Medical Image Analysis 18 (2014) 343–358

Contents lists available at ScienceDirect

Medical Image Analysis

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

Self-similarity weighted mutual information: A new nonrigid image registration metric

Hassan Rivaz *, Zahra Karimaghaloo, D. Louis Collins

McConnell Brain Imaging Center (BIC), Montreal Neurological Institute (MNI), McGill University, Montreal, QC, Canada

A R T I C L E   I N F O

Article history:
Received 31 January 2013
Received in revised form 7 October 2013
Accepted 7 December 2013
Available online 21 December 2013

Keywords:
Nonrigid registration
α-Mutual information
Nonlocal means
Ultrasound and MRI registration
Mutual information

A B S T R A C T

Mutual information (MI) has been widely used as a similarity measure for rigid registration of multi-modal and uni-modal medical images. However, robust application of MI to deformable registration is challenging mainly because rich structural information, which are critical cues for successful deformable registration, are not incorporated into MI. We propose a self-similarity weighted graph-based implementation of α-mutual information (α-MI) for nonrigid image registration. We use a self-similarity measure that uses local structural information and is invariant to rotation and to local affine intensity distortions, and therefore the new Self Similarity α-MI (SeSaMI) metric inherits these properties and is robust against signal nonstationarity and intensity distortions. We have used SeSaMI as the similarity measure in a regularized cost function with B-spline deformation field to achieve nonrigid registration. Since the gradient of SeSaMI can be derived analytically, the cost function can be efficiently optimized using stochastic gradient descent methods. We show that SeSaMI produces a robust and smooth cost function and outperforms the state of the art statistical based similarity metrics in simulation and using data from image-guided neurosurgery.

1. Introduction

Image registration involves finding the transformation that aligns one image to the second, and has numerous medical applications in diagnosis and in image guided surgery/therapy. The joint intensity histogram of two images, be they from different or the same modalities, is spread (i.e. the joint entropy is high) when they are not aligned, and is compact (i.e. the joint entropy is low) when the two images are aligned. Therefore, mutual information (MI) (Wells et al. (1996); Maes et al. (1997); Pluim et al. (2003)) and the overlap invariant normalized MI (NMI) (Studholme et al. (1999)) have been proposed and widely used for rigid registration of multi-modal images.

MI is not robust against spatially varying bias fields since they result in different intensity relations between the two images at different locations. Therefore, Studholme et al. (2006) and Loeckx et al. (2010) proposed respectively regional MI (RMI) and conditional MI (CMI) where spatial information is used as an extra channel for conditioning MI. This essentially leads to summing MI calculated for regions of the images, instead of globally estimating MI. Klein et al. (2008) proposed localized MI (LMI) where samples are randomly selected from regions in every iteration and convergence is achieved by using stochastic optimization Klein et al. (2007, 2009). Zhuang et al. (2011) proposed spatially encoded MI, which instead of giving equal weights to all pixels in a region, hierarchically weights pixel contributions based on their spatial location. These methods have shown to significantly improve the registration results in the presence of bias fields. Recently, Darkner and Sporring (2013) provided a unifying framework for NMI and other common similarity measures and shed more intuition towards local histograms.

A second difficulty rises because MI does not directly take into account local structures. Therefore, nonrigid registration, which has considerably more degrees of freedom, can distort local structures. Utilizing image gradients and their orientations was proposed by Pluim et al. (2000). Recently, De Nigris et al. (2012) proposed a gradient orientation metric that adaptively controls the trade-off between smooth or accurate cost functions. The HAMMER framework of Shen and Davatzikos (2002) sets local geometric moment invariants as attribute vectors of each voxel in the image. These attribute vectors are then used to form a cost function, which is hierarchically optimized to give the transformation parameters. Xue et al. (2004) later used wavelet-based attributes as local morphological signatures for each voxel. Recently, Ou et al. (2011) introduced Gabor attributes which can be used for different imaging modalities and tissue organs, and further utilizes mutual saliency to weight different voxels based on their local appearance. Taking a different approach, Wachinger and Navab (2012) generated entropy images independently from each image

* Corresponding author. Tel.: +1 5143981573.
E-mail address: hassan.rivaz@mcgill.ca (H. Rivaz).

1361-8415/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.media.2013.12.003

by calculating entropy in small patches around every pixel. They show that since different imaging modalities show the same tissue structure, their entropy images are similar and therefore they can be registered using monomodal registration. In addition to the entropy image representation, they show that structural information of patches can be encoded into a scalar value using manifold learning techniques. Performing the same technique on both images, they again arrive at two representations (one for each image) which can be registered using monomodal techniques.

A third problem with MI lies in the fact that the *infinite dimensional* joint and marginal probability distributions[1] are required to calculate the *scalar* parameter MI. Most MI estimation methods (Wells et al. (1996); Maes et al. (1997); Pluim et al. (2003)) substitute nonparametric density estimators, such as Parzen windows, into the MI formulation, and are called "plug-in" estimation in Beirlant et al. (1997). An inherit problem of these methods is due to the infinite dimension of the unconstrained densities. Strict smoothness constraints or lower dimensional parametrization must be enforced to estimate these densities, which can cause significant bias in the estimate (Hero et al., 2002). Graph-based entropy estimators (Hero et al., 2002; Neemuchwala and Hero, 2005) have been proposed to directly calculate entropy without the need for performing density estimation. Therefore, these methods have faster asymptotic convergence rate especially for nonsmooth densities and high dimensional feature spaces (Hero et al., 2002). Two drawbacks of these methods are their computational complexity and the discontinuity of their gradient as the graph topology changes.

Towards developing a bias invariant similarity metric for nonrigid registration that also takes into account structural information, we build on our previous work (Rivaz and Collins, 2012) to incorporate image self-similarity into MI formulation. Self-similarity estimates the similarity of a patch in one of the images to other patches in the *same* image, and attributes the similarity to the pixels in the center of the patches. Based on patches, self-similarity depends on local structures which are ignored by MI. Buades et al. (2005) first proposed exploiting repetitive regions (or patches) in the form of nonlocal means for image denoising. A recent comparative study of these methods is provided in Buades et al. (2010). Self-similarity was later used for object detection and image retrieval (Shechtman and Irani (2007)), and it has since been used successfully in denoising MR (Coupe et al. (2008); Manjon et al. (2012)) and US images (Coupe et al. (2009)), and image segmentation (Coupe et al. (2011)). Compared to our previous work (Rivaz and Collins (2012)), we present significantly more details and in-depth analysis of SeSaMI. We also provide extensive results for validation and more analysis of the results.

Recently, Heinrich et al. (2011, 2012) proposed using self-similarity for multimodal image registration. The similarity of a pixel to its neighbors, calculated using sum of square differences (SSD), are attributed to the pixel as multi-dimensional descriptors. These descriptors are calculated independently for both images. The multi-modal image similarity is then defined as the SSD of the descriptors of the two images.

Since self-similarity is calculated for pairs of points, it is natural to perceive it in a graph representation where image pixels are vertices and self-similarity is the weight of the edges. Graph-based estimators of α-mutual information (α-MI) similarity metric have recently been proposed for both rigid (Neemuchwala and Hero (2005); Sabuncu and Ramadge (2008); Kybic (2007); Kybic and Vnucko (2012)) and nonrigid (Staring et al. (2009); Oubel et al. (2012)) registration applications. These methods have been shown to outperform the traditional "plug-in" entropy estimators for MI calculation. Therefore, we choose to incorporate self-similarity into this registration framework.

We apply the method to register pre-operative magnetic resonance (MR) images to intra-operative ultrasound (US) images in the context of image-guided neurosurgery (IGNS). Previous work that registers US to other modalities is relatively rare: Roche et al. (2001) used the correlation ratio (CR) between US and MR and MR gradient, Arbel et al. (2004) and Mercier et al. (2012b) calculated a lookup table for mapping US and MR intensities and used the monomodal registration of Collins et al. (1999); Kuklisova-Murgasova et al. (2012) segmented the MR volume using a probabilistic atlas, generated a US-like volume from the segmented MR volume, and then registered the US-like volume with the US volume using robust monomodal block-matching techniques, Penney et al. (2004) generated blood vessels probability maps from from US and MR and registered these maps using cross-correlation, Ji et al. (2008) used NMI of US and MR, Zhang et al. (2011) used MI of phase information to register US to MR, De Nigris et al. (2012) optimized MI of gradient orientations to register US to MR, Wein et al. (2013) assumed a linear relationship between US intensities and MR intensities and gradient magnitudes, and finally Heinrich et al. (2013) used the self-similarity context along with a discrete optimization approach through block-wise parametric transformation model with belief propagation.

Most of the aforementioned methods simulate US images from the MR data as described. These methods cannot be readily applied to IGNS due to the variety of pathologies that the brain tissue might have, such as different grade gliomas. The appearance of these pathologies in MR and US are also highly variable (Mercier et al., 2012b; Mercier et al., 2012a), adding to the difficulty. We assume no *a priori* relationship between intensities but opt for two nonparametric MI based methods for validating our results.

Fig. 1 shows an example of the registered US and MR images. The US images suffer from strong bias field due to signal attenuation, caused by scattering (from smaller than US wavelength inhomogeneities), specular reflection (from tissue boundaries) and absorption (as heat). In addition, US beam width varies significantly with depth, and therefore the same tissue may look different at different depths. A final and important source of spatial inhomogeneities is the time gain compensation (TGC) which is manually adjusted on US machines. Hence, it is critical to exploit local structures.

Our algorithm only needs the self-similarity of one of the images. In most image guided applications, one of the images is pre-operative, and therefore the self-similarity estimation can be performed offline, resulting in a small increase in the on-line computational complexity. In addition, the pre-operative image is also usually of higher quality, making it a more attractive choice. We use a rotation invariant self-similarity metric that is also robust to bias fields, and utilize it in a graph-based α-MI method. We call our method the Self Similarity α-MI (SeSaMI) algorithm. We show that SeSaMI outperforms LMI and multi-feature α-MI in terms of producing a smooth dissimilarity function and registration accuracy.

This paper is organized as following. We first formulate the problem of image registration as an optimization problem, and provide background information for two popular similarity metrics that we use in this work for comparisons. We then elaborate on how we estimate self-similarity between patches. We explain a graph-based α-MI similarity metric, and then formulate SeSaMI by incorporating self-similarity into it. We also show how the derivative of SeSaMI can be efficiently estimated. We finally show the results on simulation and patient data for validation.

[1] The probability distributions are infinite dimensional if we assume image intensities take real continuous values. However, since intensities of digital images are discrete and finite, the probabilities distributions are finite, but still very high dimensional.