Pattern Recognition 48 (2015) 1937–1946

Contents lists available at ScienceDirect

Pattern Recognition

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

Feature Neighbourhood Mutual Information for multi-modal image registration: An application to eye fundus imaging

P.A. Legg [a,b,*], P.L. Rosin [a], D. Marshall [a], J.E. Morgan [c]

[a] School of Computer Science, Cardiff University, UK
[b] Department of Computer Science, University of Oxford, UK
[c] School of Vision Sciences and Optometry, Cardiff University, UK

ARTICLE INFO

Article history:
Received 10 May 2013
Received in revised form
2 December 2014
Accepted 17 December 2014
Available online 24 December 2014

Keywords:
Mutual information
Feature derivatives
Gauge co-ordinates
Non-rigid registration
Ophthalmology

ABSTRACT

Multi-modal image registration is becoming an increasingly powerful tool for medical diagnosis and treatment. The combination of different image modalities facilitates much greater understanding of the underlying condition, resulting in improved patient care. Mutual Information is a popular image similarity measure for performing multi-modal image registration. However, it is recognised that there are limitations with the technique that can compromise the accuracy of the registration, such as the lack of spatial information that is accounted for by the similarity measure. In this paper, we present a two-stage non-rigid registration process using a novel similarity measure, *Feature Neighbourhood Mutual Information*. The similarity measure efficiently incorporates both spatial and structural image properties that are not traditionally considered by MI. By incorporating such features, we find that this method is capable of achieving much greater registration accuracy when compared to existing methods, whilst also achieving efficient computational runtime. To demonstrate our method, we use a challenging medical image data set consisting of paired retinal fundus photographs and confocal scanning laser ophthalmoscope images. Accurate registration of these image pairs facilitates improved clinical diagnosis, and can be used for the early detection and prevention of glaucoma disease.

1. Introduction

Image registration is the task of finding the spatial transformation that gives correct matching correspondence between two images. Registration is widely used in many application areas, including medical imaging, computer vision, and satellite imagery. In particular, registration of images from different modalities has become increasingly common to combine signals from multiple sensors, where the registered images can be used to examine or explain a particular observation further, for example in patient diagnosis. However, the difficulty is that by their very nature, multi-modal image pairs may have no clearly defined relation between corresponding image intensities. Mutual Information (MI) has become a popular similarity measure for registering images of different modalities. The algorithm was simultaneously proposed by Viola and Wells [25] and Maes et al. [13]. MI differs from earlier registration methods as it is derived from information theory and is based on a statistical comparison of the images. Given two images, A and B, MI can be defined as

$$I(A;B) = H(A) + H(B) - H(A,B)$$

where $H(A)$ is the entropy of A, $H(B)$ is the entropy of B and $H(A,B)$ is the joint entropy of A and B. The transformation that maximises $I(A,B)$ *should* give the correct registration of the images. Entropy gives a measure of the amount of information that a given signal may contain, and forms the basis of MI. For a signal X consisting of n elements, Shannon's entropy [22] is defined as

$$H(X) = -\sum_{i=0}^{n} p(i)\log_2 p(i)$$

where $p(i)$ is the probability of value i occurring within the data set. The amount of information for a given value is inversely related to its probability, meaning that if the probability of a particular value occurring is low then this returns a greater amount of information than if the probability of the value is high. It can be thought of that the more rare the occurrence of an event, the more important it is when that event does occur. Despite the wide adoption of MI, it is recognised that the method is not without limitations [12], nor can it accurately register all varieties of image modalities, and so alternative methods have since been proposed [18,9,20,1,4,15,21,24,27,11].

* Corresponding author.
E-mail addresses: phil.legg@cs.ox.ac.uk (P.A. Legg), paul.rosin@cs.cf.ac.uk (P.L. Rosin), dave.marshall@cs.cf.ac.uk (D. Marshall), morganje3@cf.ac.uk (J.E. Morgan).

http://dx.doi.org/10.1016/j.patcog.2014.12.014
0031-3203/

Fig. 1. A patient's eye captured by two different image modalities, showing the retina surface and blood vessels. *Left*: confocal SLO image. *Right*: Fundus colour photograph.

For our study, we are particularly interested in the registration of a challenging data set comprised of multi-modal retinal image data, in order to improve clinician diagnosis and treatment of glaucoma. Glaucoma is the second most common cause of blindness in the West and the most common cause of irreversible blindness worldwide [23]. The effects of glaucoma are irreversible, meaning that it is crucial to detect it in the early stages in order to prevent any further progression of the condition [19]. As shown in Fig. 1, the image modalities that are to be registered together are colour fundus photographs (shown on the right) and confocal scanning laser ophthalmoscope (SLO) images (shown on the left). Both modalities capture high quality images from the eye of the optic nerve head (ONH), with the fundus photograph recording the clinical appearance and the SLO image providing quantitative information such as the retinal surface reflectivity and topographic structure [14]. From Fig. 1, it is apparent to see that there is corresponding structure present in both modalities, however the two modalities present this structure differently due to the acquisition techniques. For example, the surface reflectivity of vessels and the ONH result in a different representation compared to the colour photograph, such as the dark ring of nerve fibres at the ONH in the SLO, and the hollow appearance of large vessels in the SLO. Currently it is not typical practice to register these two modalities, however since the ONH boundary appears much clearer in the fundus photograph it seems a logic step to utilise both images effectively. Registration would provide correspondence between topographic and visible ONH damage, and early detection of glaucoma would result in better prognosis and treatment.

Fig. 1 shows an example colour fundus photograph and SLO reflectivity image captured from a patient's eye, taken from our image data set. The data set used in this study consists of 135 matching image pairs captured from the human eye. The original size of each fundus photograph is 752×490 pixels, with a resolution of 72 pixels per inch. The SLO images are captured using the Heidelberg Retinal Tomograph II (HRT II) [6] device. The field of view for each SLO image is 15×15 degrees and the original size is 384×384 pixels, with a resolution of 96 pixels per inch. The data set consists of both left and right eyes and shows various stages of the glaucoma disease ranging from no sign of infection to highly glaucomatous. The data set provides an interesting challenge for the image processing community, since there may be regions of non-uniform lighting, regions that lack textural appearance, physical changes in structure and colour over time due to degradation, and also distortion introduced by the various changes in curvature from the retina surface. In addition, whilst the SLO images are of a high clinical standard, there are some cases where slight blurring occurs in the image due to subtle movement in the eye (known as microsaccades) during acquisition. The data set comprises a wide variety of cases that a clinician would encounter when capturing these two image modalities. All images were taken by an expert clinician, who also provided ground truth data using a manual alignment tool that was developed specifically for this task. The anonymised data set is available from the authors upon request.

This paper provides a extension of our preliminary report given in [11]. Here, we now incorporate a two-stage non-rigid registration to provide greater registration accuracy, in addition to exploring a much wider set of multivariate features that can be utilised by the Feature Neighbourhood Mutual Information (FNMI) algorithm. In addition, we also provide a novel analysis of registration convergence in the transformation space to examine *why* different methods fail to provide accurate registration. This highlights a key consideration, in that a successful similarity measure not only needs to maximise the correct registration, but would also need to provide a results across the complete transformation space that converges towards the correct registration so that transformation optimisation can be performed without failure. The remainder of the paper is as follows: Section 2 provides a literature review of Mutual Information and different techniques that have extended upon the original algorithm. Section 3 describes our proposed algorithm, Feature Neighbourhood Mutual Information (FNMI). Section 4 presents the application of rigid registration. Section 5 discusses the issue of registration convergence and explores how this impacts on the different similarity measures. Section 6 extends our method for non-rigid registration and presents our final results, followed by our conclusion given in Section 7.

2. Literature review

As has already been introduced, Mutual Information (MI) is a popular similarity measure for registering images of different modalities. The algorithm was simultaneously proposed by Viola and Wells [25] and Maes et al. [13]. MI relies on the computation of entropy, which gives a measure of uncertainty for a random variable. It can be observed that by reducing the uncertainty within the joint distribution of the images, we obtain the strongest correspondence between them whilst the entropy of each individual image ensures that the image overlap contains meaningful information (rather than registering regions of little interest such as background).

One recognised issue with MI is that there is little spatial information incorporated into the measure, which means that if there is a complex correspondence between the image modalities then the standard approach can often fail [12]. Many methods have been proposed to overcome this by including additional information as part of registration. Pluim et al. [18] suggested integrating gradient information into the MI measure by multiplying MI by a gradient term. Similarly, Kubecka and Jan [9] suggested using gradient-image MI whereby MI is computed for both the original images (after performing illumination correction) and also for the corresponding gradient