Hybrid Retinal Image Registration Using Mutual Information and Salient Features

Jaeyong Ju††, Murray Loew††, Bonhwa Ku††, Nonmembers, and Hanseok Ko††, Member

SUMMARY This paper presents a method for registering retinal images. Retinal image registration is crucial for the diagnoses and treatments of various eye conditions and diseases such as myopia and diabetic retinopathy. Retinal image registration is challenging because the images have non-uniform contrasts and intensity distributions, as well as having large homogeneous non-vascular regions. This paper provides a new retinal image registration method by effectively combining expectation maximization principal component analysis based mutual information (EMPCA-MI) with salient features. Experimental results show that our method is more efficient and robust than the conventional EMPCA-MI method.

key words: retinal image registration, salient features, mutual information, medical imaging

1. Introduction

Image registration is the process of aligning two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors [1]. It has been widely used in a number of fields such as medical imaging, computer/robot vision, surveillance and remote sensing [1], [2]. In the case of medical imaging, a registered retinal image from the same patient facilitates the diagnosis and treatment planning for various eye diseases, including age-related macular degeneration, myopia, glaucoma, and diabetic retinopathy [2]. Retinal image registration is challenging. The images exhibit non-uniform/non-consistent contrasts and intensity distributions, as well as having large homogeneous non-vascular/textureless regions. These two major factors may limit the performance of registration. Also, non-vascular regions of unhealthy retinas exhibit a variety of pathologies over time, which makes the registration difficult. In contrast, vascular regions are relatively stable over time [3]. In recent years a number of retinal image registration methods have been proposed. Typically, these methods can be classified into intensity-based, feature-based, or hybrid-based methods. Feature-based methods extract features from a retinal image first, such as vascular bifurcation points [3], whole vasculature [4], and optic disk [5]. Then, the registration process that finds the best transform parameters is performed by maximizing a similarity measure based on correspondences of the extracted features. In a recent paper, Zheng et al. [6] proposed registration of retinal images using salient features, like vessels based on a saliency metric. It is robust to background changes and pathologies in the retina. Generally, most feature-based methods are based on vascular features. These methods are computationally efficient and robust to intensity variation, but their performance largely depends on the quality of segmentation and whether or not there are sufficient and reliable correspondences. Intensity-based methods optimize a similarity measure such as intensity difference, cross correlation, phase correlation, or mutual information (MI) [7]–[9]. The MI, a measure of the statistical correlation between intensity values of the images, has been widely used in medical imaging. However, because of the challenging characteristics of retinal image, the MI may not perform effectively in such an image. Accuracies of these methods are higher than those of feature-based methods because they deal with the whole image information. However, the computational cost of these methods is large and the registration performances are closely related to the background changes and image quality. Hybrid-based methods combine both feature- and intensity-based methods to overcome the drawbacks of both methods. Methods that combine vascular structures with entropy correlation coefficient [10] and spatial information in regional MI [11] and feature neighborhood MI [12] have been proposed. These methods use covariance matrices to reduce the data complexity instead of high-dimensional histograms, though as the spatial information increases, so commensurately does the corresponding computational cost. Recently, an approach termed “EMPCA-MI” [13], which combines spatial information with MI using Expectation Maximization for Principal Component Analysis (EMPCA) [14], has been proposed for efficient dimensionality reduction. Although reducing the dimensionality efficiently, it still has a lot of computational load because it considers neighbors of all pixels. It also has been shown to be ineffective in dealing with large homogeneous non-vascular/textureless regions exhibiting a variety of pathologies in unhealthy retinas [5].

This paper proposes a new hybrid image registration algorithm which combines EMPCA-MI with salient features from the vascular and the surrounding regions to achieve computational efficiency and robustness.

The remainder of this paper is organized as follows: Sect. 2 details the proposed algorithm, while Sect. 3 describes the experimental results and a discussion. Finally, conclusions are drawn in Sect. 4.

Manuscript received December 28, 2015.
Manuscript publicized March 1, 2016.
†The authors are with the School of Electrical Engineering, Korea University, Seoul, Korea.
††The author is with the Department of Biomedical Engineering, George Washington University, Washington DC, 20525, USA.
a) E-mail: jyju@ispl.korea.ac.kr
DOI: 10.1587/transinf.2015EDL8265
2. Proposed Retinal Image Registration

2.1 Principles of Retinal Image Registration

For a reference image $I_R$ and a sensed image $I_S$, we aim to find transformation parameters that align a moving $I_S$ onto a fixed $I_R$ by maximizing the proposed similarity measure through an optimization procedure. We provide details of the algorithm in the following subsections.

2.2 EMPCA-MI Similarity Measure

EMPCA-MI [13] is a recently introduced similarity measure for image registration that combines neighborhood regions information with MI. These methods reduce the computational cost by using EMPCA [14] that iteratively computes only the first principal component as the key feature from neighborhood regions of an image. It is robust against the presence of non-uniform intensity and noise. As illustrated in Fig. 1, it comprises of three steps that involve image rearrangement followed by EMPCA and MI calculations. Note the block colors in Fig. 1 indicate the pixels for preprocessing and not the actual pixel itself. In Step 1, $I_R$ and $I_S$ are both pre-processed in vector form for a given neighborhood radius $r$, so the spatial and intensity information is preserved. The first $P$ principal components $X_R$ and $X_S$ of the respective reference and sensed images are then iteratively computed by the EMPCA in Step 2, instead of solving the covariance matrix. The final MI value is calculated between $X_R$ and $X_S$ in Step 3, with higher MI value meaning the two images are better aligned. In the EMPCA-MI algorithm, only the first principal component is considered, i.e., $P = 1$ since this is the direction of highest variance and represents the most dominant feature. EMPCA-MI can be formally expressed as:

$$F(I_R, I_S) = \sum_{X_R, X_S} p(X_R, X_S) \log \frac{p(X_R, X_S)}{p(X_R)p(X_S)}$$  \hspace{1cm} (1)

where $p(X_R)$ and $p(X_S)$ are the individual probabilities of $X_R$ and $X_S$ respectively, while $p(X_R, X_S)$ is their joint probability.

2.3 Salient Feature Region

Retinal image registration using salient feature regions [6] like vessels based on a saliency metric that consists of both local adaptive variance and gradient field entropy was proposed recently. These salient feature regions are robust to intensity variation, background changes and pathologies. In order to extract efficiently the vasculature and the surrounding regions that have the most reliable and stable information of retinal image, we adopt Zheng’s method without the time-consuming accurate segmentation.

The measure of local intensity variance saliency of region $R$ is denoted as follows:

$$s_{li}(R) = \frac{\sigma}{\mu}$$  \hspace{1cm} (2)

where $\sigma$ is the standard deviation of $R$ and $\mu$ is the mean value of $R$. Since the main structures in retinal images, such as vessels, consist of combinations of edges, we use local gradient field entropy for measuring local structural saliency. Its measure for region $R$ is defined as follows:

$$s_{ls}(R) = -\sum_{\theta=1}^{36} p_{\theta} \log_2 p_{\theta}$$  \hspace{1cm} (3)

where $p_{\theta}$ is the probability of the gradient angle $\theta$, divided into 36 bins. To reduce unstable effects of low gradient magnitude, we use a gradient angle histogram, weighted according to its gradient magnitude. The local saliency measure of region $R$ combining local intensity variance with local gradient field entropy is given as follows:

$$s_l(R) = s_{li}(R) s_{ls}(R)$$  \hspace{1cm} (4)

We divide the whole image into $N \times N$ square sub-regions and then determine whether each sub-region is a salient feature region by thresholding the local saliency value of each sub-region. The example of salient feature region extraction is shown in Fig. 2.
2.4 Proposed Similarity Measure

This paper proposes a novel similarity measure to make retinal image registration robust to images containing non-uniform intensity, noise, background changes, or large homogeneous non-vascular regions. It computes the EMPCA-MI between salient feature regions including vasculature and surroundings of each retinal image. Let $S_R$ and $S_S$ be the salient feature regions in reference image $I_R$ and sensed image $I_S$. $S_U$ is a union of the two salient feature regions $S_R$ and $S_S$ in the overlapped region of the reference image $I_R$ and the sensed image $I_S$. First, $S_U$ in $I_R$ is pre-processed into vector forms $Q_{SR}$ and similarly the same $S_U$ in $I_S$ is also pre-processed into vector forms $Q_{SS}$ for a given neighborhood radius $r$. Then, the first principal components $X_{SR}$ and $X_{SS}$ of the respective $Q_{SR}$ and $Q_{SS}$ which represent the dominant features are iteratively computed by EMPCA. EMPCA-MI of salient feature regions is defined as:

$$F_s(I_R, I_S) = \sum_{x_{SR}, x_{SS}} p(x_{SR}, x_{SS}) \log \frac{p(x_{SR}, x_{SS})}{p(x_{SR})p(x_{SS})} \quad \quad (5)$$

By considering only salient feature regions, the method is (1) made robust to the above-mentioned challenging characteristics of retinal image, and (2) computationally more efficient. Since the entropy does not take into account information pertaining to the overlapped region size, the overlapped region may not correspond to the global maximum of the above measure. To handle the partial overlap problem, we introduce a weighting function representing the degree of overall similarity of salient features in the two images. The function is defined as follows:

$$W(I_R, I_S) = \frac{\sum_{x_{SR}, x_{SS} \in S_R \cap S_S} p(x_{SR}, x_{SS})}{N_{SR} + N_{SS}} \quad \quad (6)$$

where $N_{SR}$ and $N_{SS}$ are the total number of pixels of the salient feature regions $S_R$ and $S_S$, respectively. The weighting function $W$ represents the ratio of the number of coincident salient feature regions to that of the whole salient feature regions in the two images. Note that $W$ has a range of $0 \sim 1$ and increases when two images are registered. Finally, the proposed similarity measure $F_{WS}(I_R, I_S)$ becomes

$$F_{WS}(I_R, I_S) = W(I_R, I_S) \cdot F_s(I_R, I_S) \quad \quad (7)$$

The similarity measure $F_{WS}$ is always non-negative and should be maximized for registration. The function $W$ makes $F_{WS}$ relatively higher for partially overlapped but wrongly aligned regions. Thereby, $W$ tends to alleviate the partial overlap problem.

2.5 Iterative Optimization

A registration process is employed to find the best transformation parameters $p^*$ that maximize the similarity measure $F_{WS}$, namely

$$p^* = \arg \max_p F_{WS}(I_R, I_S; p) \quad \quad (8)$$

We adopt the similarity transformation as a model for registration; it can model rotation, translation, and isotropic scaling using four parameters $p = (S, \theta, t_x, t_y)$. This model represents retinal-image distortion caused by eye or camera motion or magnification changes due to the use of different equipment [15]. To maximize the proposed similarity measure, $F_{WS}$, by iteratively updating the transformation parameters, the Nelder-Mead simplex method [16] is used.

3. Experimental Results

To evaluate the performance of the proposed algorithm, the clinical dataset DRIVE [17], consisting of 20 color retinal images was used. Each image was acquired using a Canon CR5 non-mydriatic 3CCD camera with a resolution of $765 \times 584$ pixels and a 45 degree circular field of view. Each retinal image contained non-uniform illumination, low contrast and large homogeneous non-vascular regions making the registration challenging. Only the green channel of the retinal images is used because it has the highest contrast compared with the red and blue channels and vascular structures have the best contrast in the green channel. Since reference images were not available for this clinical dataset, the reference images are generated by mis-registering the original image considered as the sensed image using 20 sets of known transformation parameters (i.e., ground truth) for each image. These sets were characterized by the average distance between the positions of corresponding pixel pairs of the original and the transformed images and equally distributed into 20 average distance ranges of 1-20 pixels. The registration accuracy is evaluated by the average error which is the average distance between the positions of corresponding pixel pairs determined by the estimated transformation parameters and the ground truth. Additionally, standard deviation for the average errors of all sets is computed. To evaluate the registration robustness, success rate of the registration is computed for all sets of each image where the registrations are classified as “successful” if the average error is less than 2.5 pixels. We conducted a comparison between the proposed method and the EMPCA-MI [13]. The neighborhood radius $r$ was set to 2. The tolerance threshold for the Nelder-Mead simplex optimization was set to $10^{-4}$ while the maximum number of iterations was set at 200. The experimental results of average registration performance for all images are listed in Table 1.

The comparative experiments indicate that our method is more robust and accurate than the conventional EMPCA-MI method in variable contrast, illumination conditions, and

| Table 1 | Registration performance comparison of the proposed method and EMPCA-MI. |
|---------|-----------------|--------|--------|--------|
|         | Average error   | Standard deviation | Success rate (%) | Relative processing time |
| EMPCA-MI | 1.24           | 0.34              | 91.6             | 1.00               |
| Proposed | 1.18           | 0.29              | 93.8             | 0.48               |
large homogeneous non-vascular regions. In addition, the runtime performance was found to be the most attractive advantage of our method. The proposed method is about twice as fast as the conventional method. Figure 3 shows the registration results for image pair 10. In each figure, (a) and (b) are the reference image $I_R$ and the sensed image $I_S$, (c) and (d) show the checkerboard images of one pair before registration and after registration.

4. Conclusion

This paper proposed a new hybrid method for retinal image registration that effectively combines Expectation Maximization for Principal Component Analysis with Mutual Information (EMPCA-MI) with salient features. The proposed method is very effective for images with low contrast, non-uniform illumination, noise, and large homogeneous non-vascular regions. The experimental results show that our method is more effective and robust than the prevalent EMPCA-MI method.

Acknowledgments

Research was supported by Korea University.

References


