ARTICLE IN PRESS

Neurocomputing ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Nonrigid medical image registration with locally linear reconstruction

Lingfeng Wang *, Chunhong Pan

NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 24 September 2013
Received in revised form
14 April 2014
Accepted 11 May 2014
Communicated by Jungong Han

Keywords:
Nonrigid image registration
Locally linear reconstruction
Intensity distortion
Multi-modal

ABSTRACT

Nonrigid image registration plays an important role in medical imaging systems. In complex circumstances, it is still a challenging task, especially when input images are corrupted by the spatially-varying intensity distortion. To address this difficulty, we propose a novel locally linear reconstruction based dissimilarity measurement (LLR). The core idea behind the LLR is that in each local region, the aligned floating image is linearly reconstructed by the reference image. Due to LLR, we present a multi-resolution nonrigid image registration algorithm. Extensive experiments on various simulated and real medical images have demonstrated that our algorithm outperforms the state-of-the-art approaches in registration accuracy.

1. Introduction

Nonrigid image registration is a fundamental component in various medical imaging systems. Much approaches have been proposed to tackle this problem [1–9]. Many excellent work has been done, however, it is still a challenging task, because real-world images are often disturbed by two sources (see Fig. 1 for details). The first one is the spatially-varying intensity distortion [10–12] which is caused by the multiplicative bias field and the additive correction field. The second one is related to the problem of multi-modality [13–15], in which two input images are acquired from different devices. The approaches to these problems can be classified into two groups, i.e., pre-homogenization and source-invariant.

The pre-homogenization approaches solve the two problems by adding a source removal process before registration. In [16–19], the slowly-varying intensity distortion is eliminated by performing a bias field correction process. In [20], the image multi-modality is removed by using a back-projection-like process. In [7], Wachinger and Navab proposed the entropy image (EI) and the Laplacian image so that the reference and floating images can be modality invariant as much as possible. In [21], Lee et al. presented a supervised technique to learn the similarity measurement. In [22], Heinrich et al. proposed a new descriptor called MIND (Modality Independent Neighborhood Descriptor). The more information about multi-modal image registration will be presented in Section 6.2. After pre-homogenizing the input images, the image registration is performed based on the well-understood dissimilarity measurements, e.g., the sum of squared difference. Their main limitation is that the registration result largely depends on the performance of pre-homogenization.

The source-invariant approaches desire to find a similarity or dissimilarity measurement that is invariable under above two sources. The normalized correlation coefficient (NCC) and the mutual information (MI) [14,23–26] are often utilized to address the multi-modal problem. The MI based method can be regarded as a family of the information theoretic based method. The core behind MI is the way of entropy estimation. For example, in [14], the entropy is estimated via a nonparametric way. If the input images are corrupted by the bias and correction fields, these global similarity measurements, including NCC and MI, are often not adequate. To solve this problem, in [27], Klein et al. proposed the local CC (LCC) and the local MI (LMI), while in [28], Yi and Soatto incorporated the local normalized MI into the global normalized MI. The other information theoretic based methods are built upon Renyi entropy, such as [29] and [30]. The details about multi-modal image registration will be presented in Section 6.2. In [12], Myronenko and Song proposed the implicit residual complexity measurement (RC), which is derived from the analytically solution of intensity correction field. Comparative experiments reported in [12] have illustrated that the RC based registration algorithm can produce accurate results on the intensity distorted images. However, as denoted in [12,31], the RC is hard to be applied to the multi-modal images.

Motivated by the source-invariant approaches, we propose a novel intensity-based dissimilarity measurement based on the locally linear reconstruction. To solve the spatially-varying intensity distortion problem, we present a new linear synthetic model, where the aligned floating image is synthesized by the product of the

* Corresponding author. Tel.: +86 10 62612582.
E-mail addresses: lfwang@nlpr.ia.ac.cn (L. Wang), chpan@nlpr.ia.ac.cn (C. Pan).

http://dx.doi.org/10.1016/j.neucom.2014.05.030
0925-2312/

Please cite this article as: L. Wang, C. Pan, Nonrigid medical image registration with locally linear reconstruction, Neurocomputing (2014), http://dx.doi.org/10.1016/j.neucom.2014.05.030

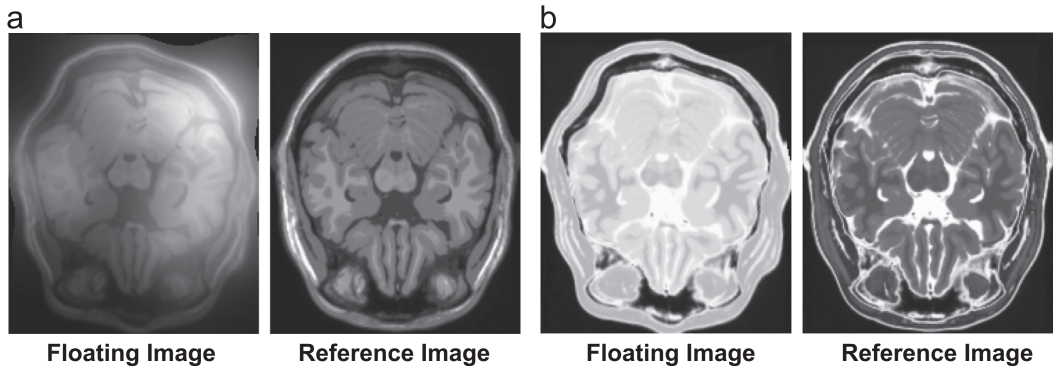

Fig. 1. Two challenging registration examples. (a) A pair of images with spatially-varying intensity distortions, (b) a pair of multi-modal images.

reference image by a multiplicative field, and with an additive field. The multiplicative and additive fields are assumed to be spatially-smooth. Based on this assumption, in each local region, we construct a linear regression model, which maps the reference image intensities to the aligned floating image intensities. After estimating the regression coefficients, the reconstruction error is obtained by substituting them into the regression model. All these local reconstruction errors are collected to form the LLR.

Specifically, the advantages or details of the proposed LLR are highlighted on as follows:

1. The proposed LLR is derived from the linear synthetic model, which incorporates both multiplicative and additive distortions. Hence, it can be applied to the images corrupted by the spatial-varying intensity distortion. The basic idea behind the LLR is that in each local region the aligned floating image can be approximately linear reconstructed from the reference image. Experimental results also indicate that it provides better adaptability to the registration of the multi-modal images.
2. The main difference between the LLR and the LCC is that the LLR further considers the variance of the aligned floating image. Thereby, the LLR based registration algorithm focuses more on salient edges rather than on smooth regions, which make our algorithm to preserve edge features better.

The rest of this paper is organized as follows. Section 2 describes the linear synthetic model. In Section 3, we present the LLR. In Section 4, we first formulate registration problem, and then describe our registration algorithm. Some experimental results are provided in Section 5. The discussions and conclusions are given in Section 6.

2. Linear synthetic model

Given the floating image $\mathbf{F}$ and the reference image $\mathbf{R}$, the image registration problem is phrased as to find a reasonable transformation $\mathbf{U}$, such that the aligned floating image should be similar to the reference image, given by

$$\mathbf{F}[\mathbf{U}] = \mathbf{R} + \mathbf{N},$$

where $\mathbf{F}[\mathbf{U}]$ is the aligned floating image and $\mathbf{N}$ is noise term, which is assumed to be Gaussian noise in this work.

In practice, the input floating image could be corrupted by the intensity distortion. Or more precisely, it may be interfered by the bias and correction fields. To introduce these two fields, the intensity relationship between $\mathbf{F}[\mathbf{U}]$ and $\mathbf{R}$ is assumed to follow the linear synthetic model:

$$\mathbf{F}[\mathbf{U}] = \mathbf{B} \odot \mathbf{R} + \mathbf{C} + \mathbf{N}, \tag{1}$$

where $\mathbf{B}$ and $\mathbf{C}$ are the multiplicative bias field and the additive correction field, and $\mathbf{N}$ is the noise term assumed to be Gaussian type. $\odot$ is the element by element multiplying operation. Both $\mathbf{B}$ and $\mathbf{C}$ are spatially-smooth.

In real-world image registration, the spatially-smooth fields $\mathbf{B}$ and $\mathbf{C}$ are unknown. To estimate these two fields accurately, it requires a good alignment from the floating image to the reference image. Hence, the pre-homogenization approaches, which estimate two fields beforehand, could be easy to fall into the dilemma of the preelection of image registration or two fields removing. On the contrary, the proposed LLR needs not to estimate two spatially-smooth fields in advance, and implicitly represents them into the locally linear model.

3. Locally linear reconstruction

The multiplicative bias field $\mathbf{B}$ and the additive correction field $\mathbf{C}$ are assumed to be slowly varying in the entire image domain Ω. The smoothness of field $\mathbf{B}$ is implicitly guaranteed by the following property[1]:

Given a relative small local region $\Omega_{\mathbf{x}}$ centered at $\mathbf{x}$ (3×3 local window for example), all the values in this local region are same with each other.

The field $\mathbf{C}$ holds the same property. Based on this property, in the local region $\Omega_{\mathbf{x}}$, the values of field $\mathbf{B}$ can be regarded as a constant $\mathbf{B}(\mathbf{x})$, while the values of field $\mathbf{C}$ can be considered as a constant $\mathbf{C}(\mathbf{x})$.

Now we consider the 3×3 local region $\Omega_{\mathbf{x}}$ centered at $\mathbf{x}$.[2] The coordinates within this local region are listed as follows:

$$\Omega_{\mathbf{x}} = \{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_8\},$$

in which $\mathbf{x}_0 = \mathbf{x}$ indicates the coordinate of center pixel, and $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_8$ are the coordinates of the pixels arranged from top-left to bottom-right. Based on the linear synthetic model, each pixel in the aligned float image $\mathbf{F}[\mathbf{U}](\mathbf{x}_i)$ is linearly synthetic by the corresponding pixel in the reference image $\mathbf{R}(\mathbf{x}_i)$ with $\mathbf{B}(\mathbf{x})$ and $\mathbf{C}(\mathbf{x})$ as linear coefficients, given by

$$\mathbf{F}[\mathbf{U}](\mathbf{x}_i) = \mathbf{B}(\mathbf{x})\mathbf{R}(\mathbf{x}_i) + \mathbf{C}(\mathbf{x}) + \mathbf{N}(\mathbf{x}_i), \quad i \in 0, 1, 2, \ldots, 8. \tag{2}$$

[1] It worth noting that our model considers the spatial smoothness of two fields $\mathbf{B}$ and $\mathbf{C}$ implicitly.

[2] The size of local region is determined experimentally by considering following two aspects: (1) the computational cost and memory usage; (2) the registration result. On one hand, when the size of local region is larger, the computational cost and the memory usage will be higher correspondingly. On the other hand, we experimentally find that when the size of local region is equal to or larger than 3×3, the registration results are very similar. Hence, we set the size of local region to 3×3.

Please cite this article as: L. Wang, C. Pan, Nonrigid medical image registration with locally linear reconstruction, Neurocomputing (2014), http://dx.doi.org/10.1016/j.neucom.2014.05.030