Biomedical Signal Processing and Control 23 (2016) 52–61

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

A structure-based region detector for high-resolution retinal fundus image registration

Zeinab Ghassabi [a,*,1], Jamshid Shanbehzadeh [b,2], Ali Mohammadzadeh [c,3]

[a] Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
[b] Department of Computer Engineering, Kharazmi University, Tehran, Iran
[c] Department of Remote Sensing, K.N.Toosi University of Technology, Tehran, Iran

ARTICLE INFO

Article history:
Received 29 September 2014
Received in revised form 30 July 2015
Accepted 5 August 2015
Available online 2 September 2015

ABSTRACT

A fundamental problem of retinal fundus image registration is the determination of corresponding points. The scale-invariant feature transform (SIFT) is a well-known algorithm in this regard. However, SIFT suffers from the problems in the quantity and quality of the detected points when facing with high-resolution and low-contrast retinal fundus images. On the other hand, the attention of human visual systems directs to regions instead of points for feature matching. Being aware of these issues, this paper presents a new structure-based region detector, which identifies stable and distinctive regions, to find correspondences. Meanwhile, it describes a robust retinal fundus image registration framework. The region detector is based on a robust watershed segmentation that obtains closed-boundary regions within a clean vascular structure map. Since vascular structure maps are relatively stable in partially overlapping and temporal image pairs, the regions are unaffected by viewpoint, content and illumination variations of retinal images. The regions are approximated by convex polygons, so that robust boundary descriptors are achieved to match them. Finally, correspondences determine the parameters of geometric transformation between input images. Experimental results on four datasets including temporal and partially overlapping image pairs show that our approach is comparable or superior to SIFT-based methods in terms of efficiency, accuracy and speed. The proposed method successfully registered 92.30% of 130 temporal image pairs and 91.42% of 70 different field of view image pairs.

1. Introduction

Retinal image registration is the process of finding the geometric transformation between two or more retinal images from different times, viewpoints and sources [1,2]. This paper focuses on registration of images from different times (i.e., temporal registration) and viewpoints in the presence of high-resolution retinal fundus images. Temporal image registration facilitates estimating the diseases or therapeutic progress by better measuring of changes both on retinal vascular tree and on the color contents of the eye fundus [3]. Moreover, registration of different viewpoint images generates a wider view of the retina which is helpful for retinal tracking [4]. Accurate and real-time retinal fundus image registration is still challenging in the presence high-resolution, small-overlapping regions, time-varying intensities and contents and low quality of retinal images. High-resolution images have made it more convenient for ophthalmologists to detect very fine structures, such as hemorrhages and microaneurysms in diabetic eye screening [5,6]. However, existing retinal fundus image registration algorithms have high computational cost in the presence of high-resolution retinal fundus images. Since reducing image resolutions produces artifacts, the registration results of resized images will be inapplicable to change detection and eye-screening [5,6]. For example, rescaling of images may produce blocking artifacts, which are liable to mistake the detection of some lesions like microaneurysms, whereas a needed resolution to detect the smallest microaneurysms is 1360 × 1000 pixels [6]. Based on this reason, the need for an image registration approach that will resolve the problem is urgent.

Retinal fundus image registration methods include area-, feature- and hybrid-based methods [2]. Area-based methods optimize intensity-based similarity measures to find parameters of

* Corresponding author. Tel.: +98 9126937852.
E-mail addresses: z.ghassabi@srbiau.ac.ir (Z. Ghassabi), jamshid@khu.ac.ir (J. Shanbehzadeh), a_mohammadzadeh@kntu.ac.ir (A. Mohammadzadeh).
[1] Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran 4933-14155, Iran
[2] No. 49, Shahid Mofatteh Street, Department of Computer Engineering, Kharazmi University, Tehran, Iran, postal code: 37551-31979
[3] No. 1346, Vali-Asr Street, Mirdamad Crossing, Remote Sensing Dept, Geomatics Eng. Faculty, K.N.Toosi University of Tech., Tehran – Iran, Postal Code: 1996715433

http://dx.doi.org/10.1016/j.bspc.2015.08.005
1746-8094/

geometric transformation between image pairs [7–9]. The similarity measures include cross-correlation, sum of absolute values of differences [10], mutual information [7,9,11] and phase correlation [12]. These intensity-based similarity measures are sensitive to temporal image registration since large morphological changes (e.g., optic disc topographic changes in a glaucoma patient [7]) lead to significant changes between pixel intensity values in the input images. Therefore, many local maxima in the similarity measure function mislead the optimization methods [7]. Moreover, similarity measures may mislead by non-overlapping areas of viewpoint images [13]. Area-based methods suffer from computational burden since they employ the entire image content, whereas feature-based algorithms detect suitable locations in the images to establish feature correspondences. Then, correspondences determine the transformation parameters. Good locations are stable under content and geometric variations between image pairs. Moreover, they should generate distinctive descriptors to find correct matches [14]. In retinal fundus images, vessel bifurcations are invariant to intensity variations. However, bifurcation extraction requires reliable centerline detection of vessels. Besides, angle-based invariant descriptors, which used around bifurcations to find correspondences, may be similar and points will not be distinguishable from each other [15,16]. Hybrid methods employ the advantages of both area-based and feature-based methods. In this view, Chanwimaluang et al. [11] extracted bifurcations, and established correspondences using mutual information in regions around the bifurcations.

Nowadays, scale-invariant feature transform (SIFT) [17] is the most stable and distinctive local feature which is independent of vascular extraction. It consists of the scale-space extreme point detection, filtering spurious detections via a single global threshold and invariant descriptor generation. Existing feature-based retinal fundus image registration methods, General Dual bootstrap iterative closest point (GDB-ICP) [18] and SIFT-GTM [19] (Graph Transformation Matching), find initial matches by SIFT. The performance of those methods degrades in the presence of contrast variability within high-resolution retinal fundus images since SIFT detector suffers from the quantity, quality, and distribution of the detected points [20]. A global threshold of intensity-based SIFT detector is unable to overcome intensity variations within the retinal images and leads to detecting inadequate points in the presence of high-resolution and low-contrast retinal images [20]. Although a small global threshold produces lots of points, it hampers feature matching performance since the feature points are mainly unstable and redundant. In addition, processing of many feature points to find correspondences leads to high computational complexity and memory requirements.

The human visual system seeks similar regions in the images instead of points to establish correspondences [21,22]. In the recent years, efforts have been made to imitate such behavior, because it allows optimization of the computational resources as they can be focused on the processing of a set of selected regions only. Recently, Zheng et al. [23] employed a region saliency measure based on local entropy and variance values of image pixels to extract informative and salient feature regions on multimodal retinal images. Regions were matched by gradient based descriptors. Soon after, Gharabaghi et al. [24] extracted closed-boundray regions on segmented retinal fundus images. Then, regions were matched by the moment invariant descriptors. Both existing retinal region detectors [23,24] and SIFT found stable areas or points by analysis of local intensity cues. They focuses on the appearance of region interiors and ignored structural cues of region boundaries which are more robust to intensity variations [25,26]. Conversely, this paper proposes a structure-based region detector. It is based on a watershed segmentation of a clean vascular structure map to detect polygonal enclosed regions by vessels reliably. Descriptions of polygonal boundaries with distinct shapes are more effective than the appearance of region interiors by SIFT since retinal images contain many vessels similar to each other. In comparison to SIFT-based approaches, our retinal fundus image registration method is fast since it only focuses on the processing of regions which the number of them are fewer than thousands of SIFT feature points, but they are selected as non-redundant and highly stable regions.

The organization of this paper is as follows: Section 2 explains our methodology, including region detection, region description, region matching and mismatch elimination, and transformation parameter estimation. Section 3 presents experimental results. Conclusion and suggestions for further study are given in Section 4.

2. Proposed registration algorithm

The proposed registration framework includes five stages as follows:

- Region detection
- Region description
- Region matching and mismatch elimination
- Increasing the number of correspondences
- Transformation parameter estimation

Fig. 1 shows the flowchart of the registration framework including the proposed region detector. The proposed registration method first extracts polygonal regions of retinal fundus images. Next, it provides correct matches by computing rotation and scale invariant boundary-based descriptors, cross-matching the descriptors, and eliminating the mismatches. Since our approach may lead to few matches in some challenging cases and they are inadequate to estimate the parameters of geometric transformation between images, it is essential to increase the number of matches. Finally, the matched points are employed to find the transformation parameters between the two input images. The following subsections provide a detailed description of the main components.

2.1. Region detection

We develop a fast process that locates the enclosed regions by the vessels using the watershed segmentation [27] of a vascular structure map. Watershed transformation leads to oversegmentation due to nonhomogeneous intensity variations in high-resolution unhealthy retinal backgrounds [25]. Hence, computing a clean and reliable vascular structure map is a starting point to detect robust watershed regions. The following subsections provide a detailed description of region detection stages.

2.1.1. Vessel structure map

Finding vessel structure map consists of three steps. First, the color saliency boosting algorithm [28,29] converts an RGB color space to an RGB color-boosted space which is analogous to the color channels in human visual processing. Then, gray-scale morphological operations enhance the gray-value version of the new RGB image. Finally, the Frangi vesselness measure (FVM) [30] intensifies the vesselness areas.

Consider an RGB image $f = (R, G, B)$ and its local color-derivatives $f_x = (R_x, G_x, B_x)$. Studies show that the luminance variations in color-derivative distribution of RGB images are more probable rather than chromatic changes [28]. Hence, RGB channels are highly correlated and retinal fundus images have dull appearances. Furthermore, the gray-value versions of the RGB retinal images contain low contrast and interpretability [31]. Here, we apply the color saliency boosting algorithm [28] to transform an