

Multi-modal Image Registration by Minimising Kullback-Leibler Distance

Albert C.S. Chung¹, William M. Wells III^{2,3},
Alexander Norbash², and W. Eric L. Grimson³

¹ Dept. of Computer Science, Hong Kong University of Science & Technology, HK

² Harvard Medical School, Brigham & Women's Hospital, Boston, MA USA

³ MIT Artificial Intelligence Laboratory, Cambridge, MA USA
achung@cs.ust.hk

Abstract. In this paper, we propose a multi-modal image registration method based on the *a priori* knowledge of the expected joint intensity distribution estimated from aligned training images. The goal of the registration is to find the optimal transformation such that the discrepancy between the expected and the observed joint intensity distributions is minimised. The difference between distributions is measured using the Kullback-Leibler distance (KLD). Experimental results in 3D-3D registration show that the KLD based registration algorithm is less dependent on the size of the sampling region than the Maximum log-Likelihood based registration method. We have also shown that, if manual alignment is unavailable, the expected joint intensity distribution can be estimated based on the segmented and corresponding structures from a pair of novel images. The proposed method has been applied to 2D-3D registration problems between digital subtraction angiograms (DSAs) and magnetic resonance angiographic (MRA) image volumes.

1 Introduction

A key issue in the medical imaging field is multi-modal image registration. As the use of co-registration packages spreads, the number of the aligned image pairs in image databases (either by manual or automatic methods) increases dramatically. These image pairs can serve as a set of training data, in which the statistical joint intensity properties can be observed and learned in order to acquire useful *a priori* knowledge for future registration tasks.

In this paper, we propose a multi-modal image registration method based on the *a priori* knowledge of the expected joint intensity distribution estimated from aligned training images. One of the key features is the use of the expected joint intensity distribution between two pre-aligned, training images as a reference distribution. The goal is to align any two images of the same or different acquisitions such that the expected distribution and the observed joint intensity distribution are well matched. In other words, the registration algorithm aligns two different images based on the expected outcomes. The difference between distributions is measured using the Kullback-Leibler distance (KLD), which is a

frequently used information theoretic similarity measure in the machine learning and information theory fields. The KLD value tends to zero when the two distributions become equal. The registration procedure is an iterative process, and is terminated when the KLD value becomes sufficiently small.

Experimental results in 3D-3D registration show that the KLD based registration algorithm is less dependent on the size of the sampling region than the Maximum log-Likelihood based method. We have also shown that, if manual alignment is unavailable, the expected joint intensity distribution can be estimated based on the segmented and corresponding structures from a pair of novel images. The proposed method has been applied to 2D-3D registration problems between DSAs and MRA image volumes.

2 Description of the Registration Algorithm

2.1 The Expected and Observed Joint Intensity Distributions

Expected joint intensity distribution: there are two ways of constructing the expected joint intensity distribution. Firstly, the joint distribution can be constructed by manual alignment, which can be done by experienced clinicians with the help of external or internal markers.

Let I_1 and I_2 be the intensity values of two training images of the same or different acquisitions, and X_1 and X_2 be their image domains respectively. Assume that the values of image pixels are independent of each other. Since the two images have been already aligned, samples of intensity pairs $\hat{\mathcal{I}} = \{i_1(x), i_2(x) | i_1 \in I_1, i_2 \in I_2\}$ can be drawn from I_1 and I_2 , where x are the pixel coordinates, $x \in X$ and $X = X_1 = X_2$. The expected joint intensity distribution $\hat{P}(I_1, I_2)$ can be approximated by either Parzen windowing or histogramming [1]. Histogramming is employed in this paper because the approach is computationally efficient, and the intensity histogram size is practical (the histogram has only 2 dimensions in this case). To achieve sub-voxel accuracy, histogram partial volume (PV) interpolation [7] can be used. A smooth histogram can be obtained by convolving with a Gaussian density function, given by

$$G_\psi(z) = (2n)^{-\frac{n}{2}} |\psi|^{-\frac{1}{2}} e^{-\frac{1}{2} z' \psi^{-1} z}, \quad (1)$$

where ψ is the co-variance of the Gaussian function and z can be a vector or scalar value.

If manual alignment is unavailable, a second method of constructing the expected joint intensity distribution is to perform segmentations separately in the two images, I_1 and I_2 , such that the internal anatomical structures are labelled. Let s_k , $k = 1 \dots M$, be the internal structures, where M represents the number of anatomical structures. Then, samples of intensity pairs $\hat{\mathcal{I}} = \{i_1(x), i_2(y) | i_1 \in I_1, i_2 \in I_2, x, y \in s_k, k = 1 \dots M\}$ can be drawn if x and y belong to the same structure s_k , where x and y are the pixel coordinates in X_1 and X_2 respectively. Similarly, the expected joint intensity distribution $\hat{P}(I_1, I_2)$ can be approximated by either Parzen windowing or histogramming.

Observed joint intensity distribution: given a new image pair with a hypothesized transformation T , samples of intensity pairs $\mathcal{I}_o = \{i_1(x), i_2(T(x)) | i_1 \in I_1, i_2 \in I_2\}$ can be drawn from I_1 and I_2 , where x are the pixel coordinates, $x \in \Omega$ and $\Omega \subset X_1 \cup X_2$. This means Ω represents a sampling domain that is equal to or inside $X_1 \cup X_2$. Note that the observed joint intensity distribution $P_o^T(I_1, I_2)$ is dependent on the values of the transformation T and changes during the registration. The Parzen windowing or histogramming approach can also be used to estimate the distribution P_o^T .

2.2 Kullback-Leibler Distance (KLD)

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

$$D(P_o^T || \hat{P}) = \sum_{i_1, i_2} P_o^T(i_1, i_2) \log \frac{P_o^T(i_1, i_2)}{\hat{P}(i_1, i_2)}. \quad (2)$$

According to [3,5], $D(P_o^T || \hat{P})$ has two important properties.

1. $D(P_o^T || \hat{P}) \geq 0$; and
2. $D(P_o^T || \hat{P}) = 0$ iff $P_o^T = \hat{P}$.

These properties show that, when the two images I_1 and I_2 are not perfectly registered, the values of KLD, D , will be non-zero and positive because the observed and expected joint intensity distributions are not equal, $P_o^T \neq \hat{P}$. On the other hand, if the images are well registered, then the value of KLD is equal to zero, i.e. $D = 0$.

2.3 Optimisation of the Transformation T

The goal of the registration is to find the optimal transformation \hat{T} by minimising 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 mutual information based registration method, which encourages the functional dependence between the two image random variables, I_1 and I_2 . 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 this paper, the value of KLD is minimised by Powell's method with a multi-resolution 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 minimisation) while other parameters are kept constant. The search step ∂T is relatively large in a coarse resolution and decreases as the resolution gets higher, ∂T is set to 2, 1 and 0.5mm in this paper (in Section 3.2). The iteration process stops when the change of KLD is sufficiently small (set 0.001 in this paper).

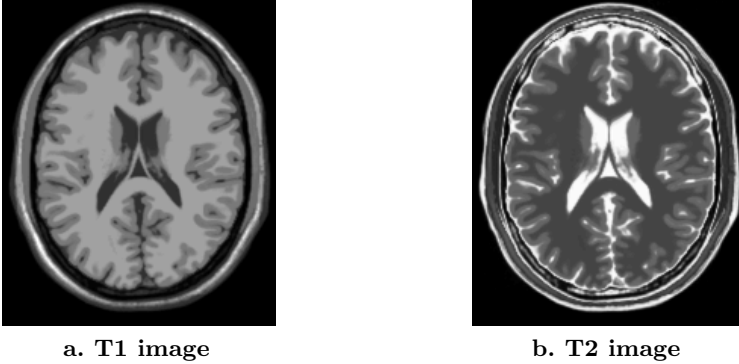


Fig. 1. (a) T1 and (b) T2 images.

3 Experimental Results

3.1 T1 – T2 (3D-3D) Registration

The T1 and T2 datasets are obtained from the BrainWeb Simulated Brain Database ($277 \times 241 \times 181$ voxels and $1 \times 1 \times 1mm^3$) [2], in which all the corresponding images have already been perfectly aligned and can be used as a testing platform for studying the performance of different objective functions. Maximum log-Likelihood (ML) [6] and Mutual Information (MI) [10] were compared with the KLD, their definitions are given by

$$ML = \sum_x \log \hat{P}(i_1(x), i_2(T(x))), \text{ and} \quad (4)$$

$$MI = \sum_{i_1, i_2} P_o^T(i_1, i_2) \log \frac{P_o^T(i_1, i_2)}{P_o^T(i_1)P_o^T(i_2)} \quad (5)$$

respectively, where $P_o^T(i_1)$ and $P_o^T(i_2)$ are the marginal distributions, x are the pixel coordinates, $x \in \Omega$ and $\Omega \subset X_1 \cup X_2$. One of the pairs of 2D T1 and T2 image slices is shown in Figs. 1a and 1b respectively, with their intensity values and image domains represented by I_1 and I_2 , and X_1 and X_2 respectively. Since these images in the datasets are aligned, the expected joint intensity distribution $\hat{P}(I_1, I_2)$ can be estimated based on the method described in Section 2.1 (only slices from positions 30 to 160 were used in order to avoid the inherent image artifacts in the dataset).

In order to study the performance of the objective functions, X_2 was shifted horizontally and rotated, whereas the position and orientation of X_1 were fixed. Given a transformation T , if any pixel x_2 in X_2 fell between the voxel positions of X_1 , then its corresponding intensity value i_1 was computed by linearly interpolating the values of its four neighbouring pixels in X_1 to achieve the sub-voxel accuracy. The observed joint intensity distribution P_o^T was then estimated according to Section 2.1. In this paper, the number of bins was set to 32 and the co-variance matrix ψ in Eq. 1 was a diagonal matrix $DIAG(\sigma^2, \sigma^2)$ and $\sigma^2 = 1$.

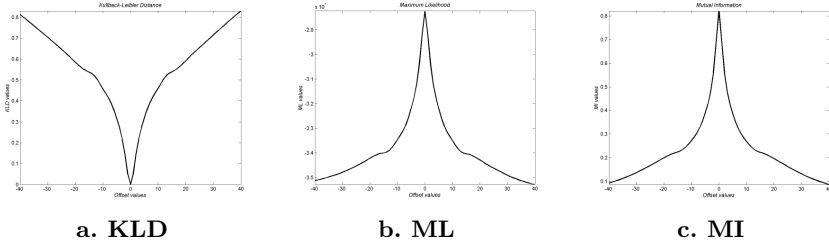


Fig. 2. T1-T2 registration performance analysis. T2 image was shifted horizontally. The offset values range from -40mm to 40mm .

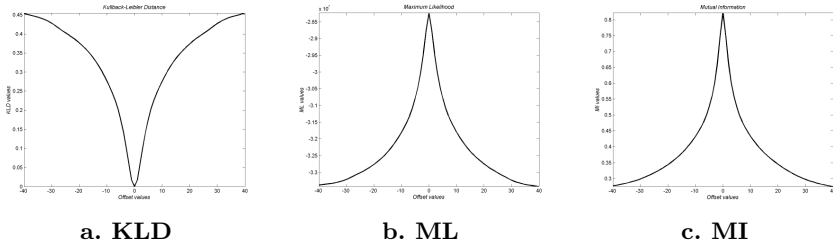


Fig. 3. T1-T2 registration performance analysis. T2 image was rotated. The offset values range from -40° to 40° .

We set $\Omega = X_2$ for ease of implementation in this paper. If x_2 fell outside the domain of image X_1 , then an arbitrary intensity value in the background of X_1 was assigned to i_1 . As plotted in Figs. 2 and 3, the performances of the three different measures (KLD, ML and MI) are comparable when the T2 image (X_2) was shifted horizontally between -40mm and 40mm , and was rotated between -40° and 40° .

However, it is also common to discard a sample $(i_1(x), i_2(T(x)))$ if it fell outside the overlapping region, i.e. $x \notin X_1 \cap X_2$. As shown in Fig. 4, when Ω was set to $X_1 \cap X_2$, the performance of ML was adversely affected when only samples drawn from the overlapping region were included in the calculation. As compared with ML, the figure shows that KLD and MI are less dependent on the size of the sampling region Ω . The major reason is that, from Eq. 4, the value of ML depends only on the observed samples x . Therefore, when the area of the overlapping region is small, fewer samples are obtained and thus the value of ML increases. In contrast, given the same set of observed samples, the value of KLD consists of the contributions of the observed samples and, most importantly, the penalties of the unobserved samples from the expected joint intensity distribution \hat{P} . Therefore, the entire distribution \hat{P} is utilised in the KLD measure. Finally, the value of MI depends mostly on the randomness of the observed samples. The decrease in overlapping area increases the sample randomness and, hence, the value of MI decreases.

In terms of computational efficiency, comparing Eq. 2 with Eq. 4, it is observed that, since KLD does not require the calculation of the marginal distri-

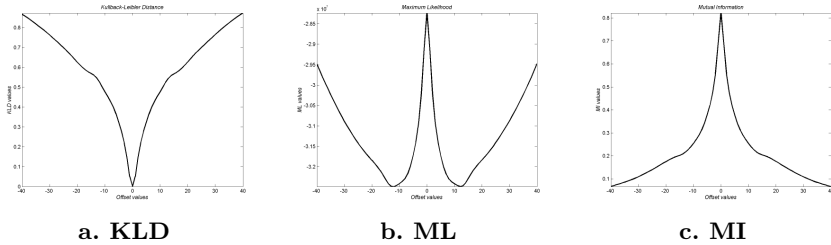


Fig. 4. T1-T2 registration performance analysis. T2 image was shifted horizontally. However, only samples, which fell in the overlapping region of the two images, were included in the calculations.

butions, $P_o^T(i_1)$ and $P_o^T(i_2)$, it can be more computationally efficient than MI. From Eq. 4, the efficiency of ML is directly proportional to the number of samples drawn. On the other hand, the efficiency of KLD is directly proportional to the product of the number of bins B_1 and B_2 partitioning I_1 and I_2 respectively. As such, the efficiencies of ML and KLD are related to different parameters, and their comparison is parameter dependent.

3.2 DSA - MRA (2D-3D) Registration

The proposed method was applied to 2D-3D registration problems and tested in two clinical datasets, which were acquired at the Department of Radiology, Brigham and Women's Hospital, Boston, USA. Each dataset consists of a pre-interventional 3D magnetic resonance angiographic (MRA) image volume ($256 \times 256 \times 60$ voxels and $0.78 \times 0.78 \times 1.3mm^3$), and a 2D digital subtraction angiogram (DSA) during the interventional treatments. Figs. 5a and 5d show the two cropped DSAs. The DSAs were distortion corrected using a distortion correction object with a uniform grid pattern [4].

A maximum intensity projection (MIP) of each MRA volume was generated using the projective geometry and ray casting method [8,11], in which there were six rigid body transformation parameters (three translational and three rotational). The initial transformations were obtained from the machine readings of the C-arm X-ray systems, as shown in Figs. 5d and 5h.

For each dataset, the expected joint intensity distribution was estimated based on the segmented and corresponding structures from the novel DSA and the initial non-registered MIP. These structures consist of vessel and background regions, in which each region was defined by a manually selected intensity range for the two datasets (more advanced methods can be applied but they are not the focus of this paper). The expected distribution \hat{P} was estimated by randomly drawing samples of the same structures from the DSA and MIP, as described in the Section 2.1. Then, the observed distribution P_o^T was generated during the registration and used to guide the rigid body transformation using the KLD measure, as defined in the Eqs. 2 and 3. The optimal transformation was searched using Powell's method with a multi-resolution strategy, as described in Section

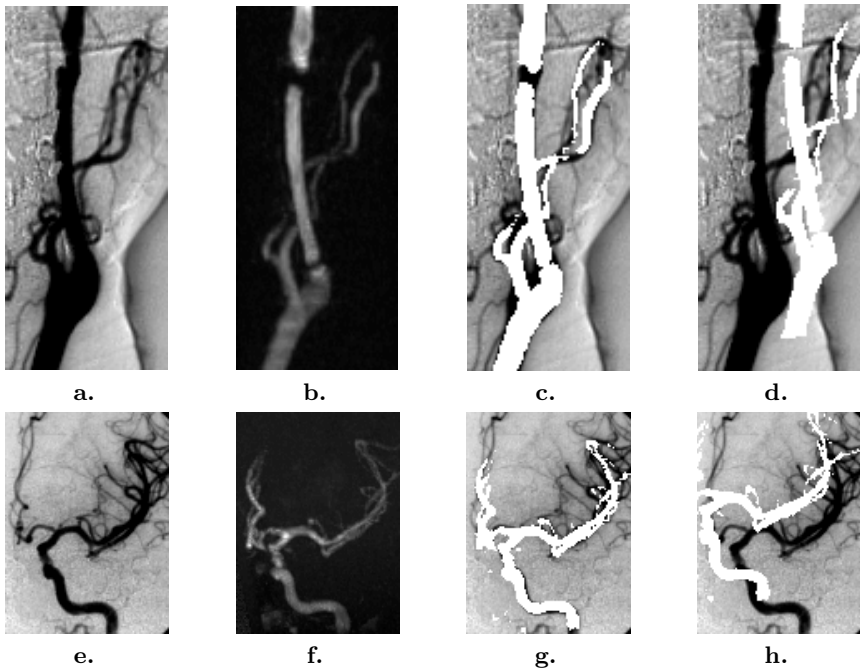


Fig. 5. 2D-3D registration results: (a,e) digital subtraction angiograms (DSA) (vessels are black in colour), (b,f) final image alignments, maximum intensity projections (MIP) of the magnetic resonance angiographic (MRA) image volumes (vessels are white in color and their intensity is directly proportional to the flow speed), (c,g) segmented MIPs are overlaid on their corresponding DSAs and (d,h) initial image alignments.

2.3. Figs. 5b and 5f show the MIPs of the registered MRA volumes and the results are promising. Segmented vessel regions of the MIPs are overlaid on the corresponding DSAs, as shown in Figs. 5c and 5g. Note that the remaining discrepancy between the DSA and MIP may be caused by (a) some vessels that are visible in one image and are not visible in another image due to different vessel delineation properties in different acquisitions and different regions of interest selected, (b) signal loss in the MRA images (e.g. turbulent or eddy flow), or (c) the geometric distortion due to the MR gradient field nonlinearity.

4 Summary and Conclusions

In this paper, we have proposed a multi-modal image registration method based on the *a priori* knowledge of the expected joint intensity distribution estimated from the aligned training images. The difference between the expected and observed joint intensity distributions is measured by the Kullback-Leibler distance (KLD), which has non-zero and positive value when there is any discrepancy be-

tween the two distributions. The KLD-based registration algorithm guides the transformations by minimising the KLD value until the two datasets are aligned.

The results based on T1-T2 (3D-3D) registration experiments show that, as compared with the Maximum log-Likelihood (ML) based registration method, the KLD-based registration algorithm is less dependent on the size of sampling region. In DSA-MRA (2D-3D) registration experiments, we have shown that the expected joint intensity distribution can also be estimated based on the segmented and corresponding structures (vessel and background regions) from the novel DSA and the initial non-registered MIP. The DSA-MRA registration results are promising and demonstrate the applicability of our method in 2D-3D registration. Future work will include a further validation of the proposed algorithm by applying it to a large number of datasets.

Acknowledgements

We would like to thank K. Rhode and D. Hawkes at Guy's Hospital, London, U.K. for sharing the DSA image distortion correction software. W. M. Wells III would like to acknowledge support from the NSF ERC grant (JHU Agreement #8810-274) and the NIH (grant #1P41RR13218).

References

1. C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford U. Press, 1995.
2. 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.
3. T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.
4. P. Haaker, E. Klotz, and et al. Real-time distortion correction of digital X-ray II/TV-systems: an application example for digital flashing tomosynthesis (DFTS). *International Journal of Cardiac Imaging*, 6(1):39–45, 1990-91.
5. S. Kullback. *Information Theory and Statistics*. Dover Publications, Inc., 1968.
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. 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.
9. W.H. Press, S.A. Teukolsky, and et al. *Numerical Recipes in C, 2nd Edition*. Cambridge University Press, 1992.
10. 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.
11. L. Zöllei. *2D-3D Rigid-Body Registration of X-Ray Fluoroscopy and CT Images*. MIT Masters Dissertation, 2001.