

# Multimodal Brain Image Registration Based on Wavelet Transform Using SAD and MI

Jue Wu<sup>1</sup> and Albert C.S. Chung<sup>2</sup>

<sup>1</sup> Bioengineering Program, School of Engineering, the Hong Kong University of Science and Technology, Hong Kong, China

<sup>2</sup> Department of Computer Science, the Hong Kong University of Science and Technology, Hong Kong, China

**Abstract.** The multiresolution approach is commonly used to speed up the mutual-information (MI) based registration process. Conventionally, a Gaussian pyramid is often used as a multiresolution representation. However, in multi-modal medical image registration, MI-based methods with Gaussian pyramid may suffer from the problem of short capture ranges especially at the lower resolution levels. This paper proposes a novel and straightforward multimodal image registration method based on wavelet representation, in which two matching criteria are used including sum of difference (SAD) for improving the registration robustness and MI for assuring the registration accuracy. Experimental results show that the proposed method obtains a longer capture range than the traditional MI-based Gaussian pyramid method meanwhile maintaining comparable accuracy.

## 1 Introduction

Image registration is a process of finding a transformation that maps one image onto another same or similar object by optimizing certain criterion. It is an essential step in medical image processing if clinicians require complementary information obtained from different images. Registration aims to fuse data about patients from more than one medical image so that doctors can acquire more comprehensive information related to pathogenesis.

Most of rigid registration methods require an iterative process to go from an initial estimate position to a final optimal one[1]. There are many factors that affect the process of registration. The capture range is one of them; it directly determines the success or failure of the alignment. Regarding intensity similarity measures, many local optima can trap the optimization method and cause it to stop at incorrect positions. Capture range can be defined as a range within which the starting algorithm is more likely to converge to the correct optimum[1]. In other words, if the starting point is not within the capture range, the registration algorithm is most likely to converge to a wrong optimum, thus leading to misregistration. Logically, the size of capture range is positively correlated to the success rate of a registration task. As such, the capture range is a very important property that influences the robustness of registration algorithms and

longer capture ranges are what registration methods pursue [2,3]. The size of a capture range depends on the features in the images and similarity metrics used to measure them. In this paper, the capture range is mathematically defined as the range in which the value of matching criterion monotonically decreases (or increases) when at a distance away from the maximal (or minimal) position.

Mutual information (MI) is one of the most popular matching criteria that are used in multi-modal image registration. Many researches have shown that MI produced satisfactory results in terms of accuracy [2,3,4]. Due to its high computational complexity, scientists have proposed the multiresolution scheme to accelerate MI-based registration. Though some researchers believe that a multiresolution scheme can also increase the capture range for there is less tendency to be trapped in local minima [2], our experiments show that the capture range is still not good enough especially in lower resolution registration. This is supported by the conclusion drawn in [5], i.e. the hope that a multiresolution approach to matching would be better equipped to avoid local optima seems unfounded. The statistical relation of image intensities that MI measures tends to decline when the image resolution decreases [6]. Thus, MI does not naturally extend to the multiresolution framework. In order to improve this situation, we propose the combination of the sum of absolute difference (SAD) and MI as similarity metrics rather than MI alone.

We do not use SAD and MI directly on the original image intensities. Instead we make use of the wavelet coefficients to calculate the two similarity measures. Scientists have tried to apply wavelet transform to image registration due to its inherent multiresolution characteristics and ability to preserve significant structural information of images. J.L. Moigne [7] registered remote sensed data using maxima of HL and LH wavelet coefficients. P. Moon *et al.* [8] looked into the applicability of the wavelet transform to digital stereo image matching. Turcujova *et al.* [9] tested various orthogonal and biorthogonal wavelets (they used LL coefficients) together with cross-correlation to register affine transformed images of objects at a distance. Liu *et al.* [10] suggested the application of the Gabor filter and the Gaussian model of registration. J. Zhen *et al.* [11] proposed a wavelet-based multi-scale block matching algorithm to register DSA images.

In this paper, we utilize the LL wavelet coefficients to register two multimodal images using the sum of absolute difference (SAD) at lower resolution levels and mutual information (MI) at higher resolution levels. The originality of this idea is to apply wavelet in the field of medical image registration and the combination of two different similarity measures. Experimental results show that our method has a longer capture range than a conventional Gaussian pyramid with MI at all levels.

## 2 Proposed Approach

### 2.1 Matching Criteria

In our algorithm, two similarity metrics are utilized, namely MI and SAD. They work on the LL wavelet coefficients at different resolutions.

Originating from information theory, MI is an entropy-based concept and denotes the amount of information that one variable can offer to the other. In terms of marginal distributions  $p(a)$  and  $p(b)$  for images  $A$  and  $B$  respectively and the joint distribution  $p(a, b)$ , MI can be defined as:

$$I(A, B) = \sum_{a,b} p(a, b) \log \frac{p(a, b)}{p(a)p(b)},$$

where  $a$  and  $b$  represent the intensity of image  $A$  and  $B$  respectively. MI measures the statistical dependence between the image intensities of corresponding voxels in both images, which is assumed to be maximal while SAD is assumed to be minimal if the images are geometrically aligned. At a low resolution level, the capture range of MI is not satisfactory and we find that combined with wavelet, SAD works better than MI with a Gaussian pyramid in terms of capture ranges. MI has a shorter capture range because after the MI value drops down from the maximum, it soon increases. This is due to the fact that joint entropy is small when images are registered and it increases considerably (thus MI decreases) when the floating image shifts away from the optimal transformation. When the brain of floating image maps mainly to the background of reference image, the joint entropy decreases. If the amount of decrement is greater than the decrement of the sum of two marginal entropies, MI increases ( $I(A, B) = H(A) + H(B) - H(A, B)$ ). This results in short capture ranges of MI at low resolution levels.

SAD is defined as:

$$SAD = \frac{1}{N} \sum_{\mathbf{x} \in D} |u(\mathbf{x}) - v(T(\mathbf{x}))|,$$

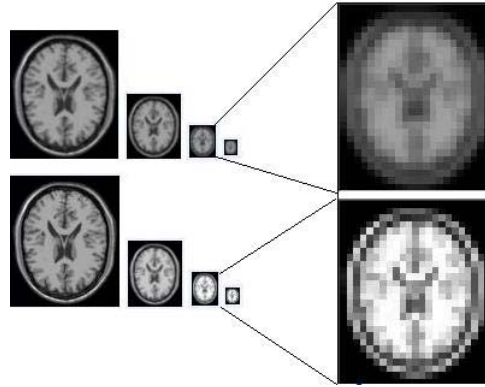
where  $\mathbf{x}$  and  $T$  refer to the position vector and the transformation respectively,  $u$  and  $v$  are the intensity functions of two images, and  $N$  is the number of pixels in the region of interest  $D$  (here refers to the overlapping part of the two images).

In a general way, SAD is not effective in multi-modal registration because even when perfectly registered, images from different modalities are different in intensity [2]. Therefore we adopt MI rather than SAD at higher resolution levels to ensure the accuracy of registration. However, SAD works well in lower resolution registration because in a low resolution scenario, all details tend to disappear and only global features are preserved, e.g., a brain image approximates to a ball. An important fact is that in almost all medical images, background is dark and objects of interest are highlighted. This fact excludes the intensity-inverted occasion that is not suitable for SAD to work. This means different medical images from different modalities are inclined to similar intensity profiles, thus SAD can be effective in this scenario. Our experiments show it outperforms MI.

## 2.2 Choice of Wavelet Coefficients

We use the standard discrete wavelet transform (with Haar basis function which is easy to implement[12]) algorithm to get wavelet decompositions of images

which include four subband images, LL, LH, HL and HH. In our method, we make use of LL subband coefficients to perform registration (See Fig. 1). Rather than the other three subbands which contain details like edges, LL subband preserves most of the significant information one image has, e.g., anatomical structures, intensity values, etc. Thus, as shown in the experiments, it is suitable for MI and SAD to get satisfying results.



**Fig. 1.** Pyramid images. Top left four are Gaussian Pyramid. Bottom left four are wavelet pyramid (LL coefficients). Right are zoomed-in images showing the difference.

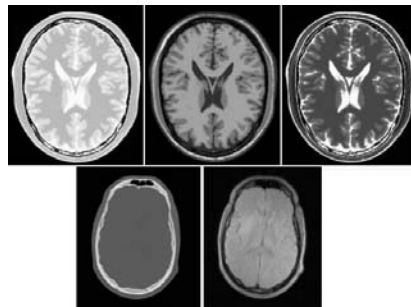
Wavelet transform provides not only frequency domain information but also spatial domain information. A considerable amount of registration work based on wavelet has been done in the remote sensing, cartography, geoscience and so forth. However in the area of medical image registration, little work has involved wavelet transform. We adopt wavelet instead of Gaussian as the representation for two main reasons. Firstly, wavelet transform can keep more significant anatomies in medical images, like the gyrus and ventricles of the brain (See Fig. 1). Secondly, wavelet does not blur the images through the hierarchical pyramid as much as Gaussian filter.

### 2.3 Multiresolution Registration

We first perform wavelet transform on each image into several levels. The total number of levels of decomposition depends on the image size. Suppose that there are  $N$  levels. Level 1 represents the highest resolution level. We start the registration process from the coarsest level, i.e. level  $N$ . For every level  $i$  ( $i > N/2$ ), we register the two LL subband images by minimization of SAD and for every level  $i$  ( $i \leq N/2$ ), we register the two LL subband images by maximization of MI. For each  $i$  ( $N \geq i \geq 2$ ), the resulting position of registration of level  $i$  is multiplied by 2 and then becomes the initial position of the next level, i.e. level  $i - 1$ . The registration process terminates when the matching criterion is optimized at the highest resolution level.

### 3 Experimental Results

The testing datasets used in our experiments were obtained from the Brainweb, a Simulated Brain Database (T1, T2 and PD-weighted MR) [13,14], and from the collaborating hospital (MR,CT), where all the corresponding images have been precisely registered(See Fig. 2). We tested the proposed algorithm and MI-based Gaussian pyramid algorithm on 4 groups of multi-modal medical images, T1-T2, T1-PD, T2-PD and MR-CT. In the experiments, we decomposed images into  $N = 4$  levels. In the MI-based Gaussian pyramid algorithm, we used MI for the registration of all resolution levels. In our method, we used SAD for the registration of Levels 4 and 3 LL subband images, and MI for the registration of Levels 2 and 1 LL subband images.



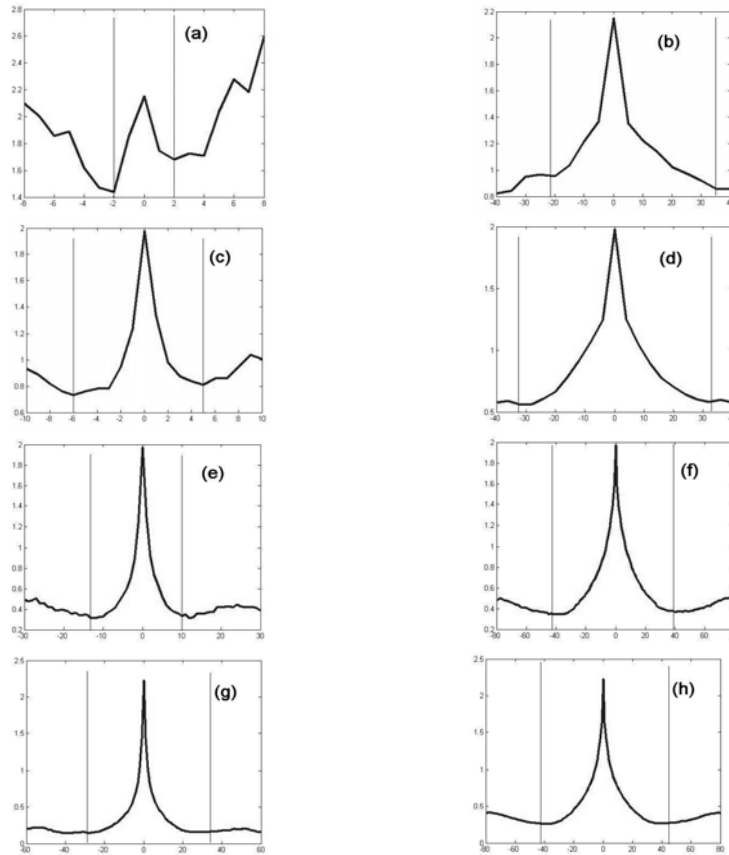
**Fig. 2.** *Datasets for the experiments. The first column are T1, T2 and PD( $181 \times 217 \times 181$ ) images from Brainweb. The second column are MR and CT images( $256 \times 256$ ) from the collaborating hospital.*

In order to study their performances, we plotted the probing graphs of our method and the MI-based method in Figs. 3 and 4 respectively. The number of bins was set to 32. To minimize the interpolation artifact, we adopted the partial volume interpolation (PVI) during the density estimation process [4]. In the figures, capture ranges are delimited by two vertical lines.

Here are some observations about the results. Firstly, in the low resolution translational shifts (Sub-figures (a) and (c) in Fig. 3 and Fig. 4), the new method has a much longer capture range (over double) than the traditional one. This is further justified by the remarkable improvements on different multi-modal image matching groups. The capture ranges along the horizontal direction are given in Table 1. For rotations (Sub-figures (b) and (d) in Fig. 3 and Fig. 4), capture ranges are similar because the overlapping regions of objects do not vary significantly.

Secondly, in high resolution translational shifts (Sub-figures (e) and (g) in Fig. 3 and Fig. 4), the new method has a slightly longer capture range.

Thirdly, in high resolution rotations (Sub-figures (f) and (h) in Fig. 3 and Fig. 4), capture ranges are similar. It is observed that, when the images are aligned,

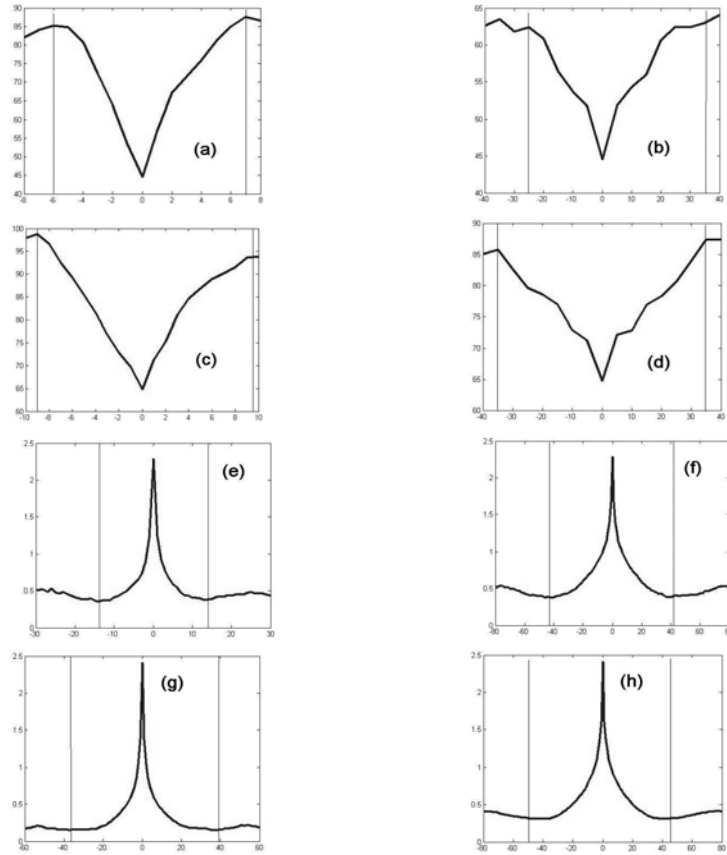


**Fig. 3.** The conventional MI-based method with Gaussian pyramid on T1-T2 images. Level 4 is at the top row of the figure while level 1 is at the bottom row. The left column gives the values of MI (All levels) across the translations along the x axis and the right column gives the corresponding values for different rotations.

the MI value of the wavelet LL subband images (around 2.3) is larger than that of the Gaussian filtered images (around 2.1). It shows that the wavelet LL subband images contain more mutual information because the anatomical structures are better preserved in the LL subband images.

#### 4 Summary

In this paper, we present a new multi-modal medical image registration method based on wavelet transform using SAD and MI. Our method extends the short capture ranges with the conventional MI-based Gaussian pyramid method and meantime achieves the same accuracy. The novelty of this idea lies in the combination of two different similarity measures (MI and SAD) in one multiresolution



**Fig. 4.** Our proposed method on T1-T2 images. Level 4 is at the top row of the figure while level 1 is at the bottom row. The left column gives the values of SAD (Levels 4 and 3) and the values of MI (Levels 2 and 1) across the translations along the x axis and the right column gives the corresponding values for different rotations.

scheme (wavelet pyramid). In principle, this framework can adopt any two metrics and any representation as long as the result is satisfying. Our experiments show that SAD together with wavelet LL subband coefficients obtains a much longer capture range than the conventional method in low resolution registration. Future work will apply wavelet transform to a large number of datasets to further study its performance compared to other existing methods.

## References

1. J.Hajnal, D.Hill, D.Hawkes, eds.: Medical Image Registration. 1st edn. CRC Press (2001)

		MI + Gaussian	Our method	Improvements
T1-T2	Level 4	6.7±1.5	13.7±1.1	104%
	Level 3	12±2.4	22.4±3.6	87%
T1-PD	Level 4	8.1±2.9	15.6±2.5	93%
	Level 3	12.9±2.5	26.9±3.3	109%
T2-PD	Level 4	7.1±1.6	14.5±2.2	104%
	Level 3	13.4±2.6	21.7±2.3	62%
MR-CT	Level 4	13.4±2.1	22.1±3.4	65%
	Level 3	30.2±3.9	40.4±7.1	34%

**Table 1.** Comparisons of capture ranges in four multimodal image matching groups. For each group, ten slices are randomly selected. Mean and standard deviation are calculated. Probing was performed across translational shifts along the  $x$  axis. Unit: pixels.

2. W.M.III.Wells, P.Viola, H.Atsumi, S.Nakajima, R.Kikinis: Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis* **1** (1996) 35–51
3. J.P.W.Pluim, J.B.A.Maintz, M.A.Viergever: Mutual-information-based registration of medical images: A survey. *IEEE Trans. Medical Imaging* **22** (2003) 986–1004
4. F.Maes, A.Collignon, D.Vandermeulen, G.Marchal, P.Suetens: Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imag.* **16** (1997) 187–198
5. J.P.W.Pluim, J.B.A.Maintz, M.A.Viergever: Mutual information matching in multiresolution contexts. *Image and Vision Computing* **19** (2001) 45–52
6. M.Irani, P.Anandan: Robust multi-sensor image alignment. In: *Proc. DARPA Image Understanding Workshop. Volume 1.* (1997) 639–647
7. J.L.Moigne, W.J.Campbell, R.F.Crompt: An automated parallel image registration technique based on the correlation of wavelet features. *IEEE Trans. Geoscience and Remote Sensing* **40** (2002) 1849–1864
8. P.Moon, G.Jager: An investigation into the applicability of the wavelet transform to digital stereo image matching. In: *Joint IEEE Communications and Signal Processing Society.* (1993) 75–79
9. R.Turcajova, J.Kautsky: A hierarchical multiresolution technique for image registration. In: *Proc. of SPIE Mathematical Imaging: Wavelet: Applications in Signal and Image Processing.* (1996)
10. J.Liu, B.C.Vemuri, J.L.Marroquin: Local frequency representations for robust multimodal image registration. *IEEE Trans. Medical imaging* **21** (2002) 462–469
11. J.Zhen, J.Yifeng, Z.Jihong: Automatic registration algorithm based on wavelet transform. In: *Proc. of International Conference on Signal Processing 2000.* (2000) 979–982
12. R.C.Gonzalez, R.E.Woods: *Digital Image Processing.* 2nd edn. Prentice Hall (2002)
13. BrainWeb: (McGill) <http://www.bic.mni.mcgill.ca/brainweb/>.
14. C.A.Cocosco, V.Kollokian, R.K.Kwan, A.C.Evans: Brainweb: Online interface to a 3d mri simulated brain database. In: *Proceedings of 3rd International Conference on Functional Mapping of the Human Brain.* (1997)