

Efficient Image Registration by Decoupled Parameter Estimation Using Gradient-Based Techniques and Mutual Information

Ramtin Shams*, Rodney A. Kennedy* and Parastoo Sadeghi*

*Research School of Information Sciences and Engineering (RSISE)

The Australian National University

Canberra ACT 0200 Australia

Abstract—We present an efficient and accurate method for similarity (rigid+scale) registration of 3D images by decoupled estimation of transformation parameters. The estimated parameters are used to initialize the search for a more accurate alignment by optimizing a mutual information-based cost function. We use computationally inexpensive methods based on image gradients to obtain the original estimates. Our method improves the efficiency of the overall registration task by allowing the more expensive optimization step to converge quickly. It also makes the registration more robust and less sensitive to the shape of the cost function by reducing the chance of the optimization algorithm being caught by a local minimum. Our experiments demonstrate up to 10-fold increase in computational efficiency and significant reduction in misalignment.

I. INTRODUCTION

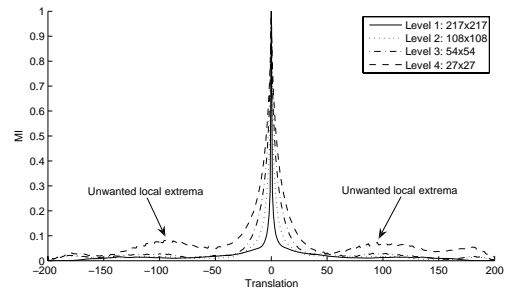
A. Background

Images obtained by various medical imaging techniques, often depict information of complementary nature. In clinical applications it is often desirable to combine useful data from different image modalities to assist in diagnosis, treatment planning, and evaluation of surgical and therapeutical procedures. To this end, a registration and a data fusion step are required.

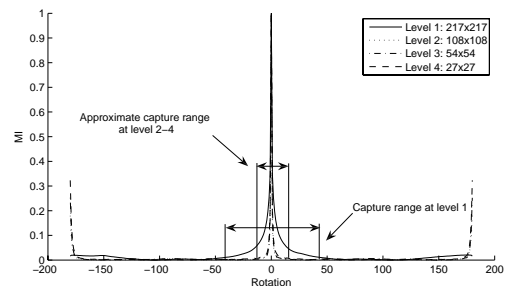
In order to register images automatically (without relying on landmarks or features), one needs a global means of measuring the similarity of the images; in other words, a cost function must be defined which attains its minimum, where the images are perfectly aligned and increases as the images move farther away. Several such cost functions have been studied in the literature such as cross correlation, variance of intensity ratios, histogram clustering, Taylor expansion of grey values, zero crossings in difference images, and finally mutual information (for a comprehensive survey refer to [1]). Since their introduction in the mid 90's by Collignon *et al.* [2] and Viola *et al.* [3], mutual information (MI)-based registration has received much attention in the literature [4] and has become the cost function of choice for registration of multi-modal images. This can be attributed to the favorable characteristics of the cost function that yields well to optimization algorithms and the accuracy with which it can perform the registration task, reaching that of a marker-based gold standard.

Given a suitable cost function such as MI¹, the registration problem reduces to an N -dimensional optimization problem, where N is the number of transformation parameters. Several optimization algorithms such as Powell, simplex, gradient descent, quasi-Newton and Levenberg-Marquardt [5] have been used for this purpose (refer to [6] for a comparison).

¹Minimization and maximization are trivially related. We refer to finding minimum anywhere we discuss optimization even though for certain functions, such as MI, the actual alignment is where the function is maximum.



(a) Translation capture range improves at first but unwanted extrema is introduced at lower resolutions.



(b) Rotation capture range worsens at lower resolutions.

Fig. 1. Typical MI functions for multiple resolutions of an MR image of the brain.

B. Motivation

The computational cost of the optimization process is indeed dominated by the cost function itself. As such, it is desirable to reduce the number of times the cost function has to be evaluated. The number of iterations and robustness of the registration task depend largely on the starting parameters used to initialize the optimization process. In the absence of any direction to choose the initialization parameters, one would use the parameter space's origin as the starting position. We refer to this approach as 'blind optimization'. This may not be a bad choice, where the misalignment between the images is small. However, with blind optimization the process takes longer for greater misalignments between the images. The chance of misregistration also increases, as the optimization process may be driven towards a local minimum along the path to the global minimum. The efficiency, robustness and performance of the optimization process can be improved by acquiring an initial estimate of the transformation parameters to guide the algorithm from an initial position close to the optimal transformation.

Despite favorable characteristics, MI functions typically suffer from a limited capture range. The optimization algorithm fails to converge, if it is initialized outside the capture range. Multi-resolution methods can extend the capture range of translation parameters (Fig. 1(a)), but are less effective for the rotation or scale and sometimes deteriorate the capture range as shown in Fig. 1(b). They can also introduce additional local extrema at lower resolutions (Fig. 1(a)).

This suggests that a method than can estimate transformation parameters has an advantage over multi-resolution, as it can improve the performance of the registration task by starting the optimization within the capture range of the cost function and without affecting the shape of the cost function or introducing additional extrema.

C. Contributions and Approach

The main contributions of the paper are as follows:

- 1) We propose a number of gradient-based methods to decouple estimating scale, rotation and translation parameters which reduce the computational complexity of parameter estimations.
- 2) We proposing a modification to Powell optimization algorithm to improve convergence of the algorithm.
- 3) We improve the efficiency of the MI-based registration by an order of magnitude through our guided optimization method.
- 4) We improve the robustness of the MI-based registration by reducing the chances of optimization algorithm being caught by local minima.

We use center of mass (centroid) and principal axes [8] of image gradients to obtain an initial estimate for the transformation parameters that register the images. Unlike [8] we prefer image gradients over intensities, due to the fact that a one-to-one mapping between intensities does not exist among multi-modal 3D medical images. In contrast, shape of the subject is maintained across different modalities of medical images [9]. This suggests that gradient images, which emphasize high frequency components of the image (e.g., corners and edges), are more suitable for our purpose.

We decouple parameters by estimating scale and rotation independently and then calculating translation using rotation and scale parameters. We use estimated parameters to initialize an optimization algorithm based on Powell's direction set method. Our guided optimization algorithm operates on an MI-based cost function. The guided optimization algorithm is specifically designed to converge quickly when initialized in close proximity of the optimal alignment. Experiments show that our method significantly improves performance and decreases computational cost by an order of magnitude compared to a conventional blind optimization approach.

We provide a brief overview of gradient image, centroid, principal axes, entropy, and mutual information in section II. Our registration and optimization methods are described in section III and the experimental results are discussed in section IV.

II. CONCEPTS

A. Gradient Image

Let \mathbf{I} be a 3D image and \mathbf{K}_x , \mathbf{K}_y and \mathbf{K}_z 3D differentiating kernels in x , y and z directions, the gradient magnitude is then calculated using (2).

$$\mathbf{G}_x = \mathbf{K}_x \otimes \mathbf{I}, \quad \mathbf{G}_y = \mathbf{K}_y \otimes \mathbf{I}, \quad \mathbf{G}_z = \mathbf{K}_z \otimes \mathbf{I}, \quad (1)$$

$$g_m(x, y, z) = \sqrt{g_x(x, y, z)^2 + g_y(x, y, z)^2 + g_z(x, y, z)^2}, \quad (2)$$

where \mathbf{G}_x , \mathbf{G}_y and \mathbf{G}_z are 3D image gradient arrays in x , y and z directions, g_m is the gradient magnitude at position (x, y, z) , and \otimes denotes 3D convolution.

B. Centroid and Principal Axes

A gradient map G is derived for 3D images using (2) and is defined as the set of vectors whose gradient magnitude is not less than a given threshold

$$G = \{\mathbf{x} \triangleq (x, y, z) | g_m(\mathbf{x}) \geq t\}, \quad (3)$$

where t is the threshold used to remove gradient vectors that do not contribute to image shape or are more sensitive to noise. We use the root mean square criterion as the gradient magnitude threshold

$$t = \left[\sum_{z=1}^l \sum_{y=1}^m \sum_{x=1}^n \frac{g_m(\mathbf{x})^2}{l \times m \times n} \right]^{\frac{1}{2}}, \quad (4)$$

where n , m and l are gradient array dimensions along x , y and z axes, respectively.

The center of the gradient map $\mathbf{x}_c = (x_c, y_c, z_c)$ is the average location of the points whose magnitude is above the selected threshold and is calculated as

$$x_c = \frac{1}{|G|} \sum_{\mathbf{x} \in G} x, \quad y_c = \frac{1}{|G|} \sum_{\mathbf{x} \in G} y, \quad z_c = \frac{1}{|G|} \sum_{\mathbf{x} \in G} z, \quad (5)$$

where $|G|$ is the number of gradient map elements.

The principal axes of the gradient image are the eigenvectors of the inertia matrix [8], [10]

$$I = \begin{bmatrix} I_{xx} & -I_{xy} & -I_{xz} \\ -I_{yx} & I_{yy} & -I_{yz} \\ -I_{zx} & -I_{zy} & I_{zz} \end{bmatrix}, \quad (6)$$

where

$$I_{xx} = \sum_{\mathbf{x} \in G} (y - y_c)^2 + (z - z_c)^2, \quad (7)$$

$$I_{yy} = \sum_{\mathbf{x} \in G} (z - z_c)^2 + (x - x_c)^2, \quad (8)$$

$$I_{zz} = \sum_{\mathbf{x} \in G} (x - x_c)^2 + (y - y_c)^2, \quad (9)$$

and

$$I_{xy} = \sum_{\mathbf{x} \in G} (x - x_c)(y - y_c), \quad (10)$$

$$I_{yz} = \sum_{\mathbf{x} \in G} (y - y_c)(z - z_c), \quad (11)$$

$$I_{zx} = \sum_{\mathbf{x} \in G} (z - z_c)(x - x_c). \quad (12)$$

C. Entropy and Mutual Information

Shannon entropy [11] is a measure of the average or expected information content of an event described by a random variable and is defined as

$$H(X) = \sum_{x \in X} p(x) \log \frac{1}{p(x)}, \quad (13)$$

where $p(\cdot)$ is the probability mass function (pmf) of the random variable X .

Mutual information of two random variables is the amount of information that each carries about the other and is calculated using the marginal and joint entropies of the random variables

$$I(X; Y) = H(X) - H(X|Y) = H(X) + H(Y) - H(X, Y),$$

where $H(X|Y)$ is the information content of random variable X if Y is known, $H(X, Y)$ is the joint entropy of the two random variables and is a measure of combined information of the two random

variables. $I(X; Y)$ can be thought of as the reduction in uncertainty of random variable X as a result of knowing Y . Mutual information is maximum, when there exists a one-to-one mapping between the random variables, and is minimum if the random variables are independent.

III. METHOD

The registration task is broken down into five steps: calculation of gradient images, decoupled estimation of scale, decoupled estimation of rotation, estimation of translation based on previously calculated scale and rotation parameters, and optimization.

A. Gradient Images

Gradient images are calculated by convolving images with the derivative of a Gaussian filter. The derivative of Gaussian has the desirable property of smoothing the image and being separable. By separating the kernel the image can be convolved with three $1 \times n$ kernels instead of a $n \times n \times n$ kernel and reduce the number of operations per voxel from $\mathcal{O}(n^3)$ to $\mathcal{O}(n)$ [12].

For our experiments we used $7 \times 7 \times 7$ x-,y- and z-derivative of Gaussian with standard deviation $\sigma = 1.5$ and separated the kernels to improve efficiency of gradient calculations.

B. Estimating Scale Parameter

Similarity transformation of a vector \mathbf{x} in 3D can be written as

$$\mathbf{x}' = s\mathbf{R}_z\mathbf{R}_y\mathbf{R}_x\mathbf{x} + \mathbf{t}, \quad (14)$$

where \mathbf{x}' is the transformed vector, s is the scale factor, \mathbf{t} is the translation vector and \mathbf{R}_x , \mathbf{R}_y and \mathbf{R}_z are the rotations matrices along x , y and z axes, respectively. The distance between two points \mathbf{x}_1 and \mathbf{x}_2 is defined as

$$\mathbf{d}(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{x}_1 - \mathbf{x}_2\|_2, \quad (15)$$

where $\|\cdot\|_2$ is the Euclidean ℓ_2 norm. It can be easily shown that

$$\mathbf{d}(\mathbf{x}'_1, \mathbf{x}'_2) = s\|\mathbf{R}_z\mathbf{R}_y\mathbf{R}_x(\mathbf{x}_1 - \mathbf{x}_2)\|_2 = s\mathbf{d}(\mathbf{x}_1, \mathbf{x}_2). \quad (16)$$

In other words, the distance between any two points under a similarity transformation remains constant as determined by the scale parameter. We need two points and their transformed locations to determine s . However, since we only have the centroid and its transformed location, we estimate the scale parameter from the average distance of the centroids from all other points in the gradient map.

$$\hat{s} = \frac{\sum_{\mathbf{x}' \in G'} \mathbf{d}(\mathbf{x}', \mathbf{x}'_c)}{\sum_{\mathbf{x} \in G} \mathbf{d}(\mathbf{x}, \mathbf{x}_c)}, \quad (17)$$

where \hat{s} is the estimated scale parameter and G and G' are the gradient maps of fixed (reference) and moving (floating) images, respectively.

C. Estimating Rotation Parameters

We estimate the rotation parameters using the principal axes method. The principal axes are the eigenvectors of the inertia matrix (6). We represent the three principal axes by finding the rotation parameters that align them with the basis axes. This can be done by equating the normalized eigenvector matrix \mathbf{E} with a 3D rotation matrix.

$$\begin{bmatrix} \cos \theta_y \cos \theta_z & \cdot & \cdot \\ -\cos \theta_y \sin \theta_z & \cdot & \cdot \\ \sin \theta_y & -\sin \theta_x \cos \theta_y & \cos \theta_x \cos \theta_y \end{bmatrix} = \mathbf{E}. \quad (18)$$

By solving (18) we find that²

$$\theta_y = \arcsin e_{31}, \quad \theta_z = \arcsin \frac{-e_{21}}{\cos \theta_y}, \quad \theta_x = \arcsin \frac{-e_{32}}{\cos \theta_y}, \quad (19)$$

where e_{ij} denotes the i^{th} and j^{th} element of the normalized eigenvector matrix. Once the rotation parameters are known for each gradient image, the estimated alignment between the images is found from the difference of corresponding parameters:

$$\hat{\theta}_x = \theta'_x - \theta_x, \quad \hat{\theta}_y = \theta'_y - \theta_y, \quad \hat{\theta}_z = \theta'_z - \theta_z. \quad (20)$$

D. Estimating Translation Parameters

Given the rotation and scale parameters, the translation can be found from (14) by replacing \mathbf{x} and \mathbf{x}' with the centroid of the fixed and moving images, respectively.

$$\hat{\mathbf{t}} = \mathbf{x}'_c - s\mathbf{R}_z\mathbf{R}_y\mathbf{R}_x\mathbf{x}_c. \quad (21)$$

E. Optimization

Powell's multi-dimensional direction set algorithm finds the minimum of the cost function by iteratively minimizing the function along a set of N directions, where N is the number of independent parameters of the cost function. One problem with the Powell algorithm is that from an observer's point of view who knows where the minimum is located, it appears to spend a lot of time, aimlessly iterating on and around the minimum. Obviously, the only way the optimization algorithm can satisfy itself that it has found the actual minimum is to spend enough time to check the surroundings.

Our implementation resembles the method described in [5] with one major difference that it takes advantage of the fact that it has been initialized not far from the minimum and as such, can quickly converge by refraining from checking the perimeter excessively. We define a minimum distance or resolution for the cost function. The cost function keeps track of each point in the N -dimensional space that it evaluates and will only evaluate a new point if falls outside all previously evaluated points by the specified minimum distance. The cost function quickly returns a previously calculated value for points inside the minimum distance of a previously evaluated point, which limits the resolution of the cost function. Obviously, this method is not suitable as a general purpose optimization tactic, and can only be used when the optimization algorithm can be properly initiated. Starting at an arbitrary position, this optimization algorithm is likely to be fatally paralyzed by the first local minimum it encounters.

IV. RESULTS

We tested our method on 3D simulated MR images of brain generated by Brain Web [13]. We experimented with various combinations of MRI modalities (T1, T2 and PD), artificially transformed with respect to one another. Fig. 2 shows an example of the images used in our experiments. Bilinear interpolation was used for transformations. We limited our experiments to transformations in the transversal plane. However, the method can be extended to out-of-plane transformations with three rotation, three translations, and a scale parameter.

Fig. 3 demonstrates that our method improves performance and efficiency of the registration tasks by giving an example for registration of T1 to T2 images. In this experiment, T2 was translated by $(-20, 20)$ mm and scaled by 0.9, the rotation parameter was varied from 0° to 55° in 5° increments. The efficiency of the registration improved on average by a factor of 9.6 (Fig. 3(a)) with the total execution time of around 25 seconds on a standard PC.

²The upper left elements of the rotation matrix do not contribute to the solution and have been removed due to space considerations.

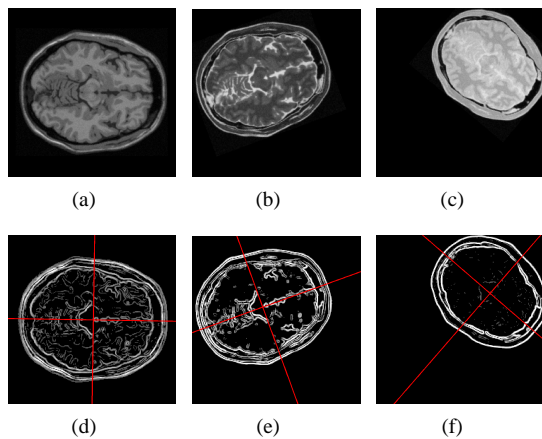


Fig. 2. A sample set of synthetic MR images of brain used in our experiments (a) MRI-T1: reference image. (b) MRI-T2: $t_x = -20$ mm, $t_y = 20$ mm, $\theta_z = 20^\circ$, $s = 0.9$. (c) MRI-PD: $t_x = t_y = 40$ mm, $\theta_z = -40^\circ$, $s = 0.8$ (d-f) Corresponding gradient images with the principal axes and the centroids overlaid.

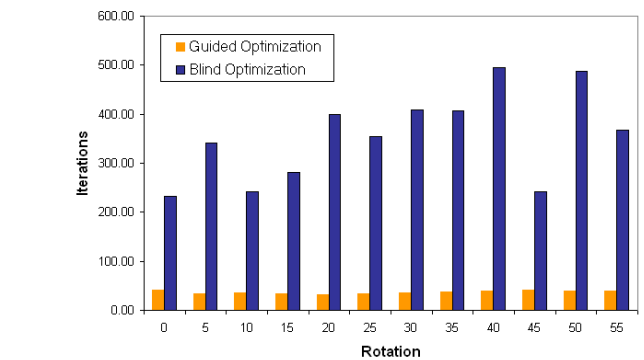
We calculated the registration error by transforming the centroid using the actual transformation parameters and under calculated transformation parameters. The registration error was then obtained from the Euclidian distance between the transformed centroids. The registration error for initially estimated parameters was around twice the voxel size. However, note that this represents an average of 3.5% error on each estimated parameter, which is close enough to the actual alignment for optimization purposes. Our guided optimization algorithm, starting off these initial estimates, is able to achieve accuracy well below the voxel size of 1 mm with an average of 0.25 mm. The registration error for the guided optimization remained almost constant as shown in Fig. 3(b). The blind optimization method slightly outperformed our lower resolution guided optimization for smaller transformation parameters but completely failed as we increased the transformation parameters.

Our method works well, since the estimated parameters are close to the actual alignment and as such, initialization of the optimization algorithm with these parameters (and with a lower resolution) results in accurate alignment as well as huge reduction in computational complexity. It should be noted that for blind optimization we used the standard Powell algorithm as described in [5], which does not impose an artificial limit on the resolution of the cost function.

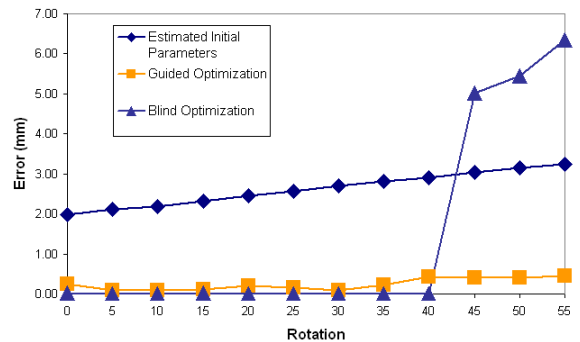
The results shown in Fig. 3 are indicative and consistent with several other experiments, which we have not included due to space limitations. We experimented by gradually increasing other transformation parameters and tested our method on other combinations of image modalities (T1 to PD and T2 to PD).

REFERENCES

- [1] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Med. Image Anal.*, vol. 2, no. 1, pp. 1–36, 1998.
- [2] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multimodality medical image registration using information theory," in *Proc. Int. Conf. Information Processing in Med. Imaging: Computational Imaging and Vision 3*, Apr. 1995, pp. 263–274.
- [3] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," in *Proc. Int. Conf. Computer Vision (ICCV)*, June 1995, pp. 16–23.
- [4] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: A survey," *IEEE Trans. on Med. Imaging*, vol. 22, no. 8, pp. 986–1004, Aug. 2003.



(a) Iterations vs. rotation (shorter bars indicate better efficiency)



(b) Registration error vs. rotation

Fig. 3. Superior robustness and improved efficiency of our method compared to the conventional method. (a) An order of magnitude improvement in efficiency (execution times are proportional to iterations). (b) The registration error for our method remains nearly constant as the rotation parameter increases, however the conventional method fails for rotations above 40° .

- [5] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, 2nd ed. Cambridge: Cambridge University Press, 1992.
- [6] F. Maes, D. Vandermeulen, and P. Suetens, "Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information," *Med. Image Anal.*, vol. 3, no. 4, pp. 373–386, 1999.
- [7] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imaging*, vol. 16, no. 2, pp. 187–198, Apr. 1997.
- [8] L. K. Arata, A. P. Dhawan, J. P. Broderick, M. F. Gaskil-Shibley, A. V. Levy, and N. D. Volkow, "Three-dimensional anatomical model-based segmentation of MR brain images through principal axes registration," *IEEE Trans. Biomedical Engineering*, vol. 42, no. 11, pp. 1069–1078, Nov. 1995.
- [9] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. on Med. Imaging*, vol. 19, no. 8, pp. 809–814, Aug. 2000.
- [10] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed. Reading, Mass : Addison-wesley Pub. Co, 2002.
- [11] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423/623–656, 1948.
- [12] M. Hopf and T. Ertl, "Accelerating 3d convolution using graphics hardware," in *Proc. Visualization*. IEEE, 1999, pp. 471–474.
- [13] *Brain Web*. <http://www.bic.mni.mcgill.ca/brainweb/>: Montreal Neurological Institute, McGill University, 2006.