. The alignment need not be perfect, but the error must be below a certain threshold. The similarity

Algorithm 1: Registration algorithm

Input: Two 3D images A and B with known dimensions and sample spacing
Output: Rigid-body transformation $optT$ that maximizes the mutual information or the normalized mutual information of images A and $optT(B)$

begin
1 scale intensity values of both images to be in $[0, 127]$;
2 generate A_l and B_l, the multiresolution Laplacian pyramid representation of the images ($l = 0, \ldots, L$);
3 let T be the identity transformation;
4 $optT = T$;
5 **for** *each pyramid level l from coarsest to finest* **do**
6 $optI = MI(A_l, optT(B_l))$;
7 **repeat**
8 $T = optT$;
9 make a change to T (Powell's method);
10 $m = MI(A_l, T(B_l))$;
11 **if** $m > optI$ **then**
12 $optI = m$;
13 $optT = T$;
endif
until $optT$ *was not changed*;
endfor
end

measure cannot be used to judge this, since it is not guaranteed that it reaches its global optimum at perfect alignment. An other method, visual inspection plays an important role. When a suitable interactive image viewing software is available, the human visual system can detect errors greater than 2 mm for CT to MR, and 4 mm for PET to MR registration [5, 15]. Although visual inspection is always necessary, since the automatic methods occassionally might fall into a nonglobal optimum producing a bad result without any warnings, a more accurate evulation procedure is necessary. An overview of such procedures can be found in [7].

To evaluate our registration method, we joined the Retrospective Registration Evaluation Project of Vanderbilt University, USA in 1999 [13]. The objective of that project was to perform blinded evaluation of retrospective image registration techniques using a prospective, marker-based registration method as a gold standard. A gold standard is a system whose accuracy is known to be high. A fiducial marker system can serve as an excellent gold standard for rigid registration, since some of these systems can provide submillimetric accuracy. The primary disadvantage is the high invasiveness i.e., bone-implanted markers [9]. In order to ensure blindness, all retrospective registrations were performed by participants who had no knowledge of the gold-standard until after their results had been submitted.

Image volumes of three modalities: X-ray computed tomography (CT), magnetic resonance (MR), and positron emission tomography (PET) were obtained from patients undergoing neurosurgery at Vanderbilt University Medical Center, on whom bone-implanted markers were mounted. These volumes had all traces of the markers removed and were provided to project collaborators outside Vanderbilt, who then performed registration on the volumes. The investigators communicated their results to Vanderbilt, where the accuracy of each registration was evaluated.

Two registration tasks were evaluated: CT to MR and PET to MR, and these tasks were broken into subtasks according to the type of MR and to whether or not the MR image was corrected (rectified) for geometrical distortion [3]. The image data set of nine patients were used, seven of which contained both CT and MR, and seven with both PET and MR.

The CT volumes have a resolution of 512 pixels in the x and y directions, and have between 28 and 34 slices in the z direction. The voxel size is 0.65 mm in x and y, and 4.0 mm in z. The MR volumes have a resolution of 256 pixels in the x and y directions, and have 20 to 26 slices. The voxel size is between 1.25 and 1.28 mm in the x and y directions, and 4.0 mm in z. The PET volumes have 15 slices with a resolution of 128 pixels in the x and y directions. The voxel size is 2.59 mm in x and y, and 8.0 mm in z.

At Vanderbilt, in collaboration with a neurological and a neurosurgical expert, a set of VOIs (Volume of Interest) representing areas of neurological and/or surgical interest was manually segmented within one of the MR image volumes for each patient. An estimate of the accuracy of the retrospective registration at the position of each VOI is computed as follows. The centroid pixel of the VOI is found, and its position is converted from voxel index to a millimetric position $\mathbf{c}$ in the reference volume using the known size for the image volume. Let T_G denote the gold-standard rigid-body transformation, and T_R the result of the retrospective registration algorithm. The point $\mathbf{c}'$ in the floating image is defined so that $\mathbf{c}$ is the mapping of $\mathbf{c}'$ under the gold-standard transformation,

$$c = T_G(c').$$

Thus,

$$c' = {T_G}^{-1}(c).$$

The point $\mathbf{c}''$ in the reference image is defined as the mapping of $\mathbf{c}'$ under the retrospective transformation,

$$c'' = T_R(c').$$

The error of the retrospective registration at the anatomical position of the VOI is defined as the Euclidean distance between the registered target position of the retrospective method and that of the gold standard, $||c'' - c||$.

4 Results

The results of the project were published in [13] and [14]. Since we joined the project later, our results were not included in those papers. Here we compare our

results against those evaluated earlier.

Ten groups of investigators applied 14 techniques to solve the registration tasks. The techniques were divided into two groups. Any technique which performs registration by making use of a relationship between voxel intensities within the images is referred to as volume based, and any technique which works by minimizing a distance measure between two corresponding surfaces in the images to be matched is referred to as surface based. Six of the 14 techniques were volume based and eight were surface based. Our methods can be classified as volume based ones.

Before the evaluation of our results, we visually inspected the quality of registration. When the normalized version of the mutual information was used, all registration results were visually acceptable. In case of mutual information, for the CT to MR task, all 41 results were visually acceptable. In case of PET to MR, for five image pairs the results of registration was visually misregistered. These pairs were PET to MR PD, MR T1, MR PD rectified of Patient 6 and PET to MR T1, and MR T2 rectified of Patient 8. The other 30 results were visually acceptable. In spite of these clear misregistrations, all results were submitted for evaluation to Vanderbilt University.

Table 1 shows the statistics of registration errors for the groups of algorithms and the rankings of our methods out of the 16 competing methods.

Modality	Surface based mean error (std.dev.)	Volume based mean error (std.dev.)	Our MI mean error (ranking)	Our NMI mean error (ranking)
CT-T1	5.7 (7.8)	2.9 (2.4)	1.6 (#2)	2.3 (#7)
CT-PD	5.8 (8.0)	2.9 (2.5)	2.2 (#2)	1.8 (#1)
CT-T2	6.3 (7.9)	2.4 (1.4)	2.0 (#5)	2.0 (#3)
CT-T1 rect.	6.1 (8.3)	2.0 (2.5)	1.7 (#5)	2.2 (#7)
CT-T2 rect.	5.7 (7.8)	1.8 (2.0)	1.4 (#3)	2.3 (#7)
CT-PD rect.	6.1 (7.6)	2.1 (1.6)	1.7 (#4)	2.4 (#7)
PET-T1	3.9 (2.0)	3.5 (2.1)	5.3 (#9)	3.0 (#2)
PET-T2	4.4 (2.1)	3.6 (1.9)	3.8 (#7)	3.5 (#4)
PET-PD	4.3 (2.6)	4.0 (2.7)	4.4 (#7)	4.2 (#10)
PET-T1 rect.	3.9 (2.3)	2.7 (1.4)	3.8 (#12)	2.7 (#3)
PET-T2 rect.	3.9 (2.0)	3.5 (1.7)	3.9 (#10)	3.3 (#5)
PET-PD rect.	3.9 (2.3)	3.5 (2.4)	4.8 (#10)	3.0 (#2)

Table 1: Mean and standard deviation of registration errors. Note that the ranking of our methods is based on the median errors of the registration methods, as it is published in [13].

5 Discussion

The results show that in case of CT to MR registration task, both of our methods produce acceptable results. For PET to MR problems, the MI method tends to fail (five failures out of 35 cases), and produces average results. The NMI method gives stable results and ranks high among the competing algorithms.

The running time was about 30-120 seconds on a 800 Mhz Pentium-III PC. More detailed results of the evaluation of our methods can be found at `http://www.vuse.vanderbilt.edu/~images/registration`.

6 Conclusion

We presented a registration algorithm, which can be successfully used to align 3D medical images from different imaging modalities. The algorithm is fully automatic, needs no user interaction. However, before using the optimal transformation determined by the algorithm, it is necessary to visually inspect it to sort out possible misregistrations.

Acknowledgements

The images and the standard transformations were provided as part of the project, "Evaluation of Retrospective Image Registration", National Institutes of Health, Project Number 1 R01 NS33926-01, Principal Investigator, J. Michael Fitzpatrick, Vanderbilt University, Nashville, TN.

References

[1] Brown, L.G.: A survey of image registration techniques. ACM Computing Surveys **24** (1992) 325–376

[2] Burt, P.J., Adelson, E.H.: The Laplacian Pyramid as a Compact Code. IEEE Trans. on Communications **31** (1983) 532–540

[3] Chang, H., Fitzpatrick, J.M.: A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities. IEEE Transactions on Medical Imaging **11** (1992) 319–329

[4] Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., Marchal, G.: Automated multi-modality image registration based on information theory. In Proceedings of Information Processing in Medical Imaging (1995) 263–274

[5] Fitzpatrick, J.M., Hill, D.L.G., Shyr, Y., West, J.B., Studholme, C., Maurer, C.R.Jr.: Visual assesment of the accuracy of retrospective registration of MR and CT images of the brain. IEEE Transactions on Medical Imaging **17** (1998) 571–585

[6] Foley, J.D., van Dam, A., Feiner, S.K., Hughes, J.F.: Computer Graphics — Principles and Practice. Addison–Wesley Publishing Company, Reading, Massachusetts (1991)

[7] Hajnal, J.V., Hill, D.L.G., Hawkes, D.J. (eds.): Medical Image Registration. CRC Press (2001)

[8] Maintz, J.B.A., Viergever, M.A.: A survey of medical image registration. Medical Image Analysis **2** (1998) 1–36

[9] Maurer, C.R., Fitzpatrick, J.M., Wang, M.Y., Galloway, R.L., Maciunas, R.J., Allen, G.S.: Registration of head volume images using implantable fiducial markers. IEEE Trans. on Medical Imaging **16** (1997) 447–462

[10] Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T.: Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press, second edition (1992)

[11] Studholme, C., Hill, D.L.G., Hawkes, D.J.: An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition, **32** (1999) 71–86

[12] Wells, W.M.III, Viola, P., Kikinis, R.: Multi-modal volume registration by maximization of mutual information. In Medical Robotics and Computer assisted Surgery, Wiley-Liss, New York (1995) 155–162

[13] West, J.B., Fitzpatrick, J.M., et al.: Comparison and evaluation of retrospective intermodality brain image registration techniques. Journal of Computer Assisted Tomography **21** (1997) 554–566

[14] West, J.B., Fitzpatrick, J.M., et al.: Retrospective Intermodality Registration Techniques for Images of the Head: Surface-Based Versus Volume-Based. IEEE Trans. on Medical Imaging **18** (1999) 144–150

[15] Wong, J.C.H., Studholme, C., Hawkes, D.J., Maisey, M.N.: Evaluation of the limits of visual detection of image misregistration in a brain fluorine-18 fluorodeoxyglucose PET-MRI study. European Journal of Nuclear Medicine **24** (1997) 642–650