

Learning-Based Multi-Modal Rigid Image Registration By Using Bhattacharyya Distances

Ronald W. K. So and Albert C. S. Chung

Abstract—Multi-modal image registration is a momentous technology in medical image processing and analysis. In order to improve the robustness and accuracy of multi-modal rigid image registration, a novel learning-based dissimilarity function is proposed in this paper. This novel dissimilarity function is based on measuring the dissimilarity between the joint intensity distribution of the testing image pair and the expected intensity distributions, which is learned from a registered image pair, with Bhattacharyya distances. Then, the aim of the registration process is to minimize the dissimilarity function. Eight hundred randomized CT - T1 registrations were performed and evaluated by the Retrospective Image Registration Evaluation (RIRE) project. The experimental results demonstrate that the proposed method can achieve higher robustness and accuracy, as compared with a closely related approach and a state-of-the-art method.

I. INTRODUCTION

Multi-modal rigid image registration is one of essential technologies which can provide many applications for medical image processing and analysis [1]. The applications, such as information fusion of diagnostic images and image guided therapy, rely on a robust and accuracy rigid registration to provide a precise diagnosis and surgery. Therefore, robustness and accuracy are two major properties to describe a rigid image registration method. Given two images, the reference image (I_r) and the floating image (I_f), the objective of image registration is to estimate the optimal transformation (T^*) such that dissimilarity of the reference image and the transformed floating image (I_f^T) is minimized. The process is classified as rigid image registration if the transformation consists of translation and rotation only. The general formulation of rigid image registration is

$$T^* = \arg \min_T D(I_f^T(X), I_r(X)), \quad (1)$$

where X is the domain of the both images, D is the dissimilarity measure and T denotes a rigid transformation. The functionality of D is to evaluate the dissimilarity of reference image and transformed floating image. Sum of Absolute Difference (SAD), Sum of Squared Difference (SSD) and Mutual Information (MI) [2], [3] are examples of dissimilarity measure.

Recently, Gan *et al.* [4] proposed a learning-based dissimilarity measure by using Kullback-Leibler distance (KLD).

Ronald W.K. So is with Lo Kwee-Seong Medical Image Analysis Laboratory, Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Hong Kong cswk@cse.ust.hk

Albert C. S. Chung is the director of Lo Kwee-Seong Medical Image Analysis Laboratory and faculty of Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Hong Kong achung@cse.ust.hk

The KLD-based approach measuring the dissimilarity of the testing image pair by comparing the observed joint intensity distribution with the prior knowledge (i.e. the expected joint intensity distribution) learned from a pre-registered image pair. Because of the longer capture range of KLD-based approach, as compared with MI-based method, the experimental results demonstrated that the prior knowledge can improve the registration robustness. To achieve higher robustness and accuracy, a novel learning-based dissimilarity measure is proposed in this paper. The novel measure is based on two Bhattacharyya distances [5] and the prior knowledge of the joint, and marginal, intensity distributions. Eight hundred randomized CT - T1 registrations were performed and evaluated by the Retrospective Image Registration Evaluation (RIRE) project [6]. The results were compared with a closely related approach (the KLD-based approach) and a state-of-the-art method (the MI-based method). It is illustrated that the registration method with the proposed dissimilarity measure can achieve higher robustness or accuracy as compared with the MI-based and KLD-based registration methods.

II. METHODOLOGY

A. Bhattacharyya Distances

Bhattacharyya distance [5] can be used to measure the distance between two probability distributions. One of the applications of Bhattacharyya distance in computer science is image segmentation, which can be achieved by maximizing the Bhattacharyya distance between the intensity distribution of the pixels inside the segmented area and the intensity distribution of the pixels outside the segmented area [7]. In this paper, we propose a dissimilarity measure for rigid registration by utilizing two Bhattacharyya distances with the prior knowledge of the joint and marginal intensity distributions of a pre-aligned image pair. Let i_f and i_r respectively be the indices of the histogram bins of the floating and reference images. We define that $P_o^T(i_f, i_r)$ and $\hat{P}(i_f, i_r)$ are the observed joint intensity distribution and expected joint intensity distribution respectively. Then, our first Bhattacharyya distance, named BD_1 , is formulated as

$$BD_1 = -\log \int \sqrt{P_o^T(i_f, i_r) \hat{P}(i_f, i_r)} d_{i_f} d_{i_r}. \quad (2)$$

BD_1 measures the distance between the observed and expected joint intensity distributions. Since the expected joint intensity distribution is the prior knowledge learned from a pre-aligned image pair, it describes the underlying statistical relationship that we expect to observe on joint intensity distribution of the testing image pair when the images are

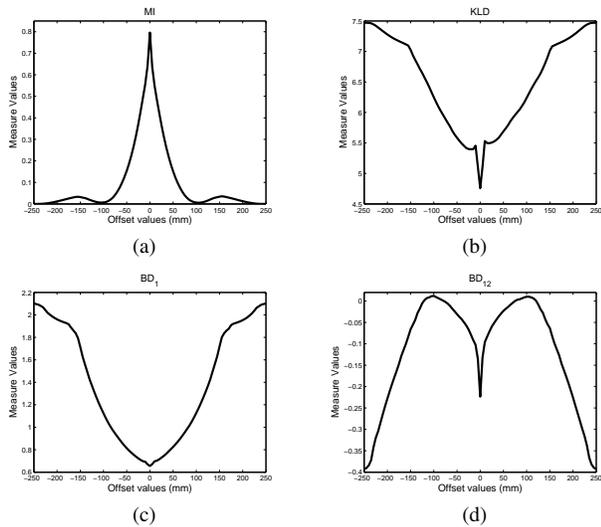


Fig. 1. The translation probes of (a) MI , (b) KLD , (c) BD_1 and (d) BD_{12} values respectively. The zero offset representing the position of ground truth.

properly aligned. Therefore, by minimizing BD_1 according to the transformation T , the floating image can be efficiently registered to the reference image. Figs. 1(a)-(c) show the translation probes of MI , KLD and BD_1 in which pt-001 and pt-005 of the RIRE project [6] were used as the testing images pair and the training image pair (for KLD and BD_1) respectively. The zero offset of the translation probes represents the position of the ground truth transformation (please refer to Section III-A for the details of the ground truth transformation). From the figure, we can notice that there are two observable local optima in the translation probes of MI and KLD , and such local optima do not exist in BD_1 . Therefore, the registration approach based on BD_1 should be able to provide higher robustness as compared with MI -based and KLD -based methods. This expectation is matched with the observation described in [8] that Bhattacharyya distance usually gives better results than the Kullback-Leibler distance. We will experimentally prove that the observation is also true for the rigid image registration in Section III.

However, as shown in Fig. 1, the translation probe of BD_1 around the zero offset is not as sharp as that of MI and KLD . It will affect the accuracy of the registration results. Therefore, we introduce the second Bhattacharyya distance which is denoted as BD_2 . The formulation of BD_2 is

$$BD_2 = -\log \int \sqrt{P_o^T(i_f, i_r)(\hat{P}(i_f)\hat{P}(i_r))} d i_f d i_r, \quad (3)$$

where $\hat{P}(i_f)$ and $\hat{P}(i_r)$ are expected marginal distributions of the floating image and the reference image respectively. BD_2 measures the distance between the observed joint intensity distribution and the product of the expected marginal distributions ($\hat{P}(i_f)\hat{P}(i_r)$). When the observed joint intensity distribution of the testing image pair is equal (or very close) to the product of the expected marginal distributions, the intensity values in the testing image pair can be recognized as statistically independent. Therefore, by maximizing BD_2 , the

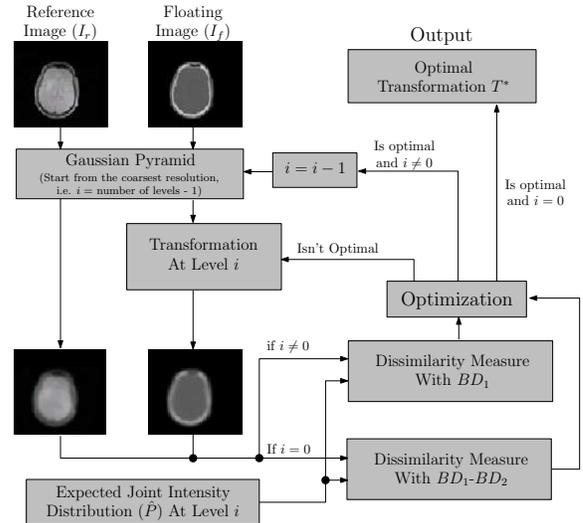


Fig. 2. Flow chat of the proposed rigid registration framework.

transformation (T) will be optimized such that the observed joint intensity distribution of the testing image pair is pushed away from the production of the expected marginal distributions. This can improve the accuracy of the registration results. Then, our proposed dissimilarity measure, which consists of BD_1 and BD_2 , is defined as

$$BD_{12} = BD_1 - BD_2. \quad (4)$$

Fig. 1(d) shows the translation probe of BD_{12} . As compared with BD_1 , we can notice that the translation probe of BD_{12} becomes very sharp in the range around the zero offset. Since the capture range of BD_{12} is much less than that of BD_1 , We suggest that the registration process should only use BD_1 as the dissimilarity measure in the early state (e.g. the coarser resolution levels of the multi-resolution registration), and BD_{12} is adopted when the testing image pair is roughly aligned (e.g. the finest resolution level of the multi-resolution registration).

B. Registration Framework

As discussed in the last section, the proposed dissimilarity measure is suggested to utilize with multi-resolution registration approach [2]. Fig. 2 is the flow chat of the multi-resolution registration framework with the proposed dissimilarity measure. The registration starts at the coarsest resolution level in which the down-sampled reference and floating images are used as the input. A rough estimation of the optimal transformation derived in this level is passed to the successive higher resolution level as the initial values for optimization. The process continues until the level of original image resolution is reached. Then, the solution of this finest resolution level is output as the optimal transformation. For the learning-based methods, in order to capture the prior knowledge about the intensity relationship of the image pair at different resolutions, the intensity distributions are estimated at each resolution. The dissimilarity measure in each level is minimized by the Powell's method [9], which

TABLE I

THE SUCCESS RATES OF *MI*-BASED, *KLD*-BASED AND *BD*-BASED METHODS TOGETHER WITH THE MEANS AND STANDARD DEVIATIONS OF THE MEDIAN TRE VALUES OF THE SUCCESSFUL REGISTRATIONS FOR DIFFERENT TESTING IMAGE PAIRS. THE VALUES OF *KLD*-BASED AND *BD*-BASED METHODS ARE AVERAGED AMONG ALL THE TRAINING IMAGE PAIRS.

Dataset	<i>MI</i>		<i>KLD</i>		<i>BD</i>	
	success%	mean \pm sd	success%	mean \pm sd	success%	mean \pm sd
pt-001	58.00%	0.6906 \pm 0.1159	87.00%	1.8647 \pm 0.4845	96.50%	0.6385 \pm 0.1345
pt-003	62.00%	0.7725 \pm 0.0373	74.00%	2.8174 \pm 0.4168	94.50%	1.0025 \pm 0.2906
pt-005	58.00%	0.6051 \pm 0.2043	88.50%	2.2618 \pm 0.4455	92.00%	0.9918 \pm 0.6258
pt-007	63.00%	1.3151 \pm 0.1496	85.50%	2.2669 \pm 0.5435	94.00%	1.1247 \pm 0.3570
Total	60.25%	0.8458 \pm 0.1268	83.75%	2.3027 \pm 0.4726	94.25%	0.9394 \pm 0.3520

does not require to calculate the gradient values as it exploits Brent's 1D line minimizations to iteratively search along a set of directions. Hereafter, we denote that the registration methods which utilize this framework with *MI*, *KLD* and the proposed dissimilarity measure as *MI*-based method, *KLD*-based method and *BD*-based method respectively. For the proposed *BD*-based method, since BD_2 will affect the capture range of BD_1 , it is suggested that BD_2 should only be adopted with BD_1 (i.e. BD_{12}) in the finest resolution level to improve the registration accuracy. In short, BD_{12} is used as the dissimilarity measure in the finest resolution level and BD_1 is used in other levels.

III. EXPERIMENTAL RESULTS

A. Image Datasets, Ground Truth and Setup of Experiments

Eight hundred randomized rigid image registration experiments on four sets of clinical brain X-ray computed-tomography (CT) and T1-weighted MRI images (T1) were performed to evaluate the proposed dissimilarity measure. The four testing image pairs are the odd-numbered (pt-001, pt-003, pt-005 and pt-007) CT and T1 images from the Retrospective Image Registration Evaluation (RIRE) Project [6]. Meanwhile, the even-numbered (pt-002 and pt-004) CT - T1 image pairs of RIRE project were used as the training images. Note that all T1 images used in the experiments have been rectified by the RIRE project. Therefore, images of pt-006 were excluded in the experiments as RIRE project does not provide the rectified T1 image of pt-006. CT images were the floating images and T1 images were the reference images for all experiments. The dimension of CT image volumes are $512 \times 512 \times z$ where z equal to 28, 29, 34, 27, 33 and 28 for pt-001, pt-002, pt-003, pt-004, pt-005 and pt-007 respectively, and the voxel size of all CT image volumes is $0.65 \times 0.65 \times 4mm^3$. For T1 image volumes, the image dimension of pt-001, pt-002, pt-003, pt-004, pt-005 and pt-007 are $256 \times 256 \times z$ where z equal to 26, 26, 20, 20, 26 and 26 respectively, and the voxel size is $1.25 \times 1.25 \times 4mm^3$ for all T1 image volumes.

The RIRE project is exploited to evaluate the registration results in which the accuracies are reported as the target registration errors (TREs) in five to ten (depends on the testing image pair) volumes of interest (VOI) [6]. Since we do not know the gold standard [6] of the RIRE project, a set of transformations (one for each image pair) were defined as the ground truth transformations for building the

expected intensity distributions and generating the initial alignments (will be discussed later in this section). In this paper, the ground truth transformations were the optimal transformations which were returned by applying the four-resolutions *MI*-based method to the six RIRE image pairs. All the ground truth transformations were evaluated by the RIRE project and had median TRE value less than $1mm$. The median TRE values of the ground truth transformations for pt-001, pt-002, pt-003, pt-004, pt-005 and pt-007 were $0.9692mm$, $0.5850mm$, $0.7577mm$, $0.7069mm$, $0.9813mm$ and $0.3337mm$ respectively.

In our experiments, we compared the *BD*-based method with the *KLD*-based method (the closely related approach) and the *MI*-based method (the state-of-the-art method) with four resolution levels. The compared methods were applied to register the testing image pairs (pt-001, pt-003, pt-005 and pt-007) with one hundred randomize initial alignments for each image pair. For the *KLD*-based and *BD*-based methods, two sets of expected intensity distributions learned from training image pairs (pt-002 and pt-004) were utilized for the registration of each testing image pair. Thus, in total, eight hundred trials (two hundred trials on each testing image pair) were performed by the proposed *BD*-based method for this evaluation. The randomize initial alignments of each testing image pair were generated by applying one hundred random rigid transformations on the corresponding ground truth transformation. The X, Y and Z translational parameters of the random rigid transformations were respectively drawn between $[-150, 150]mm$, $[-150, 150]mm$ and $[-70, 70]mm$, while all the rotational parameters were drawn between $[-0.52, 0.52]$ radians, i.e. $[-30, 30]$ degrees. These ranges were chosen based on a criterion of brain structures in CT and T1 images have at least 10% overlapping region.

All programs used in this paper were implemented based on an open source library, The Insight Segmentation and Registration Toolkit (ITK) [10]. The experiments were carried out on a PC with a CoreTM 2 Duo 2.33GHz CPU and 2GB RAM.

B. Registration Results

For each registration result, the RIRE project returned a TRE for each VOI. Depending on the testing image pair, the number of VOI varied from five to ten. In the experiments, we defined that a registration was successful if the median value of the corresponding TREs was smaller than $4mm$

TABLE II

THE SUCCESS RATES OF *KLD*-BASED AND *BD*-BASED METHODS TOGETHER WITH THE MEANS AND STANDARD DEVIATIONS OF THE MEDIAN TRE VALUES OF THE SUCCESSFUL REGISTRATIONS FOR DIFFERENT TESTING IMAGE PAIRS BY TAKING DIFFERENT TRAINING IMAGE PAIRS.

Dataset	Training image pair											
	All				pt-002				pt-004			
	<i>BD</i>		<i>KLD</i>		<i>BD</i>		<i>KLD</i>		<i>BD</i>		<i>KLD</i>	
success%	mean±sd	success%	mean±sd	success%	mean±sd	success%	mean±sd	success%	mean±sd	success%	mean±sd	
pt-001	96.5%	0.64±0.13	87%	1.86±0.48	97%	0.60±0.13	86%	1.65±0.44	96%	0.68±0.14	88%	2.08±0.53
pt-003	94.5%	1.00±0.29	74%	2.82±0.42	94%	0.91±0.18	85%	2.50±0.26	95%	1.10±0.40	63%	3.13±0.57
pt-005	92%	1.00±0.63	88.5%	2.26±0.45	96%	0.67±0.36	85%	1.87±0.30	88%	1.31±0.89	92%	2.65±0.59
pt-007	94%	1.12±0.35	85.5%	2.27±0.54	98%	0.94±0.31	83%	2.03±0.63	90%	1.31±0.41	88%	2.50±0.46
Total	94.25%	0.94±0.35	83.75%	2.30±0.47								

($4mm$ was chosen as it was the largest voxel dimension of the images). Table I lists the success rates, together with the means and standard deviations of the median TRE values for the successful registrations, of the *MI*-based, *KLD*-based and *BD*-based methods. Note that the results of *KLD*-based and *BD*-based methods shown in Table I are averaged among the two training image pairs. For the individual training image pair, the details of the experimental results can be found in the Table II. The last row of the both tables (i.e. the row named as “Total”) summarizes the results from all the testing image pairs for the corresponding method.

Since the experiments of each testing image pair were performed under different initial alignments, the success rates can illustrate the robustness of compared methods. From Table I, we can notice that *BD*-based method gives a higher success rates as compared with *MI*-based and *KLD*-based methods, in which the averaged success rates of *MI*-based, *KLD*-based and *BD*-based methods among all the experiments are 60.25%, 83.75% and 94.25% respectively. Note that the dissimilarity measure is the only different between the *MI*-based, *KLD*-based and *BD*-based methods. Thus, the higher success rate of the *BD*-based method demonstrates that the proposed *BD*-based dissimilarity measure is more robust than *MI* and *KLD*. It is also proved by Table II that the *BD*-based dissimilarity measure can achieve higher robustness than the *KLD* under different training image pairs. For the registration accuracy, as listed in the Table I, the averaged means of median TRE values of *MI*-based, *KLD*-based and *BD*-based methods are 0.8458, 2.3027 and 0.9394 respectively. This shows that the accuracy of *BD*-based method is higher than that of *KLD*-based method and is comparable to the *MI*-based method. In short, the *BD*-based method has higher robustness and similar accuracy as compared to the *MI*-based method, while it outperforms the *KLD*-based method in terms of both robustness and accuracy (noted that *KLD*-based method is more robust than the *MI*-based method as demonstrated in the tables). Meanwhile, these experimental results are matched with the observation described in [8] that Bhattacharyya distance usually gives better results than the Kullback-Leibler distance.

IV. CONCLUSIONS

This paper has proposed a novel dissimilarity measure which is based on two Bhattacharyya distances and the

prior knowledge learned from the intensity distributions of a pre-aligned image pair. The first Bhattacharyya distance measures the distance between the observed joint intensity distribution and the expected joint intensity distribution. By minimizing this distance, the testing image can be registered efficiently. The second Bhattacharyya distance measures the distance between the observed joint intensity distribution and the product of the expected marginal intensity distributions. Since the testing image pair becomes statistically independent if the observed joint intensity distribution is equal to the product of the expected marginal intensity distributions, maximizing this second Bhattacharyya distance can improve the accuracy of the registration results. Eight hundred randomized rigid image registrations were performed by using the registration method with the proposed dissimilarity measure. The results are compared with a closely related approach and a state-of-the-art method. It is demonstrated that the proposed dissimilarity measure can provide higher robustness among the compared methods, in which the accuracy outperforms the closely related approach and comparable with the state-of-the-art method.

REFERENCES

- [1] J. Maintz and M. Viergever, “A survey of medical image registration,” *Medical Image Analysis*, vol. 2, no. 1, pp. 1–16, 1998.
- [2] W. Wells, P. Viola *et al.*, “Multi-Modal Volume Registration by Maximization of Mutual Information,” *Medical Image Analysis*, vol. 1, no. 1, pp. 35–51, 1996.
- [3] J. Pluim, J. Maintz, and M. Viergever, “Mutual-information-based registration of medical images: a survey,” *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.
- [4] R. Gan, J. Wu *et al.*, “Multiresolution image registration based on Kullback-Leibler distance,” *MICCAI LNCS*, vol. 3216, pp. 599–606, 2004.
- [5] A. Bhattacharyya, “On a measure of divergence between two statistical populations defined by their probability distributions,” *Bulletin of the Calcutta Mathematical Society*, vol. 35, pp. 99–109, 1943.
- [6] J. West, J. M. Fitzpatrick *et al.*, “Comparison and evaluation of retrospective intermodality brain image registration techniques,” *J. Comput. Assist. Tomogr.*, vol. 21, no. 4, pp. 554–566, 1997.
- [7] O. Michailovich, Y. Rathi and A. Tannenbaum, “Image Segmentation Using Active Contours Driven by the Bhattacharyya Gradient Flow,” *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2787–2801, 2007.
- [8] T. Kailath, “The Divergence and Bhattacharyya Distance Measures in Signal Selection,” *IEEE Transactions on Communication Technology*, vol. 15, no. 1, pp.52–60, 1967.
- [9] W. H. Press, S. A. Teukolsky *et al.*, “Numerical Recipes in C++,” *Cambridge University Press*, second edition edition, 2002.
- [10] L. Ibanez, W. Schroeder *et al.*, “The ITK Software Guide”, 1st edn. kitware edition, 2003.