

# Multiresolution Image Registration Based on Kullback-Leibler Distance

Rui Gan<sup>1</sup>, Jue Wu<sup>2</sup>, Albert C.S. Chung<sup>1</sup>,  
Simon C.H. Yu<sup>3</sup>, and William M. Wells III<sup>4,5</sup>

<sup>1</sup> Department of Computer Science, and

<sup>2</sup> Bioengineering Program, School of Engineering,  
Hong Kong University of Science and Technology, Hong Kong.

{raygan, johnwoo, achung}@cs.ust.hk

<sup>3</sup> Dept. of Diagnostic Radiology and Organ Imaging, Prince of Wales Hospital, HK.

<sup>4</sup> Brigham & Women's Hospital, Harvard Medical School, Boston, MA, U.S.A.

<sup>5</sup> MIT CSAI Laboratory, Cambridge, MA, U.S.A.

**Abstract.** This paper extends our prior work on multi-modal image registration based on the *a priori* knowledge of the joint intensity distribution that we expect to obtain, and Kullback-Leibler distance. This expected joint distribution can be estimated from pre-aligned training images. Experimental results show that, as compared with the Mutual Information and Approximate Maximum Likelihood based registration methods, the new method has longer capture range at different image resolutions, which can lead to a more robust image registration method. Moreover, with a simple interpolation algorithm based on non-grid point random sampling, the proposed method can avoid interpolation artifacts at the low resolution registration. Finally, it is experimentally demonstrated that our method is applicable to a variety of imaging modalities.

## 1 Introduction

In this paper, we extend our prior work on multi-modal image registration method based on the *a priori* knowledge of the joint intensity distribution that we expect to obtain. This expected joint distribution can be estimated from aligned training images [3,10]. Unlike Mutual Information (MI) based image registration method [7,11], our method makes use of the expected joint intensity distribution between two pre-aligned training images as a reference distribution. Two novel images of the same or different acquisitions are aligned when the expected and observed joint intensity distributions are well matched. The difference between distributions is measured using the Kullback-Leibler distance (KLD). The registration procedure is a multiresolution iterative process. The procedure at the current image resolution is terminated when the KLD value becomes sufficiently small. Then, based on the current estimated transformation, the next higher resolution registration continues until the original image resolution is reached.

Based on the results on T1, T2, PD, CT, and 3DRA image volumes, it is experimentally shown that our method has significantly longer capture range<sup>1</sup> than that of MI based and Approximate Maximum Likelihood (MLa) [6,12] based methods, which can make the multiresolution image registration more robust. Moreover, our method can avoid the problem of interpolation artifact in the low resolution registration. Finally, the experiments demonstrate that our method can be applicable to a variety of imaging modalities.

## 2 Method

### 2.1 Estimation of the Joint Intensity Distributions

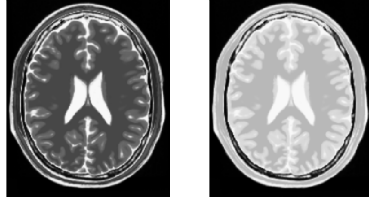
Let  $I_f$  and  $I_r$  be the intensity values of two images of the same or different acquisitions ( $f$  and  $r$  represent respectively the floating and reference images), and  $X_f$  and  $X_r$  be their image domains respectively. Assume that the intensity values of image voxels are independent of each other.

The expected joint distribution can be estimated from a pre-aligned training image pair, which can be obtained from experienced clinicians or other image registration methods (e.g., an MI based method). Given two precisely aligned training image volumes, samples of intensity pairs  $\hat{\mathcal{I}} = \{i_f(x), i_r(x^r) | i_f \in I_f, i_r \in I_r\}$  can be drawn from  $I_f$  and  $I_r$ , where  $x$  are the grid point coordinates in  $X_f$  (i.e. the sampling domain is equal to  $X_f$ ) and  $x^r$  are the corresponding coordinates of  $x$  in  $X_r$ . Histogram partial volume (PV) interpolation [7] is used to achieve subvoxel accuracy in registration results. The expected joint intensity distribution  $\hat{P}(I_f, I_r)$  can be approximated by either Parzen windowing or histogramming [1]. Histogramming is employed in this paper because the approach is computationally efficient.

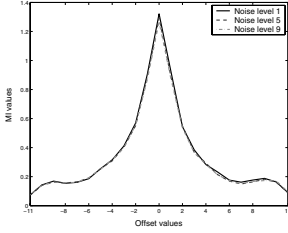
For the observed joint intensity distribution, given a novel testing image pair with a hypothesized transformation  $T$ , samples of intensity pairs  $\mathcal{I}_o = \{i_f(x), i_r(T(x)) | i_f \in I_f, i_r \in I_r\}$  can be drawn from  $I_f$  and  $I_r$ , where  $x$  are the coordinates in  $X_f$ . Note that the observed joint intensity distribution  $P_o^T(I_f, I_r)$  is dependent on the transformation  $T$  and changes during the registration. The histogramming approach is used to estimate the distribution  $P_o^T$ .

According to our experiments, interpolation artifacts may occur if we only draw samples at grid positions from  $X_f$  and apply PV interpolation in  $X_r$ , especially for the same voxel-size image pair of low resolution (see Section 3.1 for more details). Similar observations for the problem have been reported in [8]. In this paper, we propose a simple improvement of PV interpolation based on random sampling to avoid interpolation artifacts and increase the robustness of our method. It can be outlined as follows. Instead of drawing samples at grid positions, we randomly draw samples from  $X_f$ . If the sample is a grid position, we update the joint distribution by using the same method as before; otherwise, i.e. a non-grid position, we perform PV interpolation both in  $X_f$  and  $X_r$ . In

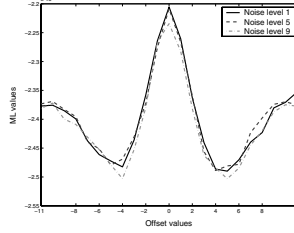
<sup>1</sup> Capture range represents the range of positions from which a registration algorithm can converge to the correct minimum or maximum.



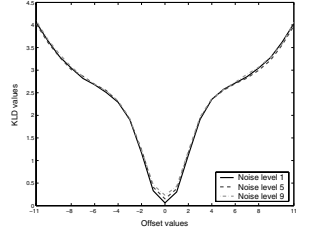
(a) A T2 image. (b) A PD image.



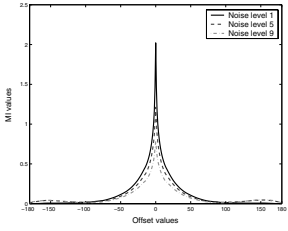
(c) MI



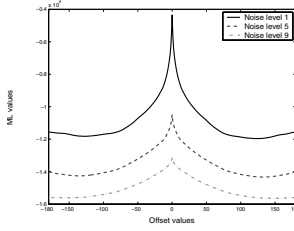
(d) MLa



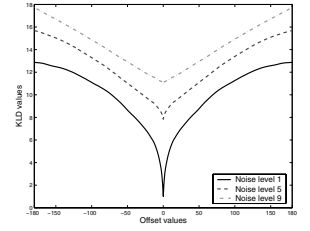
(e) KLD



(f) MI



(g) MLa



(h) KLD

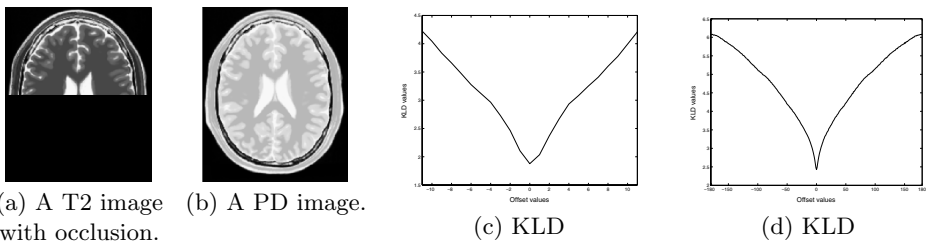
**Fig. 1.** (a) A slice of BrainWeb T2 image volume. (b) A slice of BrainWeb PD image volume. Translational probes for registering low resolution (Level 4) image pairs with different noise levels (1%, 5% and 9%): (c) MI, (d) MLa, and (e) KLD. Translational probes for registering the original resolution (Level 0) image pairs: (f) MI, (g) MLa, and (h) KLD. Training dataset was 0% noise level for all probes.

practice, the sampling rate can be set to 10% – 50% of the total number of voxels in the floating image.

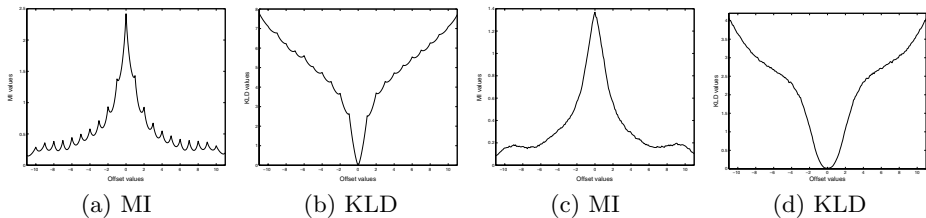
For a further illustration, suppose that  $s$  is a non-grid position in  $X_f$ . Let  $\mathcal{N}_s$  be a set of neighboring grid positions of  $s$  in  $X_f$  and  $\tilde{\mathcal{N}}_{T(s)}$  be that of  $T(s)$  in  $X_r$ . Then, the update of  $P_o^T$  is given by

$$\forall n \in \mathcal{N}_s \text{ and } \forall \tilde{n} \in \tilde{\mathcal{N}}_{T(s)} : P_o^T(i_f(n), i_r(\tilde{n}))_+ = \frac{w_n \cdot \tilde{w}_{\tilde{n}} \cdot D(T(n), \tilde{n})}{Z}, \quad (1)$$

where  $w_n$  and  $\tilde{w}_{\tilde{n}}$  are respectively the corresponding fractions of position  $n$  and position  $\tilde{n}$  in PV interpolation and can be determined via the trilinear interpolation [7],  $D(T(n), \tilde{n})$  is a decreasing function with respect to the Euclidean distance between positions  $T(n)$  and  $\tilde{n}$ , and  $Z$  is a normalizing factor which keeps  $\sum_{\tilde{n} \in \tilde{\mathcal{N}}_{T(s)}} \sum_{n \in \mathcal{N}_s} P_o^T(i_f(n), i_r(\tilde{n})) = 1$ . In this paper, we set  $D(s_1, s_2)$



**Fig. 2.** (a) An occluded slice of BrainWeb T2 image volume. (b) A slice of BrainWeb PD image volume. Translational probes for registering image pairs with different resolutions: (c) KLD level 4 (low resolution), (d) KLD level 0 (original resolution).



**Fig. 3.** Interpolation artifacts in the low resolution (level 4) image registration. Results of translational probing experiments based on sampling on grid points: (a) MI and (b) KLD. Results based on non-grid point random sampling: (c) MI and (d) KLD.

to  $(2 \cdot L_d^r - d_{s_1, s_2})$ , where  $L_d^r$  is the diagonal length of voxels in  $X_r$  and  $d_{s_1, s_2}$  represents the Euclidean distance between positions  $s_1$  and  $s_2$ .

### 2.2 Kullback-Leibler Distance (KLD) and Multiresolution Optimization

Given the expected  $\hat{P}$  and observed  $P_o^T$  joint intensity distributions, the Kullback-Leibler distance between the two distributions is given by

$$D(P_o^T || \hat{P}) = \sum_{i_f, i_r} P_o^T(i_f, i_r) \log \frac{P_o^T(i_f, i_r)}{\hat{P}(i_f, i_r)}. \tag{2}$$

KLD may be used to measure the similarity between two distributions. According to [5], the KLD value is non-negative and becomes zero if and only if two distributions are equivalent. The more similar the two distributions are, the smaller the KLD value is. Therefore, when the two testing images  $I_f$  and  $I_r$  are not perfectly registered, the value of KLD,  $D$ , will be positive and relatively large because  $P_o^T$  and  $\hat{P}$  are not similar. On the other hand, if the testing images are well registered, then the value of KLD becomes small or is equal to zero (i.e.  $P_o^T$  is very similar or equal to  $\hat{P}$ ).

The goal of registration is to find the optimal transformation  $\hat{T}$  by minimizing the difference between the observed  $P_o$  and expected  $\hat{P}$ , which is formulated as

$$\hat{T} = \arg \min_T D(P_o^T || \hat{P}). \quad (3)$$

The proposed method is conceptually different from the MI based registration method, which encourages functional dependence between the two image random variables,  $I_f$  and  $I_r$ . The KLD based registration method guides the transformation  $T$  based on the difference between the expected  $\hat{P}$  and observed  $P_o^T$  joint intensity distributions, or, in other words, based on the expected outcomes learned from the training data.

In order to accelerate the registration process and ensure the accuracy and robustness of the proposed method, we exploit a multiresolution approach based on the Gaussian Pyramid representation [2,11]. Rough estimates of  $\hat{T}$  can be found using downsampled images and treated as starting values for optimization at higher resolutions. Then, the fine-tuning of the solution can be derived at the original image resolution.

In this paper, the value of KLD for each resolution is minimized by Powell's method with a multiresolution strategy [9] because it does not require calculations of gradient and, hence, is simpler in terms of implementation. Powell's method iteratively searches for the minimum value of KLD along each parameter axis  $T$  (1D line minimization) while other parameters are kept constant.

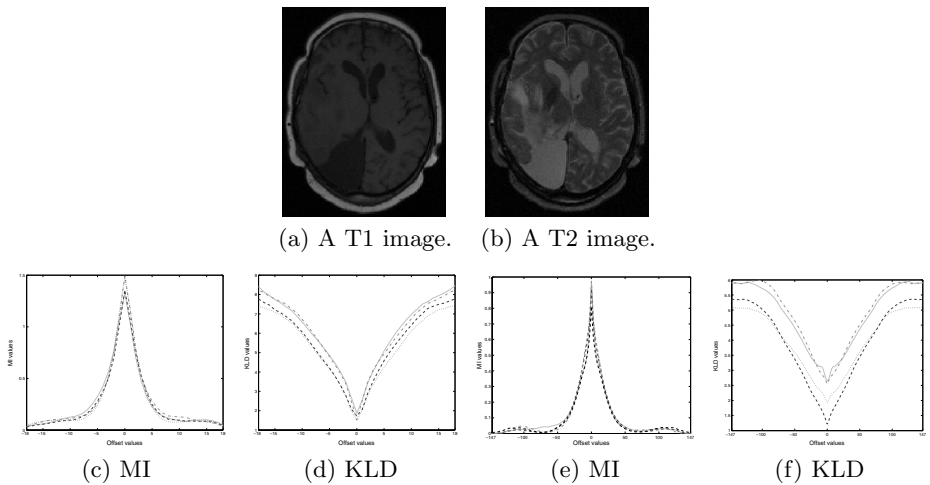
### 3 Experimental Results

#### 3.1 T2 – PD (3D – 3D) Registration

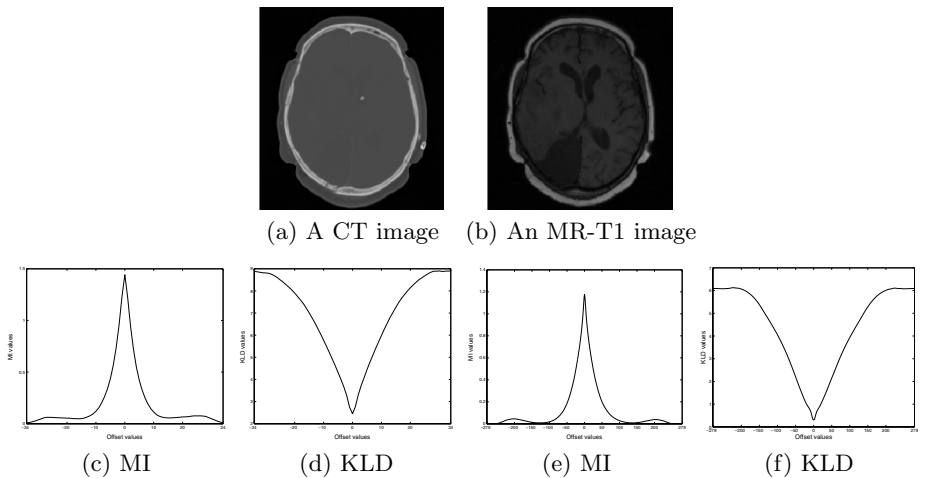
Four pairs of T2 and PD image volumes were used for the registration experiments. The images were obtained from the BrainWeb Simulated Brain Database [4] ( $181 \times 217 \times 181$  voxels,  $1 \times 1 \times 1 \text{ mm}^3$  and noise levels were 0%, 1%, 5% and 9%). Figures 1a and 1b show T2 and PD image slices respectively. An image pair with 0% noise level was used as the training image pair and the others were the testing image pairs. Figures 1c, 1d and 1e plot the translational probes for registering the low resolution<sup>2</sup> (Level 4) testing image pairs with different noise levels for the Mutual Information (MI) [7,11], Approximate Maximum Likelihood (MLa) [6,12] and Kullback-Leibler Distance (KLD) respectively. Similarly, Figures 1f, 1g and 1h plot the translational probes for registering the original resolution (Level 0) image pairs. Figures 1e and 1h show that the capture ranges of KLD are significantly longer than those of MI and MLa.

To study the effect of occlusion, as shown in Figure 2a, only half of the T2 image volume was used for registration experiments. Another PD image volume (Figure 2b) and the training image pair were unchanged. Results (Figures 2c and 2d) show that our method works when the testing image is occluded.

<sup>2</sup> The definition of resolution levels in the Gaussian Pyramid representation follows the same line as in [2]. The smoothing filter was  $\{1, 4, 6, 4, 1\}$  in our experiments.

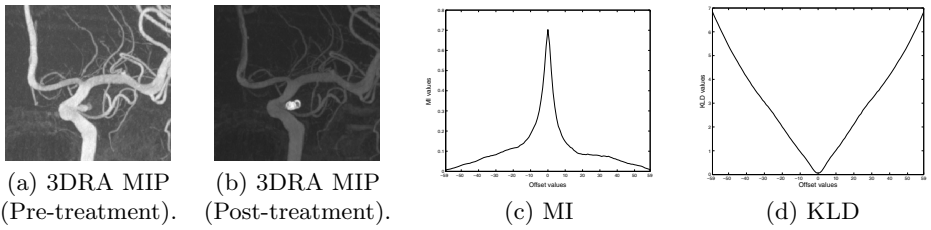


**Fig. 4.** (a) A slice of Vanderbilt T1 image volume. (b) A slice of Vanderbilt T2 image volume. Translational probes for registering the low resolution (Level 3) image pairs from different testing datasets: (c) MI and (d) KLD. Translational probes for registering the original resolution (Level 0) image pairs: (e) MI and (f) KLD. Training datasets were from patient #001.



**Fig. 5.** (a) A slice of Vanderbilt CT image volume. (b) A slice of Vanderbilt MR-T1 image volume. Translational probes for registering the low resolution (Level 3) image pair: (c) MI and (d) KLD. Translational probes for registering the original resolution (Level 0) image pair: (e) MI and (f) KLD.

Figures 3a and 3b illustrate that, when we only draw samples at grid positions for the estimation of joint distributions, interpolation artifacts occur for both MI and KLD based methods at the low resolution (level 4). With the random sam-



**Fig. 6.** (a) Maximum intensity projection (MIP) of a pre-treatment 3D rotational angiography (3DRA). (b) MIP of a post-treatment 3DRA. Translational probes for registering the image volume pair: (c) MI and (d) KLD.

pling on non-grid positions, Figures 3c and 3d show that interpolation artifact can be avoided. This can make the low resolution registration more robust.

### 3.2 T1, T2, and CT (3D – 3D) Registration

Five pairs of T1 and T2 image volumes were obtained from the Retrospective Registration Evaluation Project <sup>3</sup> ( $160 \times 160 \times 26$  voxels and  $1.25 \times 1.25 \times 4$  mm<sup>3</sup>). Image volumes from patient #001 were used as the training image pair.

Figures 4a and 4b show T1 and T2 slices respectively. Figures 4c (MI) and 4d (KLD) plot the translational probes for registering four low resolution (Level 3) image pairs. At the original image resolution (Level 0), Figures 4e and 4f plot the translational probes based on MI and KLD respectively. Figures 4d and 4f illustrate that the KLD based method consistently gives longer capture range than that of the MI based method (Figures 4c and 4e).

Two pairs of T1 and CT image volumes from the same project were used for the registration experiments. Figures 5a and 5b show CT and T1 image slices respectively. Probes along the horizontal direction are shown in Figures 5c and 5d for the low resolution (Level 3) registration and in Figures 5e and 5f for the original resolution (Level 0) registration. Similar results for the capture range were obtained as compared with the T1 and T2 registration as shown above.

### 3.3 Pre-treatment 3DRA – Post-treatment 3DRA Registration

Two pairs of 3D rotational angiographic (RA) image volumes of the same patient were obtained from the Prince of Wales Hospital, Hong Kong for the registration experiments ( $256 \times 256 \times 256$  voxels and  $0.19 \times 0.19 \times 0.19$  mm<sup>3</sup>). Figures 6a and 6b show two maximum intensity projections (MIP) of the pre-treatment and post-treatment image volumes respectively. (Note that the MIPs are for visualization purposes only, not part of the registration process.) The figures show that our method can be applied to 3DRA image registration and useful for interventional treatment assessments because 3DRA volumes can be compared quantitatively.

<sup>3</sup> Images were provided as part of the project, “Evaluation of Retrospective Image Registration”, National Institutes of Health, Project Number 1 R01 NS33926-01, Principle Investigator, J. Michael Fitzpatrick, Vanderbilt University, Nashville, TN.

## 4 Summary and Conclusions

This paper has proposed a multiresolution multi-modal image registration method based on minimizing the Kullback-Leibler distance (KLD) between the observed and expected joint intensity distributions until the two testing datasets are aligned. The results consistently show that our method has a longer capture range than that of MI and MLa based methods. The proposed simple non-grid point random sampling method has been experimentally shown that it can avoid the problem of interpolation artifact in the low resolution registration based on histogram partial volume (PV) interpolation. Future work will include a further validation of the proposed algorithm by applying it to a large number of datasets.

**Acknowledgements.** W. Wells would like to acknowledge support from the NSF ERC grant (JHU Agreement #8810-274) and the NIH (grant #1P41RR13218). A. Chung would like to acknowledge support from the HK RGC grants (HKUST6155/03E and HKUST6209/02E).

## References

1. C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford U. Press, 1995.
2. P.J. Burt and E.H. Adelson. The Laplacian Pyramid as a Compact Image Code. *IEEE Trans. Comms.*, 31(4):532–540, 1983.
3. A.C.S. Chung, W.M. Wells II, and et. al. Multi-Modal Image Registration by Minimising Kullback-Leibler Distance. In *MICCAI*, pages 525–532, 2002.
4. D.L. Collins, A.P. Zijdenbos, and et al. Design and Construction of a Realistic Digital Brain Phantom. *IEEE Trans. Med. Img.*, 17(3):463–468, 1998.
5. T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.
6. M.E. Leventon and W.E.L. Grimson. Multi-Modal Volume Registration Using Joint Intensity Distributions. In *MICCAI*, pages 1057–1066, 1998.
7. F. Maes, A. Collignon, and et al. Multimodality Image Registration by Maximization of Mutual Information. *IEEE Trans. Med. Img.*, 16(2):187–198, 1997.
8. J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. Interpolation Artefacts in Mutual Information-Based Image Registration. *CVIU*, 77:211–232, 2000.
9. W.H. Press, S.A. Teukolsky, and et al. *Numerical Recipes in C, 2nd Edition*. Cambridge University Press, 1992.
10. S. Soman, A.C.S. Chung, and et. al. Rigid Registration of Echoplanar and Conventional MR Images by Minimizing Kullback-Leibler Distance. In *WBIR*, pages 181–190, 2003.
11. 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.
12. L. Zöllei, J.W. Fisher III, and W.M. Wells III. A Unified Statistical and Information Theoretic Framework for Multi-modal Image Registration. In *IPMI*, pages 366–377, 2003.