Information Sciences 372 (2016) 16–31

Contents lists available at ScienceDirect

Information Sciences

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

Self-similarity inspired local descriptor for non-rigid multi-modal image registration

Fei Zhu[a], Mingyue Ding[a,b], Xuming Zhang[a,b,*]

[a] Department of Biomedical Engineering, School of Life Science and Technology, Image Processing and Intelligent Control Key Laboratory of Education Ministry of China, Huazhong University of Science and Technology, Wuhan 430074, PR China
[b] Ministry of Education Key Laboratory of Molecular Biophysics, Huazhong University of Science and Technology, Wuhan 430074, PR China

ARTICLE INFO

Article history:
Received 27 November 2015
Revised 5 July 2016
Accepted 11 August 2016
Available online 12 August 2016

Keywords:
Non-rigid multi-modal registration
Self-similarity
Zernike moments
Similarity metric
Target registration error

ABSTRACT

Non-rigid multi-modal image registration plays an important role in medical image analysis. It remains a challenging problem due to the significant intensity distortion and the non-rigid transformation between images. Existing registration methods based on the information theoretic measures and image representations cannot address this problem effectively. In this paper, we have proposed a novel self-similarity inspired local descriptor to determine the similarity metrics, a key component in image registration. The self-similarity is determined based on the Zernike moments of image patches in a local neighborhood, and it is utilized to generate the local descriptor. The Euclidean distance between the local descriptors computed in the reference and moving images is used as the similarity metrics. Distinctively, the proposed local descriptor can provide an effective representation of complicated image features due to its robustness to noise and rotational invariance, which provides the proposed method with good registration performance. Extensive experiments on both simulated and real multi-modal image datasets demonstrate that the proposed method has higher registration accuracy appreciated by the target registration error (TRE) than the state-of-the-art registration methods based on the normalized mutual information (NMI), the sum of squared differences on entropy images (ESSD), the Weber local descriptor (WLD) and the modality independent neighborhood descriptor (MIND).

1. Introduction

Medical imaging technologies can assist doctors in the diagnosis and treatment of diseases. Despite the great advancement of such medical imaging technologies as ultrasound (US), computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET), they generally provide complementary information about the human body due to the different imaging principles. For example, US, CT and MRI could provide the anatomical information of the organs but fail to demonstrate their functional information. PET is good at demonstrating the metabolism information but it cannot provide morphological structures of organs clearly. Doctors often need to fuse a variety of information from different modalities for the diagnosis of diseases. In this scenario, the multi-modal medical image registration is essential for the effective image fusion.

* Corresponding author.
E-mail addresses: zhufei@hust.edu.cn (F. Zhu), myding@hust.edu.cn (M. Ding), zxmboshi@hust.edu.cn (X. Zhang).

http://dx.doi.org/10.1016/j.ins.2016.08.031
0020-0255/

The objective of image registration is to find correspondence between points that are present in the reference and moving images. In general, the image registration framework involves three main components including the deformation model, the similarity metric and the optimization method [38]. Among these components, the similarity metrics have attracted much attention recently. In the mono-modal image registration, there are many widely-used similarity metrics such as the sum of squared or absolute differences (SSD and SAD, respectively), cross correlation and correlation coefficient [3,8,24]. However, these metrics cannot be directly applied to the multi-modal image registration because of intensity distortion and non-rigid transformations between images caused by respiration and body motion in patients and some other factors. To overcome the disadvantageous influence of these factors, two approaches have been proposed to address the non-rigid multi-modal image registration problem.

The first approach is to use the information theoretic measures as the similarity metrics. Mutual information (MI) [11,29,47,48] has been widely applied to multi-modal image registration because of its generality without assuming the relationship between image intensities [33]. However, the spatial information has been ignored in the time-consuming computation of MI. Meanwhile, it is likely to produce misregistration in that MI is not overlap invariant, and it is easy to get trapped in the local optimum because MI is generally not a smooth function but one containing many local maxima. To overcome these disadvantages, the normalized MI (NMI) [39], the regional MI [40], the conditional MI [28] and the localized MI [25] have been proposed. But these improved MI metrics do not directly take into account the local image structures, which may result in distortion of local structures due to considerably more degrees of freedom in the non-rigid registration than those in rigid registration [35].

The second approach is the reduction of the multi-modal registration problem to the mono-modal one. Along this line, some researchers have tried to transform one modality into another. Arbel et al. [2] segmented the MR image and assigned different intensity transformations to different MR regions to generate pseudo-US images. These images were then registered to the real US ones using the automatic nonlinear image matching and anatomical labeling (ANIMAL) based registration technique [12]. Wein et al. [46] simulated the pseudo-US images from CT by taking into account the physical principles of ultrasound and then registered the pseudo-US images to real ones by using the correlation ratio as the similarity metric. Other researchers have tried to firstly map the multi-modal images to a common space using the image representation methods based on the assumption that there exists the same anatomical structure in multi-modal images, and then use such simple similarity metrics as SSD to perform the registration. Wachinger and Navab [44] proposed the entropy image and Laplacian eigenmaps based structural representation methods for image registration. In the first method, the entropy of the image patches was calculated and the entropy images based SSD (ESSD) was used as the similarity metric. The disadvantage of this method is that the entropy images seem to be blurry, thereby leading to the inaccurate similarity metrics. In the second method, all the image patches were utilized to build a neighborhood graph to approximate the manifold embedded in high dimensional patch space. The low-dimensional embedding was then calculated with the graph Laplacian. Embeddings from different modalities were aligned to obtain the final representation. The Laplacian images seem to be sensitive to noise in the image and their creation involves high computational complexity. To overcome the disadvantages of the ESSD method, Yang et al. [50] introduced a new similarity metric based on the Weber local descriptor (WLD). The WLD tends to generate artefacts and it is vulnerable to noise because of the involved differential operation implemented on the individual pixels. The modality independent neighborhood descriptor (MIND) [19] was proposed by Heinrich et al. for multi-modal image registration. The MIND was computed based on the similarities between neighboring patches, and thus it is robust to non-functional intensity relations and image noise. However, the MIND is not rotationally invariant and it only utilizes the image intensities for similarity computation. Therefore, it will generate inaccurate registration results for corners, edges and complicated textured regions where rotations may exist between features. Piella [34] used the diffusion maps to obtain a unified representation that captures the geometric and spectral properties of the data for the multi-modal registration. This method can represent the complicated image features, but it is computationally intensive and sensitive to image noise.

To address the problems of the above-mentioned image representation based registration methods, we have proposed a new similarity metric based on the self-similarity inspired local descriptor. The notion of self-similarity has been explored by the nonlocal means for image denoising [51], segmentation [18], super-resolution [15] and so on. In this paper, self-similarity is calculated based on the Zernike moments [10,14,23,32] of image patches and utilized to construct an effective local descriptor for the determination of similarity metrics. This descriptor will be used to extract such hybrid image features as image intensities as well as edge and texture features rather than only intensity information. Because of the robustness of self-similarity computation to image noise and the rotational invariance of the Zernike moments, the proposed Zernike moments based local descriptor (ZMLD) is invulnerable to image noise and it is rotationally invariant, thereby leading to its effectiveness in extracting the features of corners, edges and textures in the case of image noise. The L2 distance between the local descriptors of two images is computed and used as the similarity metric. By using the free-form deformation (FFD) [36,37] model as the transformation model, the similarity metric is optimized using the Limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) [31] method to obtain the registration results. Extensive experiments on simulated and real multi-modal medical images demonstrate that the proposed ZMLD based method outperforms the registration methods based on the NMI, ESSD, WLD and MIND in terms of registration accuracy.

The paper is organized as follows. Section 2 presents the computation of similarity metric based on ZMLD and the implementation of the proposed non-rigid multi-modal registration method. The experimental results and discussions are provided in Section 3. The conclusion is given in Section 4.