

Robust Optimization Using Disturbance for Image Registration

Rui Gan and Albert C.S. Chung

Lo Kwee-Seong Medical Image Analysis Laboratory,
Department of Computer Science,
The Hong Kong University of Science and Technology, Hong Kong
{raygan, achung}@cs.ust.hk

Abstract. This paper exploits the different properties between the local neighborhood of global optimum and those of local optima in image registration optimization. Namely, a global optimum has a larger *capture neighborhood*, in which from any location a monotonic path exists to reach this optimum, than any other local optima. With these properties, we propose a simple and computationally efficient technique using transformation disturbance to assist an optimization algorithm to avoid local optima, and hence to achieve a robust optimization. We demonstrate our method on 3D rigid registrations by using mutual information as similarity measure, and we adopt quaternions to represent rotations for the purpose of the unique and order-independent expression. Randomized registration experiments on four clinical CT and MR-T1 datasets show that the proposed method consistently gives much higher success rates than the conventional multi-resolution mutual information based method. The accuracy of our method is also high.

1 Introduction

An important component in the medical imaging field is multi-modal image registration, which can integrate complementary image information acquired from different modalities. The task of image registration is to reliably identify a geometric transformation to accurately align two images. General promising results have shown that mutual information (MI) as a voxel intensity-based similarity measure is well-suited for multi-modal image registration [8, 14]. However, it has been suggested that the conventional mutual information based registration method can result in misalignment for some cases [9, 11]. One possible causation is that the mutual information based method can get trapped into local maxima during transformation optimization process. Most of the existed optimization methods, e.g. Powell's method, simplex methods, gradient descent and so on [12], are only suitable for local optimization. Several attempts have been made to improve the optimization performance, such as multi-resolution approaches [2], extension and combination of different optimization methods [10, 5]. However there is still no guarantee to find the global solution in general. On the other hand, global stochastic optimization methods, e.g. Simulated Annealing [4], need

to take a large number of iterations to converge, which limit their application in image registration.

In this paper, in an attempt to obtain a robust optimization result, a simple and computationally efficient technique is proposed to assist an optimization method to avoid local optima. The technique is based on the different properties between the local neighborhood of global optimum and those of local optima. Specifically, the global optimum has a larger *capture neighborhood*, in which from any location a monotonic path exists to reach this optimum, than any other local optima. When an optimization procedure gets trapped into a local optimum, a proper disturbance on the obtained transformation can pull it out of the undesirable location and consequently provide a further chance for a following up optimization to achieve the global optimum. The proposed method is demonstrated on 3D rigid registrations. We adopt quaternions to represent rotations for the purpose of the unique and order-independent expression. Based on the randomized registration experiments on four clinical 3D CT and MR-T1 image volumes, it is demonstrated that the new method consistently gives much higher successful registration rates than the conventional multi-resolution mutual information based method. The results also implies that our method can obtain acceptably high registration accuracy.

2 Methods

2.1 Mutual Information as Similarity Measure

Mutual information (MI) is an useful concept from information theory [3] and measures the amount of information shared between two random variables. Specifically, mutual information quantifies the Kullback-Leibler distance [7] between the joint distribution of two random variables, A and B, and the product of their marginal distributions, that is

$$MI(A, B) = \sum_{a,b} p(a, b) \log \frac{p(a, b)}{p(a) \cdot p(b)}, \quad (1)$$

where $p(a, b)$ is the joint distribution of A and B, and $p(a)$ and $p(b)$ are the individual marginal distributions respectively.

Mutual information was proposed independently as similarity measure for 3D rigid registration of medical images by Wells et al. [14] and Maes et al. [8]. To utilize mutual information, the intensity values of the corresponding voxel pair in the two images to be registered are considered as random variables, and the joint and marginal distributions can be the normalization of the joint and marginal histograms of the sampling set. The mutual information of two images measures the amount of information conveyed by one image that is shared by the other image, and it is assumed to be maximum when the images are aligned. Therefore, the mutual information based registration method identifies a geometrical transformation $\hat{\mathbf{T}}$ as follows,

$$\hat{\mathbf{T}} = \arg \max_{\mathbf{T}} MI(\mathbf{T}) = \arg \max_{\mathbf{T}} MI(\mathbf{T}(I^f), I^r),$$

where I^r and I^f are the intensity domains for the *reference* and *floating* images respectively.

2.2 Robust Optimization Using Transformation Disturbance

General promising results have shown that the mutual information based registration method works well on multi-modal images [8, 14]. However, it has been suggested that it can result in misalignment for some cases [9, 11]. One possible causation is that mutual information based method can get trapped into local maxima during transformation optimization process. This motivates us to propose a technique to assist an optimization method to avoid those local maxima.

From any initial transformation \mathbf{T}_0 , it is obvious that there exists a *monotonic path*, $\overrightarrow{\mathbf{T}_0\mathbf{T}_1 \cdots \mathbf{T}_l}$, to a local (or global) maximum \mathbf{T}_l , such that $MI(\mathbf{T}_0) \leq MI(\mathbf{T}_1) \leq \cdots \leq MI(\mathbf{T}_l)$. Then, the *capture neighborhood* of an arbitrary local (or global) maximum \mathbf{T}_{l_i} , $\mathcal{N}_{\mathbf{T}_{l_i}}$, can be defined as a maximum local neighborhood in which any transformation has such a monotonic path to reach \mathbf{T}_{l_i} . Alternatively, $\mathcal{N}_{\mathbf{T}_{l_i}}$ can be expressed as follows,

$$\mathcal{N}_{\mathbf{T}_{l_i}} = \{ \mathbf{T} \mid \exists \overrightarrow{\mathbf{T} \cdots \mathbf{T}_{l_i}}, \text{ such that } MI(\mathbf{T}) \leq \cdots \leq MI(\mathbf{T}_{l_i}) \},$$

and its radius is given by

$$R_{\mathbf{T}_{l_i}} = \max_{\mathbf{T} \in \mathcal{N}_{\mathbf{T}_{l_i}}} d(\mathbf{T}, \mathbf{T}_{l_i}),$$

with $d(\cdot)$ being the distance of two locations. With this definition, the whole transformation space can be treated as the union of all such capture neighborhoods, i.e. $\bigcup_i \mathcal{N}_{\mathbf{T}_{l_i}}$.

The foundation of the proposed method is an observation that the radius of the capture neighborhood of an arbitrary local maximum \mathbf{T}_{l_i} is much smaller than that of the global maximum \mathbf{T}_o , i.e. $R_{\mathbf{T}_{l_i}} \ll R_{\mathbf{T}_o}, \forall \mathbf{T}_{l_i} \neq \mathbf{T}_o$. Thus, if an optimization procedure gets trapped into $\mathbf{T}_{l_i} \neq \mathbf{T}_o$, a relatively small offset added on the obtained transformation can pull it out of $\mathcal{N}_{\mathbf{T}_{l_i}}$, and then a further chance for a following up optimization to converge to \mathbf{T}_o is given. On the other hand, when an optimization procedure achieves \mathbf{T}_o , the same offset will not pull it out of $\mathcal{N}_{\mathbf{T}_o}$, and thus a following up optimization process still achieves \mathbf{T}_o . Practically, when an optimization procedure converges, in order to pull it out of $\mathcal{N}_{\mathbf{T}_{l_i}}$ ($\mathbf{T}_{l_i} \neq \mathbf{T}_o$), or to validate whether it reaches \mathbf{T}_o or not, we propose to randomly add an reasonable disturbance (e.g. larger than $R_{\mathbf{T}_{l_i}}$ and much less than $R_{\mathbf{T}_o}$) on the obtained transformation and resume the optimization process. (For the determination of the range of disturbance, please refer to Section 3.) As a consequence, the potentiality of obtaining a robust optimization result increases. This process iterates until the change of obtained transformations at two successive iterations becomes sufficiently small.

2.3 Rotation Representation

The technique proposed in Section 2.2 is demonstrated on 3D rigid registration. We use quaternion algebra to represent 3D rotations in rigid transformations.

This is because any rotation can be uniquely represented by quaternion as a complex number:

$$\begin{aligned} q &= a + ib + jc + kd \\ &= \cos\left(\frac{\alpha}{2}\right) + i\left(x \sin\left(\frac{\alpha}{2}\right)\right) + j\left(y \sin\left(\frac{\alpha}{2}\right)\right) + k\left(z \sin\left(\frac{\alpha}{2}\right)\right), \end{aligned} \quad (2)$$

where $(x, y, z)^T$ is an unit vector representing the axis of rotation, α is the angle of rotation. Note that other rotation representations may have ambiguity in some cases, e.g. Euler angle representation has ambiguity when the attitude (or elevation) is equal to $\pm\frac{\pi}{2}$. Moreover, quaternion has another advantage of encoding rotation in an order-independent manner.

In addition, in Equation 2, we have $a^2 + b^2 + c^2 + d^2 = 1$. In order to decrease the degree of freedom of quaternion for optimization, Equation 2 is revised as follows,

$$\begin{aligned} q &= 1 + i\tilde{b} + j\tilde{c} + k\tilde{d} = 1 + i\frac{b}{a} + j\frac{c}{a} + k\frac{d}{a} \\ &= 1 + i\left(x \tan\left(\frac{\alpha}{2}\right)\right) + j\left(y \tan\left(\frac{\alpha}{2}\right)\right) + k\left(z \tan\left(\frac{\alpha}{2}\right)\right). \end{aligned}$$

Furthermore, during optimizations of brain image registrations, large rotations from optimal position, e.g. $\geq \frac{\pi}{2}$ and $\leq -\frac{\pi}{2}$, are fatal and most likely to invert brains in floating images. To avoid this, a hard constraint, namely $\tilde{b}^2 + \tilde{c}^2 + \tilde{d}^2 < 1$, may be added to limit the search space of rotation angle to $\alpha \in (-\frac{\pi}{2}, \frac{\pi}{2})$.

3 Implementation Details

In our implementation, in order to accelerate the registration process, a multi-resolution approach based on the Gaussian Pyramid representation [1, 14, 2] is exploited. Four resolution levels are used and the definition of resolution levels in the Gaussian Pyramid representation follows the same convention as in [1], i.e. Level 0 image represents the highest and original resolution and Level 3 image represents the lowest resolution. Smoothing is performed via the binomial filter with coefficients [1, 4, 6, 4, 1] [14]. For the ease of implementation, all voxels in the downsampled floating volumes are used at Levels 1 – 3. At Level 0, 1/4 (25%) of all voxels are sampled (one voxel randomly picked from every 2×2 matrix in each slice). To construct the joint and marginal histograms, the image intensity values are linearly scaled to 64 bins, which have been commonly used in the mutual information based registrations. For optimization at each resolution, we use the Powell's direction set method [12] with Brent's 1D line minimization, where the fractional precision convergence parameters are set to 10^{-4} and 10^{-3} respectively.

The transformation disturbance across progressive optimizations as discussed in Section 2.2 is only performed at Level 3 for reasons of speed, since the computational burden at higher levels can be more significant. On the other hand,

it is empirically observed that, although from large initial misalignments, the majority of optimization results can converge close to the ground truth at Level 3, with the help of transformation disturbance. (Please refer to Section 4.2 for details.) For an optimization at Level 3, a disturbance, which consists of 3 disturbances for translational parameters and 3 disturbances for quaternion parameters, is randomly generated based on Gaussian distributions with zero means. To determine the standard deviations of these Gaussian distributions, a set of experiments with different standard deviation values have been performed. Empirical results suggest that a relative high success rate is given by the following set of standard deviation values: for translation, they are equal to $1/16$ of image dimension sizes (in millimeter), and for quaternion, they are equal to $\tan(\pi/16)$. The convergence criterion is that, at two successive iterations, the change of obtained transformations for an individual degree of freedom is less than $1/5$ of the corresponding standard deviation.

4 Experimental Results

4.1 Image Datasets and Ground Truth

In the experiments described below, we used a set of real CT – T1 data obtained from the Retrospective Image Registration Evaluation (RIRE) project¹. Note that all the T1 images have been rectified for intensity inhomogeneity and scaling. In general, the size of a CT image volume is $512 \times 512 \times 34$ voxels and the voxel size is $0.65 \times 0.65 \times 4\text{mm}^3$, and a T1 image contains $256 \times 256 \times 26$ voxels of dimensions $1.25 \times 1.25 \times 4\text{mm}^3$.

With regard to the data, we determined the “ground truth” for registration experiments as follows. First, the multi-resolution mutual information based and normalized mutual information (NMI) [13] based methods were used to register the image pairs. The evaluations of accuracy were obtained from the RIRE project. By examining the median errors, four datasets (Datasets pt-001, pt-003, pt-005 and pt-007) with less than 1mm registration error were selected and used in the experiments. Then the corresponding optimal transformations, whose median errors were 0.5077 (for pt-001), 0.7200 (for pt-003), 0.7807 (for pt-005) and 0.6179 (for pt-007) respectively, were used as the ground truth registrations.

4.2 Justification and Determination of Disturbance

In order to justify the capability of the proposed transformation disturbance on helping optimization to avoid local maxima, especially at Level 3, randomized experiments were performed for the mutual information based registration method with disturbance and quaternion (hereafter referred to as MI-d), and also for the conventional mutual information based registration method (hereafter referred

¹ Images were provided as part of the project, “Evaluation of Retrospective Image Registration”, National Institutes of Health, Project Number 8R01EB002124-03, Principle Investigator, J. Michael Fitzpatrick, Vanderbilt University, Nashville, TN.

to as MI) as a comparison. The testing image pair was the aforementioned pt-001 CT – T1 dataset and the experiment for either method took 100 trials. At each trial, the ground truth registration parameters of pt-001 were perturbed by six uniformly distributed random offsets for all translational and rotational axes. The perturbed parameters were then treated as the starting alignment. In order to show the high optimization capability of MI-d with respect to initial alignment, random offsets for X and Y translational axes were drawn between around $[-150, 150]$ mm and those for Z translational axis were drawn between around $[-70, 70]$ mm. (Note that these ranges were set so that two brains in CT and T1 images have at least 10% overlapping region.) While, random offsets for each rotational axis were respectively drawn between $[-0.52, 0.52]$ radians, i.e. $[-30, 30]$ degrees. (The perturbed rotations were first converted to quaternions prior to MI-d.) As a fair comparison, the same set of randomized starting alignments was used for MI and MI-d.

In addition, as mentioned in Section 3, in order to determine a suitable setting for generating random disturbances, we have tested several different sets of standard deviation values for the Gaussian distributions: $1/n$ of image dimension sizes (in millimeter) for translation, and $\tan(\pi/n)$ for quaternion, $n \in \{8, 16, 32, 64\}$. Then, we selected the one set, which produced the most *successful* results, for further experiments. Since the registrations were only performed to the downsampled image pair at Level 3, an optimization result was judged to be *successful* when the individual translational errors w.r.t. the ground truth were less than 10mm and the individual rotational errors (measured by Euler angles) were less than 5° . Such thresholds were selected because, based on our empirical experiences, a starting alignment within them definitely can be fine-tuned to converge to the ground truth.

Table 1 presents the number of successful optimization results for MI and MI-d with different sets of standard deviations (i.e. MI-d-8, MI-d-16, MI-d-32, MI-d-64) on pt-001 dataset at Level 3. In the ”# Success” column, it is suggested that, as compared with MI, MI-d with different sets of standard deviations can give much more successful optimization results at Level 3. It is also noted that, amongst the four different sets, MI-d-16 has the best performance and thus we determine to adopt it for further experiments. The average running time at each trail for each method is also listed in the table. Obviously, all MI-d methods take

Table 1. The number of successful optimization results and the average processing time (in seconds) of MI, MI-d with different sets of standard deviations on pt-001 CT – T1 dataset at Level 3

	# success	Time (sec)
MI	47	28
MI-d-8	90	227
MI-d-16	96	193
MI-d-32	83	311
MI-d-64	81	321

relatively longer time to converge than MI. However, as we do not perform transformation disturbance at higher levels (i.e. Levels 0 – 2), where the computational burden of conventional MI is already significant, the difference between the overall processing time of MI-d and MI will be much less than that at Level 3.

4.3 Performance Comparisons on Robustness and Accuracy

In this section, we further study and compare the registration performance (w.r.t. robustness and accuracy) of MI-d-16 and MI. A series of similar randomized experiments as described in Section 4.2 was performed on the aforementioned four clinical CT – T1 image pairs (i.e. pt-001, pt-003, pt-005 and pt-007) with all four resolution levels. Therefore, for either method, an experiment with 100 trials was performed on each dataset with a set of randomized starting alignments generated as described in Section 4.2.

To evaluate each derived registration with respect to the corresponding ground truth registration, similar to [8, 6], a tight bounding box was fitted around the brain for each T1 images. For each of the eight corner points, the Euclidean distance between the ground truth position and the position transformed by our solution was computed. The median value of the eight distances was then taken for assessing registration success. A registration was judged to be successful if the median error was smaller than or equal to 4mm, which was the largest voxel dimension of the CT – T1 image pair; otherwise, it was considered a misregistration.

Table 2 lists the success rates for MI and MI-d-16 for all testing image pairs (pt-001, pt-003, pt-005 and pt-007), together with the means and standard deviations of the median errors (in millimeters) for the *successful* registrations. It is shown in the table that MI-d-16 consistently gives much higher success rates as compared with MI. For registration accuracy, it is observed that the median errors of the successful registrations for MI-d-16 are comparable to those for MI, and are acceptably low.

Table 2. The success rates of MI and MI-d-16, and the means and standard deviations of the median errors (in millimeters) for different testing image pairs

	MI		MI-d-16	
	success%	mean \pm sd	success%	mean \pm sd
pt-001	51%	0.5295 \pm 0.0459	96%	0.5438 \pm 0.0530
pt-003	45%	0.1983 \pm 0.1103	99%	0.2134 \pm 0.2854
pt-005	40%	0.5182 \pm 0.7830	93%	0.3959 \pm 0.6452
pt-007	44%	1.6842 \pm 0.4790	92%	1.6959 \pm 0.4138

5 Conclusion and Discussion

To conclude, this paper has proposed a simple and computationally efficient technique based on transformation disturbance to assist an optimization method

to avoid local optima, and hence to achieve a robust optimization result. The new technique has been demonstrated on 3D rigid registrations, and quaternions have been adopted to represent rotations for the purpose of the unique and order-independent expression. Randomized registration experiments on four clinical CT and MR-T1 datasets have revealed that the success rates of our method are consistently much higher than those of the conventional multi-resolution mutual information based method. It has been also shown that the registration accuracy of the new method is acceptably high.

Finally, we would like to note that there is no practical guarantee for our method to achieve global maximum. However, empirical observations have shown that the optimization performance of our method is much better than the conventional multi-resolution approach. A theoretical justification of the work is desirable and remains a topic for future research. Furthermore, although the illustration and demonstration in this paper just concentrate on mutual information as similarity measure and the Powell's optimization method, the proposed technique is quite general and can be applied to other similarity measures (e.g. normalized mutual information and so on) and optimization methods (e.g. simplex methods and so on).

Acknowledgements

The authors would like to acknowledge the support from the Research Grants Council (RGC) of Hong Kong under Grant HKUST6155/03E.

References

1. P.J. Burt and E.H. Adelson. The Laplacian Pyramid as a Compact Image Code. *IEEE Trans. Comms.*, 31(4):532–540, 1983.
2. A.A. Cole-Rhodes, K.L. Johnson, and et al. Multiresolution Registration of Remote Sensing Imagery by Optimization of Mutual Information Using a Stochastic Gradient. *IEEE Trans. Image Processing*, 12(12):1495–1511, 2003.
3. T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.
4. S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721 – 741, 1984.
5. M. Jenkinson and S. Smith. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2):143 – 156, 2001.
6. Z.F. Knops, J.B.A. Maintz, M.A. Viergever, and J.P.W. Pluim. Normalized mutual information based registration using k-means clustering and shading correction. *Medical Image Analysis*, 2005 (To appear).
7. S. Kullback. *Information Theory and Statistics*. Dover Publications, Inc., 1968.
8. F. Maes, A. Collignon, and et al. Multimodality Image Registration by Maximization of Mutual Information. *IEEE Trans. Med. Img.*, 16(2):187–198, 1997.
9. G.P. Penney, J. Weese, and et al. A Comparison of Similarity Measures for Use in 2D-3D Medical Image Registration. *IEEE Trans. Med. Img.*, 17(4):586–595, 1998.

10. D. Plattard, M. Soret, J. Troccaz, P. Vassal, J.-Y. Giraud, G. Champleboux, X. Artignan, and M. Bolla. Patient set-up using portal images: 2d/2d image registration using mutual information. *Comput Aided Surgery*, 5(4):246 – 262, 2000.
11. 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. Med. Img.*, 19(8):809–814, 2000.
12. W.H. Press, S.A. Teukolsky, and et al. *Numerical Recipes in C, 2nd Edition*. Cambridge University Press, 1992.
13. C. Studholme, D.L.G. Hill, and D.J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32:71–86, 1999.
14. W.M. Wells, P. Viola, and et al. Multi-Modal Volume Registration by Maximization of Mutual Information. *Medical Image Analysis*, 1(1):35–51, 1996.