

# Multi-modal image registration by minimizing Kullback-Leibler distance between expected and observed joint class histograms

Ho-Ming Chan<sup>1</sup>, Albert C.S. Chung<sup>1\*</sup>, Simon C.H. Yu<sup>2</sup>, Alexander Norbash<sup>3</sup> and William M. Wells III<sup>3,4</sup>

<sup>1</sup>Department of Computer Science, Hong Kong University of Science and Technology, Hong Kong.

<sup>2</sup>Department of Diagnostic Radiology and Organ Imaging, Prince of Wales Hospital, Hong Kong.

<sup>3</sup>Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.

<sup>4</sup>Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA.

hmchan@cs.ust.hk    achung@cs.ust.hk    simonyu@cuhk.edu.hk    anorbash@partners.org    sw@bwh.harvard.edu

## Abstract

*In this paper, we present a new multimodal image registration method based on the a priori knowledge of the class label mappings between two segmented input images. A joint class histogram between the image pairs is estimated by assigning each bin value equal to the total number of occurrences of the corresponding class label pairs. The discrepancy between the observed and expected joint class histograms should be minimized when the transformation is optimal. Kullback-Leibler distance (KLD) is used to measure the difference between these two histograms.*

*Based on the probing experimental results on a synthetic dataset as well as a pair of precisely registered 3D clinical volumes, we showed that, with the knowledge of the expected joint class histogram, our method obtained longer capture range and fewer local optimal points as compared with the conventional Mutual Information (MI) based registration method. We also applied the proposed method to 2D-3D rigid registration problems between DSA and MRA volumes. Based on manually selected markers, we found that the accuracies of our method and the MI-based method are comparable. Moreover, our method is more computationally efficient than the MI-based method.*

## 1. Introduction

Images acquired by different medical imaging modalities provide useful complementary information. For example, magnetic resonance (MR) images provide anatomical information; PET and SPECT images provide functional information. Multi-modal image registration is an important technique for integrating the complementary image information from different modalities by aligning the images.

To correctly align two images, we need a similarity measure to determine how well the images match with each other through a hypothesized spatial transformation. Most of the similarity measures can be classified into two

categories: feature-based or intensity-based. In the following subsections, we will introduce these similarity measures, and discuss their advantages and disadvantages. In general, the intensity-based method is more accurate but the feature-based method is more computationally efficient [12]. Since accuracy is important in medical diagnosis, we compare our proposed method with Mutual Information (MI) [4],[5], which is one of the most commonly used intensity-based methods in medical image analysis.

We present probing experimental results on a synthetic dataset and 3D clinical magnetic resonance angiograms (MRA) in Section 3 and results on clinical datasets consisting of MRA volumes and DSA in Section 4. It is shown in Section 5 that the accuracies of our proposed method and MI-based method are comparable. Moreover, our method is more computationally efficient than the MI-based method.

### 1.1. Intensity-Based Methods

Over the past few years, many intensity-based similarity measures have been introduced, for example Euclidean distance, correlation coefficients, correlation ratio [11] etc. The most popular and widely used one is the information-theoretic similarity measure, Mutual Information (MI) [4],[5]. It only makes use of the assumption of statistical dependence between the two images, and has been successfully applied to many multi-modality combinations.

Let  $X_1$  and  $X_2$  be image domains,  $T: X_1 \rightarrow X_2$  be the rigid transformation between the two input images (in 3D, the rigid transformation has three translational and three rotational parameters), and  $I_1$  and  $I_2$  be the intensities of images  $X_1$  and  $X_2$  respectively. Then Mutual Information  $MI(T(X_1), X_2)$  is defined by

$$MI(T(X_1), X_2) = \sum_{i \in I_1} \sum_{i_2 \in I_2} P_T(i_1, i_2) \log \frac{P_T(i_1, i_2)}{P_{I_1}(i_1)P_{I_2}(i_2)}, \quad (1)$$

where  $P_T$  is the joint intensity probability density

\* Contact author (email address: achung@cs.ust.hk)

function, and  $P_{I_1^T}$  and  $P_{I_2^T}$  are the corresponding marginal probability density functions given the hypothesized transformation.

The value of  $MI(T(X_1), X_2)$  is expected to be maximized when the images  $X_1$  and  $X_2$  are correctly aligned.

## 1.2. Feature-Based Methods

The feature-based registration algorithms extract features such as surfaces, curves [10] or skeletons [8],[9] at the preprocessing steps.

They are faster than intensity-based methods but comparatively less accurate [12]. Also, the extraction of skeletons, for example, can be sensitive to noise [13]. The surface or curve construction can be complicated and would also affect the registration accuracy.

## 1.3. Related Works

Chung et al. [1] used a pair of precisely registered or segmented images to build the expected joint intensity histogram and used Kullback-Leibler distance (KLD) [2],[3] to measure the discrepancy between observed and expected joint intensity histograms. Usually, the images are aligned by experienced clinicians so that the expected joint intensity histogram can be obtained. If manual alignment is not available, the expected joint intensity histogram will be estimated from segmented images.

Instead of using a joint intensity histogram, we explore the possibility of using a joint class histogram, where each bin of the histogram corresponds to a class label. The idea of using a joint class histogram instead of a joint intensity histogram was inspired partly by Wyatt and Noble [6], who use entropies, similar to MI, as the similarity measure. By using a joint class histogram, we expect that the proposed method would be more efficient than the MI-based method as the number of bins is typically smaller.

Although our method requires segmentation in preprocessing, we do not need to extract any further features such as surfaces, curves or skeletons as required by other feature-based methods.

## 2. Our Registration Algorithm

We will present some background information about our similarity measure in the following subsections before introducing our registration algorithm.

### 2.1. Estimation of Joint Class Histograms

In addition to the notations defined in Section 1.1, let  $\Phi_1$  and  $\Phi_2$  be the sets of possible class labels of the

two images (a class can represent an anatomical structure in medical images); and  $L_1: X_1 \rightarrow \Phi_1$  and  $L_2: X_2 \rightarrow \Phi_2$  be the mappings from a voxel coordinate to its class label in images  $X_1$  and  $X_2$  respectively.

Assuming that the mappings between class labels of two images are known, there exists a mapping relation  $R: \Phi_1 \times \Phi_2$  such that  $R = \{(l_1, l_2) \mid \text{there is a class label mapping between } l_1 \text{ and } l_2, \text{ where } l_1 \in \Phi_1, l_2 \in \Phi_2\}$ . This mapping relation incorporates domain knowledge about the valid correspondences among classes.

In order to estimate the expected joint class histogram, for all pixels  $x_2 \in X_2$ , we select  $x_1'$  randomly from  $X_1$  such that  $(L_1(x_1'), L_2(x_2)) \in R$ , and increase the value of  $\text{bin}(L_1(x_1'), L_2(x_2))$  by 1. Then every bin value is divided by the sum of all bin values in order to normalize the histogram.

Similarly, to estimate the observed class histogram, for all pixels  $x_2 \in X_2$ , if  $T^{-1}(x_2) \in X_1$ , we then increase the value of  $\text{bin}(L_1(T^{-1}(x_2)), L_2(x_2))$  by 1. Otherwise, we select  $x_1'$  randomly from  $X_1$  such that  $(L_1(x_1'), L_2(x_2)) \in R$ , and increase the value of  $\text{bin}(L_1(x_1'), L_2(x_2))$  by 1. Then, every bin value is divided by the sum of all bin values in order to normalize the histogram.

An example of estimating the expected joint class histogram is shown in Figure 1.

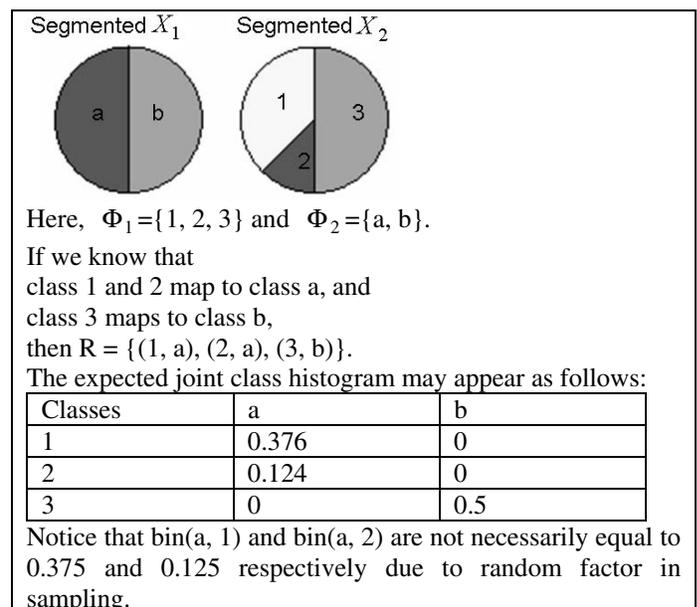


Figure 1: An example of estimation of the expected joint class histogram.

## 2.2. Kullback-Leibler Distance as Similarity Measure

Kullback-Leibler distance [2],[3], which is a frequently used measure in the field of information theory, is employed to measure the difference between the expected joint class histogram, denoted by  $\hat{H}$ , and observed joint class histogram, denoted by  $H_o^T$ . The KLD between the two histograms is given by

$$KLD(H_o^T | \hat{H}) = \sum_{l_1 \in \Phi_1} \sum_{l_2 \in \Phi_2} H_o^T(l_1, l_2) \log \frac{H_o^T(l_1, l_2)}{\hat{H}(l_1, l_2)}. \quad (2)$$

When the two images  $X_1$  and  $X_2$  are perfectly aligned, the value of KLD is expected to be minimum.

## 2.3. Outline of Our Registration Algorithm

Our registration algorithm is summarized as follows.

On both floating and reference images, segmentation is performed to label each voxel/pixel (Segmentation methods can vary and depend on the applications, which is not the focus of this paper). Mappings between their class labels can be determined manually. The expected joint class histogram  $\hat{H}$  can then be estimated as mentioned in Section 2.1.

During the registration process, given the current transformation  $T$ , we can estimate the observed joint class histogram  $H_o^T$ , as described in Section 2.1.

The goal is to find the optimal transformation  $\hat{T}$  by minimizing the value of KLD (Equation (2)) between the expected joint class histogram and observed joint class histogram.

$$\hat{T} = \arg \min_T KLD(H_o^T | \hat{H}). \quad (3)$$

This proposed KLD-based registration method is conceptually different from the MI-based method. It makes use of the *a priori* knowledge of the expected joint class histogram to guide the transformation towards the expected outcome.

Powell's method [7] is used to iteratively search for the minimum value of KLD along each parameter using Brent's method [7]. The algorithm halts when the percentage change of KLD values is below a user-specified threshold. We set this threshold to 0.001% in our program.

The flow chart of our registration algorithm is shown in Figure 2.

## 3. Comparative Studies Between KLD and MI Similarity Measures

In this section, we will compare our similarity

measure with the Mutual Information (MI) based similarity measure. The first probing experiment used synthetic data. The second probing experiment used a pair of precisely registered 3D clinical datasets.

## 3.1. Experiment on Synthetic Data

We constructed two images for this experiment, as shown in Figure 3. Both of them were 100 x 100 pixels. The first image consists of 3 vertical strips, where the left and right strips are 40 pixels wide and the middle one is 20 pixels wide. The intensity values of the middle strip are normally distributed with mean = 204 and standard deviation = 5, while the intensities of left and right strips were all zero. The second image consisted of 5 vertical strips. Each of them was 20 pixels wide. The intensity values in each strip were normally distributed with the same standard deviation = 5. The mean intensities of each strip were 20, 44, 204, 44 and 20 from left to right respectively.

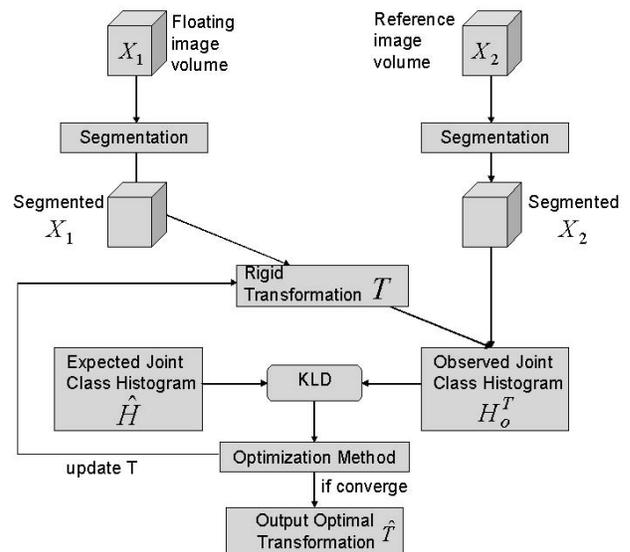


Figure 2: A flow chart of our registration algorithm.

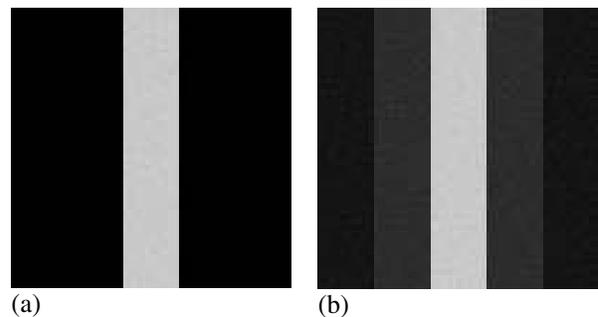


Figure 3: (a) Synthetic floating image, and (b) synthetic reference image.

There are 2 classes in the floating image (Figure 3a),

one for middle strip (class 1) and another one for strips on both sides (class 2). Also, the reference image (Figure 3b) consists of 5 classes where each strip corresponds to one class. We label them as classes a, b, c, d and e from left to right accordingly. It is expected that class c maps to class 1 and classes a, b, d and e map to class 2. This information is used as our *a priori* knowledge for our KLD-based similarity measure.

We computed the values of KLD and MI between these two images at different horizontal translations ranging from -40 pixels to +40 pixels.

The probing results are plotted in Figure 4, which reveals that the capture range of KLD ( $\pm 20$ ) is longer than that of MI (about  $\pm 14$ ). The number of local optimal points is fewer for KLD as well. The global optimum occurs at null transformation for KLD but not for MI, which has global optimal values at -40 or 40 away from null transformation.

It shows that, in this synthetic dataset, the KLD-based similarity measure can provide a longer capture range and fewer local optimal points.

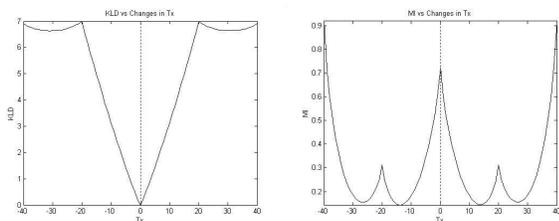


Figure 4: Probing results of KLD (left) and MI (right) values against changes in horizontal translation.

### 3.2. Experiments on 3D Clinical Data

We used a pair of precisely registered 3D phase contrast magnetic resonance image volumes (512 x 512 x 50 voxels of 0.352mm x 0.352mm x 0.8mm) for the experiments. In the experiments, the floating and reference images were the MRA speed images and the tissue images respectively, which were acquired in the same scan. Both speed images and tissue images show the vascular structures clearly, thus they can be segmented into two classes (vessel and non-vessel) by global thresholding (the threshold was selected by the experienced clinicians so that vessels were clearly segmented), and the expected joint class histogram can then be estimated as mentioned in Section 2.1. We used the entire image volume for the probing experiment. The values of KLD and MI are plotted against the six rigid body transformation parameters.

The plots of KLD and MI values (Figure 5) against the changes in translation in x-axis (Tx) and y-axis (Ty) as well as the changes in rotation about x-axis (Rx) and z-axis (Rz) are shown. The ranges of Tx and Ty were

$\pm 18$ mm away from the null transformation, while the ranges of Rx and Rz were from  $\pm 30^\circ$  away from null transformation. The global optima of all the graphs shown are correctly located at the null transformation. Notice that MI has more local optimal points than KLD for the changes in Ty due to the overlapping of vessels and other anatomical structures.

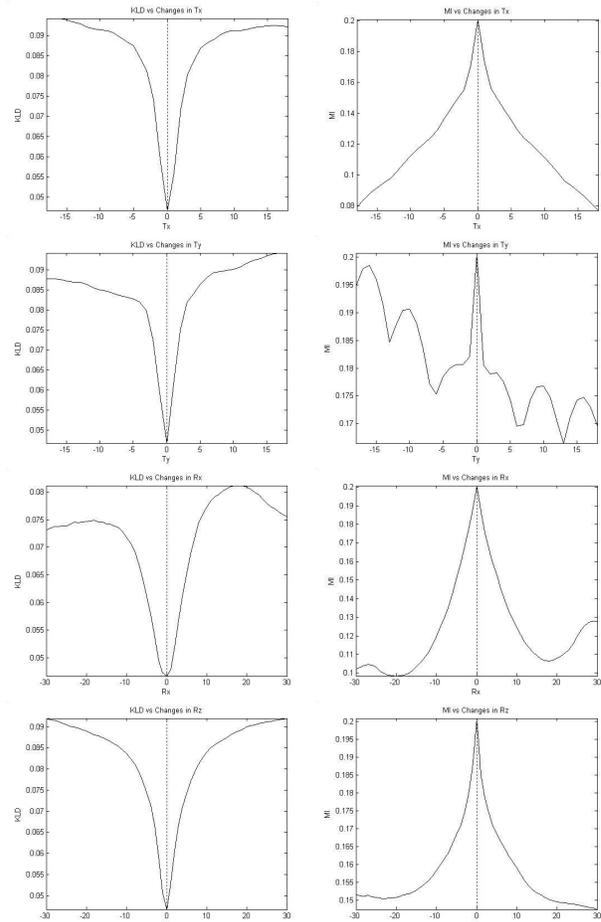


Figure 5: Probing results of KLD (left) and MI (right) values, in which Tx, Ty, Rx and Rz were varied.

The computational time of all the evaluations in the probing experiments presented in this subsection is summarized as follows. For the KLD-based method, the mean and standard deviation are 12.04s and 1.29s respectively while for the MI-based method, the mean and standard deviation are 15.00s and 2.26s respectively. Although the difference of mean computational time between two objective functions is only about 3 seconds, it is essential for the registration process because a registration process generally requires a lot of iterations, e.g., usually more than 1000 iterations.

## 4. 2D-3D Rigid Registration Using the Proposed Algorithm

The proposed KLD-based method has been applied to 2D-3D rigid registration of DSA and MRA, in which the floating image was the 3D MRA volume while the reference image was the 2D DSA image. The outline of the 2D-3D registration and reasons of using the proposed method are given in the following subsections.

### 4.1. Outline of the 2D-3D Registration

There are several additional points to notice when the proposed method was applied to the specific 2D-3D registration problem.

The distortion of the DSA image is corrected by bilinear interpolation using an accurately machined acrylic calibration grid with 20 x 20 embedded ball bearings. On each 3D MRA volume and 2D DSA image, a rectangular region of interest containing the main vessels is selected in order to shorten the computational time. Then, the segmentation is performed to label each voxel/pixel as one of the following two classes: vessel or non-vessel. There are only two mappings between these class labels: vessels map to vessels and non-vessels map to non-vessels. The expected joint class histogram can then be estimated as mentioned in Section 2.1.

At each iteration, the ray-casting technique [14] is used to generate a binary projection of the segmented MRA volume for the current pose. The observed joint class histogram is based on using this projection and the segmented DSA image. The optimization scheme described in Section 2.3 is used to find the optimal transformation.

### 4.2. Reasons for Using the Proposed Method on the 2D-3D Registration of DSA and MRA Images

- i. The KLD-based method has fewer local optimal points than the MI-based method as shown in the plots in Section 3.
- ii. As segmentation is done for both MRA volume and DSA image in the preprocessing steps, the MIP required for building the joint class histogram is only a binary image. Therefore, during the ray casting process for MIP image generation, when the ray hits a vessel, the casting process can be stopped for that particular ray immediately. This technique speeds up the MIP generation process by about 3-4 times which is not feasible for the MI-based method because whether the ray hits a vessel or not cannot be determined without segmentation.
- iii. KLD may be less sensitive to noise compared with feature-based methods using skeletons [13].

## 5. Results and Comparisons

In this section, we demonstrate the results of the 2D-3D rigid registration algorithm presented in Section 4 and also compare its accuracy and computational efficiency with the MI-based method.

### 5.1. Accuracy Study of Our Registration Algorithm

We use two pairs of 3D MRA and 2D DSA datasets for this experiment. The dimensions and the resolutions of the two MRA datasets are 512 x 512 x 50 (0.293mm x 0.293mm x 1mm) and 256 x 256 x 80 (1.25mm x 1.25mm x 2.4mm) respectively.

In order to evaluate the accuracy of the registrations, seven and five target points were chosen by an experienced user using an interactive tool for cases 1 and 2 respectively. After the optimal transformation was found, for every user-selected point on a DSA image, the registration error was defined by the Euclidean distance between the point and the projected corresponding 3D MRA point. The results of cases 1 and 2 are shown in Figure 6 and Figure 7 respectively. Figure 6c and Figure 7c show the MIPs of the registered MRA volumes using the KLD-based method. Notice that full MIPs are shown here for visualization purpose only. During the registration, only binary projection images were generated and used. Segmented vascular regions of the binary projection image (at final alignment using the KLD-based method) are overlaid on the corresponding DSA (Figure 6d and Figure 7d) and the results are promising. The registered images using the MI-based method are also shown for comparison (Figure 6f and Figure 7f).

The summary of errors is tabulated in Table 1. It shows that the registration accuracy of our method is comparable to the MI-based method in terms of the mean error values. The mean error is less than 2mm, which is acceptable in our application.

Case No.		Min Error (mm)	Max Error (mm)	Mean Error (mm)	S.D. (mm)
1	KLD	0.378417	1.60155	1.13038	0.428727
	MI	0.843802	1.55816	1.092088	0.240371
2	KLD	1.47155	2.1964	1.784515	0.364806
	MI	0.942206	3.42202	1.947052	1.148561

Table 1: Results of experiments on 2D-3D registration.

Table 2 lists the final transformation using both methods. Notice that the projection plane for cases 1 and 2 is parallel to xz-plane and yz-plane respectively; therefore, the registration results are insensitive to the changes of  $T_y$  in case 1 and the changes of  $T_x$  in case 2.

Case No.		Tx	Ty	Tz	Rx	Ry	Rz
1	KLD	-0.5442	25.85	29.04	2.174	3.090	-5.834
	MI	-0.1596	20.81	28.57	1.229	3.174	-7.224
2	KLD	-109.94	-19.51	106.0	0.753	-2.267	-1.660
	MI	-43.001	-14.06	105.9	-1.088	-0.524	-13.78

Table 2: Final transformation parameters. Translations are in mms and rotations are in degrees.

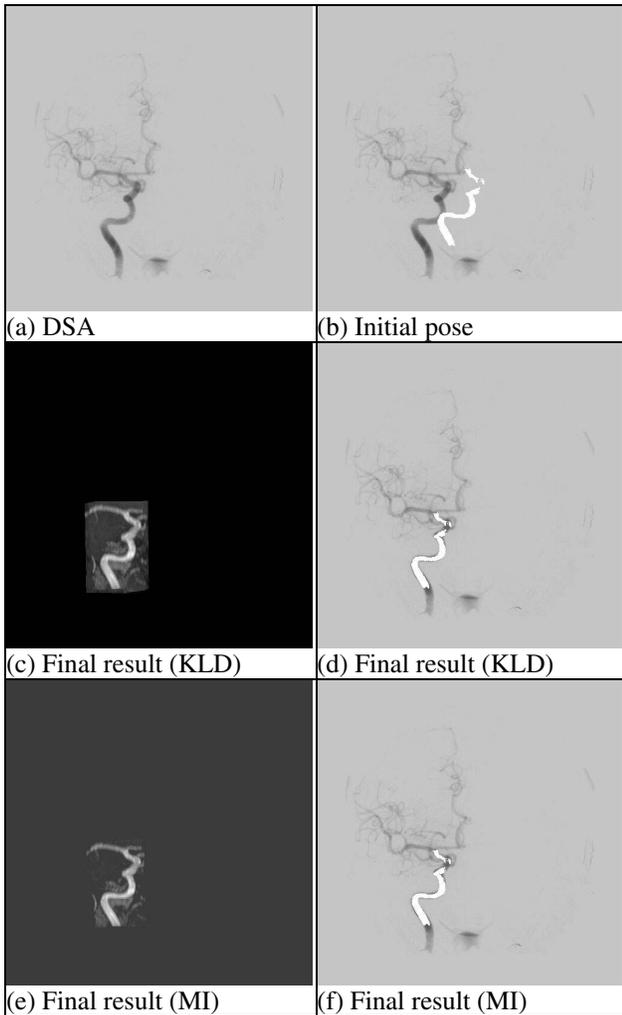


Figure 6: Results of case 1. (a) DSA image, (b) Initial image alignment, (c) Final image alignment by using the KLD-based method, MIP of the MRA volume, (d) Final image alignment by using the KLD-based method, MIP is overlaid on the DSA, (e) Final image alignment by using the MI-based method, MIP of the MRA volume, (f) Final image alignment by using MI-based method, MIP is overlaid on the DSA.

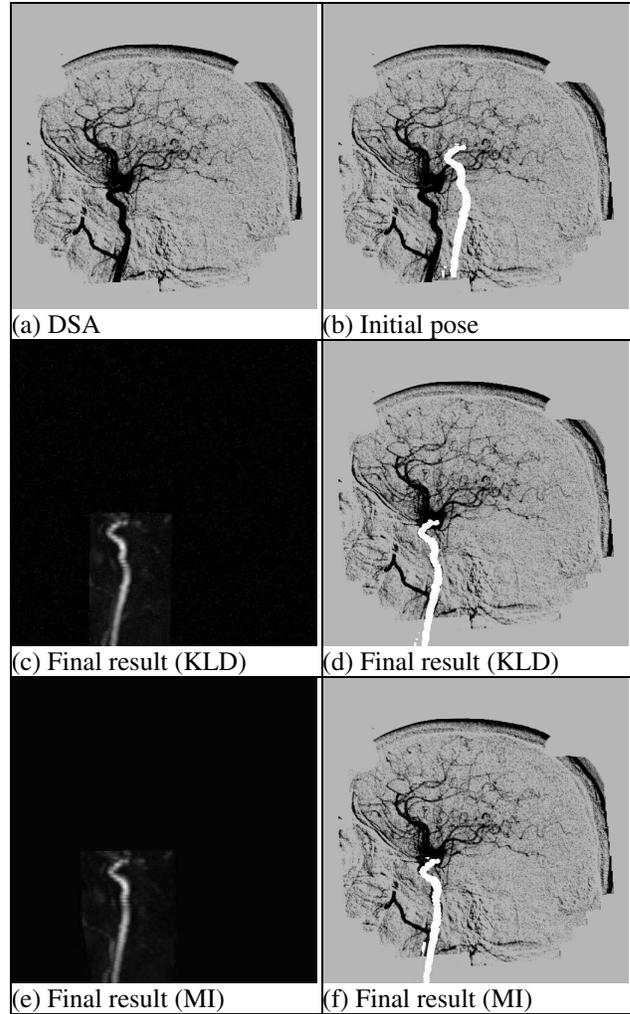


Figure 7: Results of case 2. (a) DSA image, (b) Initial image alignment, (c) Final image alignment by using the KLD-based method, MIP of the MRA volume, (d) Final image alignment by using the KLD-based method, MIP is overlaid on the DSA, (e) Final image alignment by using the MI-based method, MIP of the MRA volume, (f) Final image alignment by using the MI-based method, MIP is overlaid on the DSA.

## 5.2. Computational Time Study of Our Registration Algorithm

According to the Equations (1) and (2), the computational time of evaluation of KLD is expected to be faster than that of MI for the following reasons:

- i. KLD does not require the calculation of marginal distributions, which is required by MI.
- ii. The number of bins in the histogram for the KLD-based method is typically less than the number of bins for the MI-based method. The reason is that the former uses the joint class

histogram, whereas the later uses the joint intensity histogram.

- iii. As mentioned in Section 4.2, for the 2D-3D registration problem, the generation of binary projection image could be faster than the generation of projection image with a full intensity range.

The computational time of the registration process for both KLD-based and MI-based methods is listed in Table 3. All experiments done in this paper are carried out using a 1.7GHz Pentium IV PC, with 768M RAM. Based on the results of these two datasets, our method can run faster than the MI-based method by about 2-3 times.

Case No.	Methods	Time (sec)
1	KLD	115.019
	MI	335.547
2	KLD	86.081
	MI	216.947

Table 3: Comparisons of the computational time for the registration algorithms.

## 6. Conclusions

In this paper, we have proposed a new multi-modal image registration method based on the *a priori* knowledge of the expected joint class histogram. The difference between the observed and expected joint class histograms is measured by using Kullback-Leibler distance (KLD).

This objective function, KLD, is not only computationally more efficient than the MI-based similarity measure, but also produces comparable accuracy. It is particularly suitable for 2D-3D registration of DSA image and MRA volume. Further work will include a large number of datasets for validation of our registration algorithm.

## 7. Acknowledgements

HC and AC would like to acknowledge the support from RGC HKUST6209/02E, HKUST HIA, DAG 01/02 EG04 and SSRI01/02.EG22 grants. WW would like to acknowledge the support from NIH 5 P41 RR13218-01, NIH 5 P01CA67165-03 and NSF EEC 9731748 grants.

## 8. References

- [1] Albert C.S. Chung, William M. Wells III, and et al, "Multi-modal Image Registration by Minimising Kullback-Leibler Distance", *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, p. 525-532, 2002.
- [2] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.

- [3] S. Kullback. *Information Theory and Statistics*. Dover Publications, Inc., 1968.
- [4] W. Wells, P. Viola, and et al, "Multi-modal Volume Registration by Maximization of Mutual Information", *Medical Image Analysis*, 1996; 1:32-52.
- [5] A. Colignon et al, "Automated multi-modality image registration based on information theory", *Proceedings of Information Processing in Medical Imaging Conference*, p.263-274; Kluwer Academic Publisher, 1995.
- [6] P.P. Wyatt and J.A. Noble, "MAP MRF Joint Segmentation and Registration", *MICCAI'02*, p.580-587.
- [7] W.H. Press, S.A. Teukolsky, and et al, *Numerical Recipes in C*, 2nd Edition, Cambridge University Press, p.402-405,412-420; Kluwer Academic Publisher, 1995.
- [8] Y. Kita, D.L. Wilson and et al, "Real-time Registration of 3D Cerebral Vessels to X-ray Angiograms", *MICCAI*, p.1125-1133, 1997.
- [9] A. Liu, E. Bullitt and et al, "3D/2D Registration via skeletal near projective invariance in tubular objectives", *MICCAI*, p.952-963, 1998.
- [10] J. Feldmar, N. Ayache and et al, "3D-2D projective registration of free-form curves and surfaces", *INRIA*, Report # 2434, 1994.
- [11] A. Roche, G. Malandain, and et al, "Multimodal image registration by maximization of the Correlation Ratio", *INRIA*, Report # 3378, 1998.
- [12] R.A. McLaughlin, J.Hipwell, and et al, "A comparison of 2D-3D intensity-based registration and feature-based registration for neurointerventions", *MICCAI*, p.517-524, 2002.
- [13] R. Fisher, S. Perkins, and et al, <http://www.dai.ed.ac.uk/HIPR2/skeleton.htm>, 2000.
- [14] A. Watt, M. Watt, *Advanced Animation and Rendering Techniques: Theory and Practice*, Addison-Wesley, Reading, Massachusetts, 1992.