# MULTI-MODAL NON-RIGID IMAGE REGISTRATION BASED ON SIMILARITY AND DISSIMILARITY WITH THE PRIOR JOINT INTENSITY DISTRIBUTIONS

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

Lo Kwee-Seong Medical Image Analysis Laboratory,
Department of Computer Science and Engineering,
The Hong Kong University of Science and Technology, Hong Kong

## ABSTRACT

Multi-modal non-rigid image registration is widely used in different areas, including medical image analysis and image processing. In this paper, we introduce a new learning-based method for non-rigid image registration. The proposed method is based on $a\ priori$ knowledge of the joint intensity distribution of a pre-aligned image pair. The similarity and dissimilarity of the expected and observed joint intensity distributions are measured by two Kullback-Leibler distances (KLD). Free-Form Deformation (FFD) is employed as the transformation model along with the L-BFGS-B optimizer. The derivatives of KLDs are derived to work with the L-BFGS-B optimizer. Moreover, we have tested our method with CT-T1 image pairs and compared the results obtained by using the mutual information based FFD and the conventional KLD based FFD. The experimental results show that our method gives remarkable improvement on the registration quality.

***Index Terms***— Non-rigid image registration, Kullback-Leibler distances

## 1. INTRODUCTION

Non-rigid image registration is an essential technology of medical image analysis [1]. According to the nature of transformation, image registration can broadly be categorized into two classes: rigid and non-rigid. 3D Rigid image registration involves the optimization of transformation with six degrees of freedom which consists of three angles of rotation and three translation vectors. Different from rigid image registration, transformation of non-rigid image registration has higher degrees of freedom which allows all voxels to transform freely if no constraint is added to the deformation to keep the transformation smooth. Therefore, due to the requirements of smoothness and high degree of freedom, non-rigid image registration is a challenging research problem.

Given a floating image $I_f$ and a reference image $I_r$, the goal of image registration is to find an optimal transformation $t^*$ that realigns $I_f$ to $I_r$. A dissimilarity function $C(I_r, t(I_f))$ is usually used to measure the dissimilarity of the transformed floating image $t(I_f)$ and the reference image. The whole process is generally formulated as:

$$t^* = \arg\min_t (C(I_r, t(I_f)) + \lambda S(t)), \tag{1}$$

where $S(t)$ is a penalty term that encourages the smoothness of the deformation field and $\lambda$ is a positive constant that governs the strength of the penalty.

In this paper, we mainly consider the dissimilarity function $C(I_r, t(I_r))$. Sum of Squared Differences (SSD), Sum of Absolute Differences (SAD), Mutual Information (MI) and Normalized Mutual Information (NMI) are widely used as dissimilarity function in both rigid and non-rigid registration tasks. Those four functions merely depend on the intensity of the testing image pair. To incorporate the intensity relationship learned from the training image pair into non-rigid image registration, we extend the conventional Kullback-Leibler distance (KLD) based method [2], which was originally proposed for rigid image registration. This conventional KLD based method has been utilized in rigid [2] and non-rigid registrations [3, 4]. Recently, for rigid registration, Chung $et\ al.$ [5] proposed a novel KLD based dissimilarity measure and it was demonstrated that the combined use of the similarity and dissimilarity measures improves the robustness and the quality of the rigid registration tasks. In this paper, the derivatives of the similarity measure [2] and the dissimilarity measure [5] are derived. We exploit the derivatives along with the "Limited memory Broyden Fletcher Goldfarb Shannon minimization with simple Bounds"(L-BFGS-B) optimizer [7] to perform Free-Form Deformation (FFD) [6] based non-rigid registration. To the best of our knowledge, it is the first time to apply the two similarity and dissimilarity measures together with their derivatives into non-rigid registration (details will be discussed in the next section). The experimental results show that our method gives better registration quality than MI based FFD and the conventional KLD based FFD.

The organization of this paper is as follows. In Section 2, we will discuss our registration framework, including the transformation model, our dissimilarity function and the derivatives. Without lost of generality, our discussion is based on three dimensional cases. The experimental results of our method will be presented in Section 3. Finally, we conclude our paper with future research in Section 4.

## 2. METHODOLOGY

### 2.1. Transformation Model

We use Free-Form Deformation (FFD) described in [6] as our transformation model along with the L-BFGS-B optimizer. In FFD, a set of equal spacing control points is defined on the floating image. Those control points can be moved freely in 3D space during the registration process. The deformations of other points are calculated by B-Spline interpolation on the displacements of nearby control points. With more control points, the process has better ability to handle higher frequency local deformation, but the deformation field will be rougher. Therefore, we should keep a balance on the accuracy and the smoothness of the deformation field by controlling the number of control points.

### 2.2. Kullback-Leibler Distances

With $a\ priori$ knowledge of the joint intensity distribution $\hat{P}(i_f, i_r)$ between floating and reference images, Chung $et\ al.$ [5] proposed a learning-based rigid registration method with a novel dissimilarity function, which is constructed by three Kullback-Leibler distances. Those three KLDs are denoted as $D_1$, $D_2$ and $D_3$. In this paper, we assume that all floating images are registered to the reference image by a rigid transformation and then an affine transformation before applying our proposed method. Under this assumption, we only employ $D_1$ and $D_2$ into our dissimilarity function, and $D_3$ is not included since it is useful only if the overlap region of the floating image and the reference image is small. The formulations of $D_1$ and $D_2$ are:

$$D_1(P_o^t(i_f, i_r)||\hat{P}(i_f, i_r)) = \sum_{i_f, i_r} P_o^t(i_f, i_r) \log \frac{P_o^t(i_f, i_r)}{\hat{P}(i_f, i_r)}, \quad (2)$$

$$D_2(P_o^t(i_f, i_r)||\hat{P}(i_f, i_r)) = \sum_{i_f, i_r} P_o^t(i_f, i_r) \log \frac{P_o^t(i_f, i_r)}{\hat{P}(i_f)\hat{P}(i_r)}, \quad (3)$$

where $i_f$ and $i_r$ are the indices of the histogram bins of the floating image and the reference image respectively, $\hat{P}(i_f, i_r)$ is the expected joint intensity distribution, $P_o^t(i_f, i_r)$ is the observed joint intensity distribution, and $\hat{P}(i_f)$ and $\hat{P}(i_r)$ are the expected marginal distributions of the training image pair. Note that $D_1$ is the conventional KLD used in [2, 3, 4, 5] and the product of expected marginal distributions, i.e. $\hat{P}(i_f)\hat{P}(i_r)$, represents the distribution when the intensity values of the input images are statistically independent.

Fig. 1 illustrates the functionalities of $D_1$ and $D_2$. The cross symbols representing $P_o^t(i_f, i_r)$ of different image pairs before the registration. The similarity measure $D_1$ and the dissimilarity measure $D_2$ like a pull force and a push force respectively. When minimizing $D_1$, $t$ is optimized to pull the observed joint distributions toward the expected joint distribution. When maximizing $D_2$, $t$ is optimized to push $P_o^t(i_f, i_r)$ away the condition that the testing image pair are statistically independent. By utilizing the pull force and the push force together, we can decrease the possibility that the registration process is trapped in local minimum and also improve the registration quality. We denote the resulting dissimilarity function as $D_{12} = D_1 - D_2 = \sum_{i_f, i_r} P_o^t(i_f, i_r) \log \frac{\hat{P}(i_f)\hat{P}(i_r)}{\hat{P}(i_f, i_r)}$. In the meanwhile, the objective of our non-rigid registration process becomes finding the optimal transformation,

$$\mu^* = \arg\min_{\mu}(D_{12}(P_o^{t(\mu)}(i_f, i_r)||\hat{P}(i_f, i_r))). \quad (4)$$

where $\mu = [\mu_1\ \mu_2\ \ldots\ \mu_{3n}]^T$ are $3n$ independent transformation parameters of $t$ under FFD and $n$ is the number of control points. Note that the penalty term $S(t)$ is omitted as $t$ is restricted by $\mu$ in FFD.

### 2.3. Derivatives of KLDs

Some optimizers, for instance, L-BFGS-B which is employed in this paper, require the analytical form of gradient of the energy function. These optimizers usually converge faster than other optimizers that do not require the analytical form of gradient. It is because the extra information given by the gradient can drive them to the global minimum or strong local minimum.

To compute the derivatives of $D_1$ and $D_2$ efficiently, we apply the Parzen window model using separable B-Spline kernels [4, 8, 9] to model the joint intensity distributions.

$$\begin{aligned} P_o^{t(\mu)}(i_f, i_r) &= \frac{1}{\text{card}(V)} \sum_{\vec{x} \in V} \beta^{(0)} \left( i_r - \frac{I_r(\vec{x}) - I_r'}{\Delta b_r} \right) \\ &\quad \times \beta^{(3)} \left( i_f - \frac{I_f(t(\vec{x}|\mu)) - I_f'}{\Delta b_f} \right), \end{aligned} \quad (5)$$

where $\beta^{(0)}$ is a zero-order B-Spline kernel used for the reference image, $\beta^{(3)}$ is a cubic B-Spline kernel used for the floating image, $V$ is a subset of voxels involved in the computation of the joint intensity distributions, $I_r'$ and $I_f'$ are minimum intensity values of the input images, and $\Delta b_r$ and $\Delta b_f$ are intensity ranges of each histogram bin of $I_f$ and $I_r$ respectively. The $k^{th}$ derivative of Eq. 5 can be obtained by applying the chain rule [8],

$$\begin{aligned} \frac{\partial P_o^{t(\mu)}(i_f, i_r)}{\partial \mu_k} &= \frac{1}{\text{card}(V)\Delta b_f} \sum_{\vec{x} \in V} \beta^{(0)} \left( i_r - \frac{I_r(\vec{x}) - I_r'}{\Delta b_r} \right) \\ &\left( \frac{\partial \beta^{(3)}(u)}{\partial u}\Big|_{u = i_f - \frac{I_f(t(\vec{x}|\mu)) - I_f'}{\Delta b_f}} \right) \left( \frac{\partial I_f(t(\vec{x}|\mu))}{-\partial(t(\vec{x}|\mu))} \right)^T \frac{\partial t(\vec{x}|\mu)}{\partial \mu_k}. \end{aligned} \quad (6)$$

There are three derivative terms in Eq. 6. The first derivative term can be differentiated explicitly by a subtraction of two second-order B-Spline kernels [8]. The second derivative term is the gradient of the floating image on a point that is transformed by $t(\mu)$. This term can be computed by finite element methods. The last derivative term is the derivative of the transformation. As the transformation model FFD is linear with respect to the transformation parameter $\mu_k$, the last term is the coefficient of the parameter.

By Eq. 5 and Eq. 6, the derivative of a dissimilarity function $C(\cdot)$ is given as $\nabla C = [\frac{\partial C}{\partial \mu_1}\ \frac{\partial C}{\partial \mu_2}\ \cdots\ \frac{\partial C}{\partial \mu_{3n}}]^T$. For $D_1$ and

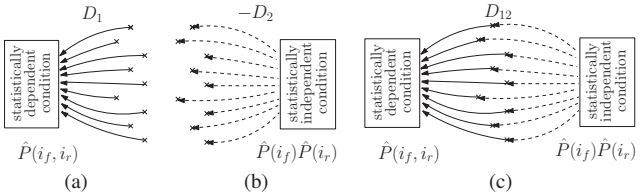


**Fig. 1**. Illustration of the functionalities of (a) $D_1$, (b) $D_2$ and (c) $D_{12}$. The cross symbols representing the observed joint distributions of different image pairs before the registration.

$D_2$, $\frac{\partial D_1}{\partial \mu_k} = \sum_{i_f,i_r} \log \frac{P_o^{t(\mu)}(i_f,i_r)}{\hat{P}(i_f,i_r)} \cdot \frac{\partial P_o^{t(\mu)}(i_f,i_r)}{\partial \mu_k}$ and $\frac{\partial D_2}{\partial \mu_k} = \sum_{i_f,i_r} \log \frac{P_o^{t(\mu)}(i_f,i_r)}{\hat{P}(i_f)\hat{P}(i_r)} \cdot \frac{\partial P_o^{t(\mu)}(i_f,i_r)}{\partial \mu_k}$ respectively. Therefore, the $k^{th}$ derivative of $D_{12}$ is:

$$\frac{\partial D_{12}}{\partial \mu_k} = \sum_{i_f,i_r} \log \frac{\hat{P}(i_f)\hat{P}(i_r)}{\hat{P}(i_f,i_r)} \cdot \frac{\partial P_o^{t(\mu)}(i_f,i_r)}{\partial \mu_k}. \quad (7)$$

## 3. EXPERIMENTAL RESULTS

All images used in experiments were obtained from the Retrospective Image Registration Evaluation (RIRE) Project [10] and have been normalized with the intensity within 0 and 255. In order to compare our $D_{12}$ based FFD with the MI based FFD and the conventional KLD based FFD [2], we have done 21 experiments, 7 experiments on each method. All experiments were 3D to 3D multi-modal non-rigid registration and run under the same configuration where $15 \times 15 \times 15$ control point grid was used in FFD. We chose MR-T1 of pt-001 as reference image, and the pre-deformed CT of pt-001 and CT of other 6 patients (pt-002 to pt-007) were floating images. All image pairs were pre-registered with a rigid transformation followed by an affine transformation. The leftmost columns of Figs. 2 shows slices of checker boards of image pairs used in the experiments before non-rigid registration was performed. For the experiments of the $D_{12}$ based FFD and the conventional KLD based FFD, we chose pre-aligned CT-T1 image pair of pt-002 as the training image pair. All programs used in experiments were implemented by using an open source library, Insight Toolkit ITK [11]. For MI and its derivative, we used the implementation described in [8] which was provided by ITK. All experiments were performed on a PC with an Intel 2.13GHz dual-core CPU with 3GB RAM.

Columns 2 to 4 of Figs. 2 shows the registration results of different CT-T1 image pairs. The images shown were slices of checker boards which were generated by the reference image (T1 of pt-001) and the registration results of image pairs. As described in Section 2.1, we used FFD as the transformation model. Since only control points can be moved freely, whereas the movements of other voxels are calculated by their neighborhood control points via B-Spline function, the checker boards of the result images are similar. However, it is easy to notice that our method gives obvious improvements on registration quality. In general, inside the brain region, the bright structures of CT should be mapped to the dark structures of MR-T1. To show the differences, we have circled observable misaligned regions. Among all cases, our dissimilarity function $D_{12}$ can obtain more precise results than MI and conventional KLD, especially in pt-001, pt-004 and pt-005. On the other hand, the results of MI and conventional KLD are comparable. For pt-001, pt-005 and pt-006, the results of MI are more precise than the conventional KLD whereas conventional KLD can work better in other cases. Note that $D_{12}$ is composed of the conventional KLD, i.e. the similarity measure $D_1$, and the newly added KLD, i.e. the dissimilarity measure $D_2$. Therefore, the contribution of $D_2$ can be proved by comparing the results of $D_{12}$ and conventional KLD. It is observed that $D_2$ improves the registration quality especially in pt-001, pt-005 and pt-006 where conventional KLD performed worse in those cases.

## 4. CONCLUSION

In this paper, we have employed both similarity and dissimilarity measures in the dissimilarity function, $D_{12}$, for multi-modal non-rigid image registration. We have also derived the derivative of $D_{12}$ so that it can be used in the optimizers, which require the gradient of the energy function. This is essential as many optimizers work on the basis of gradient and those optimizers are usually faster than others. $D_{12}$ has been tested on 7 CT-T1 image pairs with FFD and the L-BFGS-B optimizer, and compared with MI and conventional KLD based non-rigid registration methods. Experimental results show that the proposed method exhibits surpassing registration quality as compared to other two dissimilarity function based registration methods, among all 7 image pairs. The future research directions include applying $D_{12}$ to other non-rigid registration models like Markov Random Field (MRF) model, and testing $D_{12}$ with different types of images and modalities.

## 5. REFERENCES


[1] D. Rueckert, "Non-rigid registration: Techniques and applications," in *Medical Image Registration, CRC Press*, 2001.

[2] A. Chung *et al.*, "Multi-modal image registration by minimising Kullback-Leibler distance," *MICCAI*, 525–532, 2002.

[3] C. Guetter *et al.*, "Learning based non-rigid multi-modal image registration based on Kullback-Leibler distance," *MICCAI*, 255–262, 2005.

[4] A. Gholipour *et al.*, "Kullback-Leibler distance optimization for non-rigid registration of echo-planar to structural mr brain images," in *ICIP*, 2007.

[5] A. Chung *et al.*, "Robust multi-modal image registration using Kullback-Leibler distance based measures of intensity distribution similarity and dissimilarity." in *HKUST CS Technical Report*, HKUST-CS10-01, 2010.

[6] D. Rueckert *et al.*, "Non-rigid registration using free-form deformations: Application to breast mr images," *TMI*, V18 (8), 712–721, 1999.

[7] C. Zhu, and J. Nocedal, "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization," *ACM Trans. on Math. Software*, V23 (4), 550–560, 1997.

[8] D. Mattes *et al.*, "PET-CT image registration in the chest using free-form deformation," *TMI*, V22 (1), 120–128, 2003.

[9] P. Thevenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," *TIP*, V9 (12), 2083–2099, 2000.

[10] J. West *et al.*, "Comparison and evaluation of retrospective intermodality image registration techniques," *SPIE*, V2710, 332–347, 1996.

[11] L. Ibanez *et al.*, *The ITK Software Guide*, 2003.

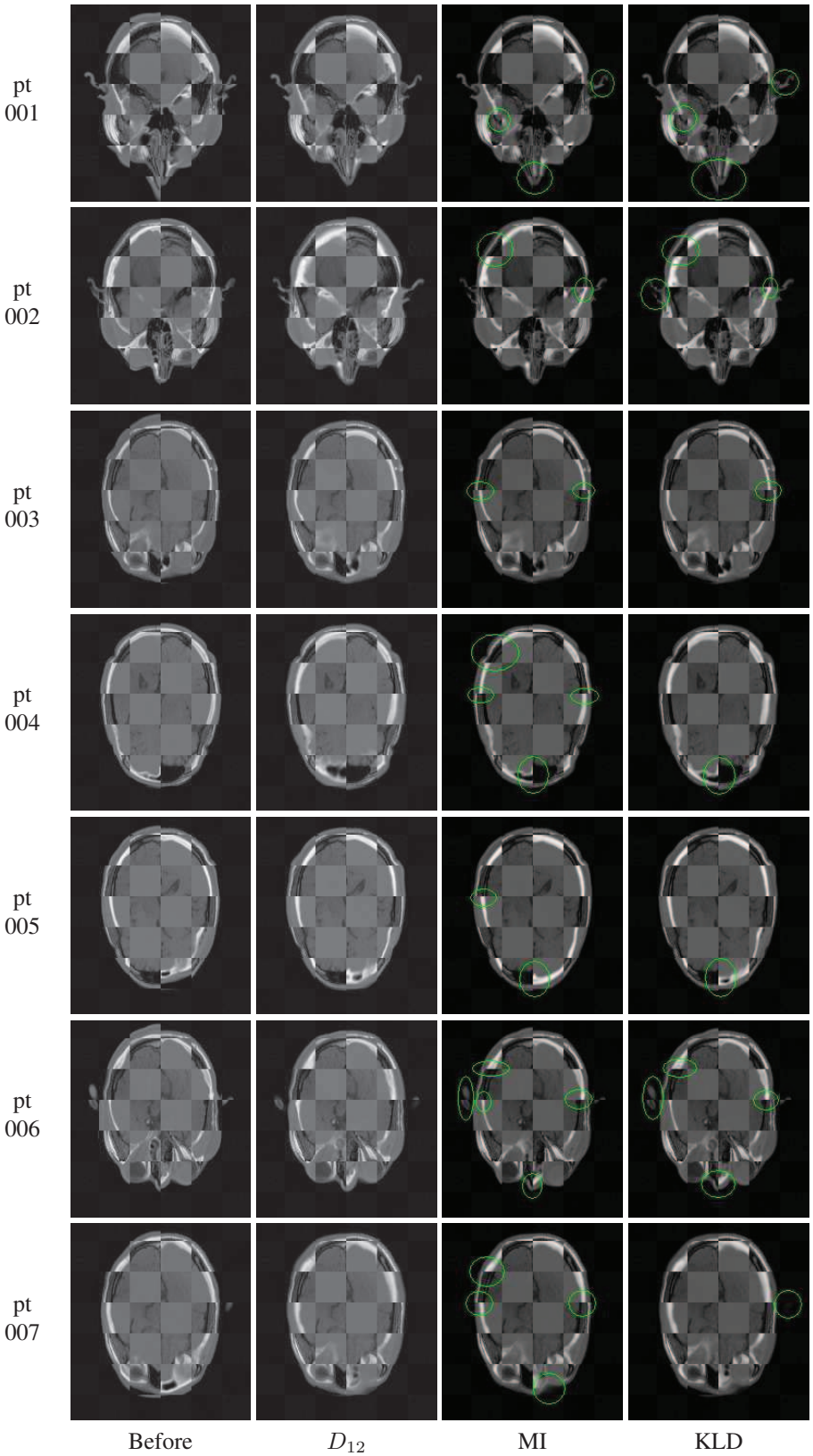


**Fig. 2**. Results of pt-001 to pt-007. (For viewing details, zoom-in the electronic file.)