Landmark Matching Based Automatic Retinal Image Registration with Linear Programming and Self-similarities

Yuanjie Zheng[1], Allan A. Hunter III[2], Jue Wu[1], Hongzhi Wang[1], Jianbin Gao[3], Maureen G. Maguire[2], and James C. Gee[1]

[1] PICSL, Department of Radiology, University of Pennsylvania, Philadelphia, PA
[2] Department of Ophthalmology, University of Pennsylvania, Philadelphia, PA
[3] School of Computer Science & Engineering, University of Electronic Science and Technology of China

Abstract. In this paper, we address the problem of landmark matching based retinal image registration. Two major contributions render our registration algorithm distinguished from many previous methods. One is a novel landmark-matching formulation which enables not only a joint estimation of the correspondences and transformation model but also the optimization with linear programming. The other contribution lies in the introduction of a reinforced self-similarities descriptor in characterizing the local appearance of landmarks. Theoretical analysis and a series of preliminary experimental results show both the effectiveness of our optimization scheme and the high differentiating ability of our features.

1 Introduction

Retinal images are widely used in diagnosing and monitoring the progress of a variety of eye diseases, such as diabetic retinopathy, age-related macular degeneration, and glaucoma. As a process of establishing spatial correspondences between two retinal images, retinal image registration is fundamental to applications of as diverse as detecting locations of scars and burns, tracking the progress of diseases, mosaicing to provide a complete view of the retina, and creating tools to assist ophthalmologists during laser surgery or other procedures [4].

However, automatic retinal image registration has remained a difficult problem due to several challenges. First, the transformation between two retinal images to be aligned can be very complicated. Retina is a curved surface and retinal images are projections of this surface from a wide range of view points with an un-calibrated camera. Second, there exist large homogeneous areas in retinal images. Nonvascular surface occupies an overwhelming area of the retina, which is textureless for healthy retinas while exhibits a variety of pathologies for unhealthy retinas. Finally, large appearance variations may be observed between the images to be registered. It is because the pathologies can appear and disappear over time, the effects of disease and poor image quality can obscure the vasculature [16], and images may be captured under vastly different illumination

G. Székely and H.K. Hahn (Eds.): IPMI 2011, LNCS 6801, pp. 674–685, 2011.

conditions or acquired from different modalities. To solve this hard problem of retinal image registration, there are two keys to a successful algorithm: selection of an optimization strategy to more efficiently estimate the transformation model no matter the model is simple or complicated, and extraction of reliable features robust to image variations.

Landmark matching [4,17,5,16] is more attractive in retinal image registration compared with pixel-wise registration (a brief review to previous registration techniques is included in Sec. 2). It is because that the large homogenous areas and appearance variations in retinal images make the pixel-wise registration very hard and that landmarks are also used in practice by ophthalmologist [4]. Disregarding certain encouraging results produced by some previous methods [4,16,5], two major weaknesses were recognized in our research from many previous works. *First*, optimization strategies for estimating the transformation model are less efficient. Most previous methods perform an iterative optimization process by repeating the estimations of correspondences and transformation model, which requires a good initialization and results in an inferior local estimation. As discovered by us and also mentioned in [4], joint estimation of the correspondences and an appropriately predefined transformation model can result in a superior estimation. Unfortunately, no effective optimization approach has been described yet to accomplish this. *Second*, features describing landmarks are not always distinguishable. Most previous features (e.g. [5]) were extracted from a binary vascular image in which a lot of information of the local image pattern is lost. Moreover, landmarks are not uniquely identified by the widely used features characterized with vessel orientations [4,17,16] because branching angles at bifurcations can be similar across the retina.

This paper addresses the problem of landmark matching based automatic retinal image registration and bears two major contributions. One is a novel landmark-matching formulation proposed to accomplish a joint estimation of the correspondences and transformation model. This new formulation enables the matching problem to be solved by a very efficient combinatorial optimization technique - linear programming (LP) [11]. The other contribution is the introduction of a reinforced local self-similarities descriptor [14] in characterizing the local appearance of landmark. This descriptor captures the geometric layout in nearby image locations of the repetitions of local intensity-pattern. The features from this descriptor are invariant not only to local affine deformations but also large rotations of image. Theoretical analysis and a series of preliminary experimental results show that this joint estimation strategy using the LP techniques is very effective with the ability to produce superior accuracy and robustness, and that the proposed features bear a high differentiating ability.

2 Previous Work

Image registration has been extensively involved in a wide variety of practical tasks in the fields of retinal image analysis, medical image analysis, computer vision, etc. There exist a huge number of related publications. These

previous methods can basically be categorized as pixel/voxel-wise based, landmark/feature based, or their combinations. Detailed review to all of them is beyond the scope of this paper. We below only provide a brief review to the specific methods of retinal image registration and to the optimization techniques of the landmark matching based image registration.

For retinal image registration, both landmark matching based methods [4,17,5,16] and pixel-wise strategies [10,12] were explored. However, landmark matching has attracted more interests due to several reasons as explained in Sec. 1. In most previous works, landmarks were placed at cross points or bifurcations of vessel and characterized by for example coordinates, orientations or widths of the surrounding vessels. However, it is well known that these features are not robust and not always well distinguishable in practice, introducing severe matching ambiguities. The inter-image motion was described by various models such as weak affine [17], affine [4,10,12], bilinear/projective [12] and quadratic [4]. Optimization techniques include iterative closest point (ICP) algorithm [5], dual-bootstrap ICP [16], Hough Transform [17], and traditional global optimization techniques such as simulated annealing and genetic algorithm [12].

In those more general fields of medical image analysis and computer vision, landmark matching based image registration has been treated as a fundamental task [7,2]. To obtain a high matching accuracy, optimization plays a vital role. There are basically three classes of optimization techniques previously used. The first class assumes the correspondences to be pre-known and only resolves the transformation model [15,3,13]. The second class does not explicitly assume any specific transformation model but solely solves the correspondences with various optimization techniques such as the Hungarian method [1] and the quadratic programming [2]. The third class handles the joint estimation of the correspondences and transformation model. However, for simplicity, an iterative process by repeating the estimations of each of them was more widely used, e.g. the EM like technique in [6] and the iteratively updating process in [18]. With this iterative strategy, only a local optimization can be attained and the accuracy is highly conditioned on the initialization. In contrast, recently explored joint estimation with combinatorial optimization techniques [9] is free from the iterative process and can simultaneously solve the correspondences and transformation model. These methods are free from initialization and able to produce a global optimization. Unfortunately, only the simple rigid spatial transformation has been investigated [9].

3 Landmark Matching with Linear Programming

We provide a new formulation for the landmark matching problem. It bears at least two benefits compared with previous work. First, correspondences and the transformation model can be solved simultaneously. This joint estimation can result in superior accuracies because the correspondences and transformation model can assist each other in the optimization process compared with estimating them independently. The other benefit is that the new formulation enables

the optimization solved by linear programming (LP) - an important class of combinatorial optimization techniques. LP is attracting more and more interests in computer vision and medical image analysis due to its ability to obtain a globally optimal solution and the availability of more and more efficient and free solvers online. However, few work with LP can handle the joint estimation.

3.1 Problem Formulation

Suppose we have one reference-landmark set $\Omega^r = \{\xi_i^r, i = 1, 2, \cdots, N_r\}$ and one floating-landmark set $\Omega^m = \{\xi_i^m, i = 1, 2, \cdots, N_m\}$ where N_r and N_m represent the number of reference-landmarks and floating-landmarks, respectively. Taking the reference-landmark i as an example, ξ_i^r is expressed by the features (denoted by vertical vector $\mathbf{v}_i^r$) extracted to describe the visual appearance of this point and its coordinates $\chi_i^r = [x_i^r \ y_i^r]^T$, i.e. $\xi_i^r = \left[\mathbf{v}_i^{rT} \ \chi_i^{rT}\right]^T$ where T means transpose. For later usages, we provide as well the vector $\chi'^r_i = [x_i^r \ y_i^r \ 1]^T$ to denote the corresponding homogeneous coordinates, vector $\chi''^r_i = \left[x_i^{r2} \ x_i^r y_i^r \ y_i^{r2} \ x_i^r \ y_i^r \ 1\right]^T$ to represent the 12-parameter transformation model in [4], matrix $\chi^r = \left[\chi_1^r \ \chi_2^r \ \cdots \ \chi_{N_r}^r\right]^T$, matrix $\chi'^r = \left[\chi'^r_1 \ \chi'^r_2 \ \cdots \ \chi'^r_{N_r}\right]^T$, and matrix $\chi''^r = \left[\chi''^r_1 \ \chi''^r_2 \ \cdots \ \chi''^r_{N_r}\right]^T$. We omit the definitions of similar symbols for the floating-landmarks for brevity.

Point matching is defined as estimating a mapping σ which indicates that a floating-landmark i_m corresponds to a reference-landmark i_r, i.e. $i_r = \sigma(i_m)$. Mapping σ can be represented by a correspondence/assignment matrix (denoted by E) [9] or a transformation model (denoted by $\mathcal{T}$). $\mathcal{T}$ is usually represented by a parametric model in order for the robustness to noise/outliers and for reducing the involved computational burden.

Our point matching scheme is formulated as jointly estimating the correspondences and transformation model by maximizing the feature matching quality and the transformation model compliance quality. Feature matching quality represents how similar the features $\mathbf{v}^m$ of each floating-landmark are to $\mathbf{v}^r$ of its corresponding reference-landmark. Transformation model compliance quality evaluates how well the coordinates $\mathbf{x}^m$ of each floating point and $\mathbf{x}^r$ of its corresponding reference-landmark comply with the estimated transformation model.

Correspondence Matrix. We first define the correspondence matrix E as a binary matrix and then relax it later. E is of size $N_m \times N_r$, for which each row contains exactly one 1. If a floating-landmark i_m is matched to the reference-landmark i_r, then $E_{i_m,i_r} = 1$ and other elements of the i_mth row all equal to 0. These can be stated alternatively as the following two constraints:

$$E \in \{0, 1\}^{N_m \times N_r}, \tag{1}$$

$$E\mathbf{1} = \mathbf{1}, \tag{2}$$

where $\mathbf{1}$ is a vertical vector of ones.

Discreteness of the values taken by the elements of E in Eq. (1) introduces hardships into designing an efficient optimization algorithm for landmark matching. Many papers [6,9] tried to relax it to a continuous value within $[0, 1]$. This softassign strategy [6] can make the resulting energy function better behaved. This relaxation can be guaranteed by Eq. (2) together with the below constraint

$$E \geq 0. \tag{3}$$

Unfortunately, we found through our experiments that the softassign strategy can also cause ambiguities in the matching process. Imagine that one floating-landmark is matched to five reference-landmarks and the matching likelihood value equals to 0.2 for each of them. We then have no idea to decide the assignment. To deal with this dilemma, we propose to penalize the to-centroid spatial deviations of the matched reference-landmarks to enforce a spatial concentration.

To define the to-centroid spatial deviations, we first locate the centroid of the matched reference-landmarks by computing its coordinates with

$$\bar{\chi}^r = E\chi^r. \tag{4}$$

We then use $\bar{\chi}^r_x$ and $\bar{\chi}^r_y$ to denote the vertical vectors concatenated by the x-axis and y-axis coordinates of all centroids, respectively, and denote $\bar{\chi}^r = \left[\bar{\chi}^r_x \ \bar{\chi}^r_y\right]$. Then, the to-centroid deviation matrix is written as

$$D = \sqrt{\left(\bar{\chi}^r_x \mathbf{1}^T - \mathbf{1}(\chi^r_x)^T\right)^2 + \left(\bar{\chi}^r_y \mathbf{1}^T - \mathbf{1}(\chi^r_y)^T\right)^2} \tag{5}$$

where lengths of the $\mathbf{1}$s are set appropriately to get a matrix D in size $N_m \times N_r$.

Penalizing the to-centroid deviations amounts to minimizing the function

$$\mathcal{O}_c = tr(D^T E), \tag{6}$$

where $tr()$ denotes the trace of matrix. Minimization of the objective function in Eq. (6) with the constraints in Eq. (2) and Eq. (3) and other criteria explained below will result in a correspondence matrix E taking a continuous values in $[0, 1]$ but with a significant bias towards 0 or 1.

Feature Matching Quality. Feature matching quality measures how similar the visual appearance of each floating-landmark is to the matched reference-landmark(s). We use the reinforced self-similarities (to be proposed in Sec. 4) to describe the landmarks, which is invariant to local affine deformation, radially increasing non-rigid deformation, and rotation. We use the negative of the correlation between two series of features as the matching cost and then obtain a feature matching cost matrix C in size $N_m \times N_r$ for which the element $C_{i,j}$ denotes the cost matching the ith floating-landmark to the jth reference-landmark. Note that C can be computed beforehand.

Maximization of feature matching quality is then expressed as the minimization of the below objective function:

$$\mathcal{O}_f = tr(C^T E). \tag{7}$$

Transformation. Most transformation models can be incorporated in our landmark matching scheme, such as the widely used general models as diverse as affine, elastic, quadratic, deformable represented by Thin-plate Spline (TPS) or others, and some specific models designed for retinal imaging, e.g. the model in [4]. In the followings, we provide a discussion on the model in [4] and the TPS based deformable model as examples.

We propose to apply the constraints of the predefined transformation model on the location relations of each floating-landmark and the corresponding centroid of its matched reference-landmarks. As an example, the 12-parameter transformation model in [4] can be enforced with the following equation

$$\bar{\chi}^r = \chi''^m \mathbf{\Theta}^T, \tag{8}$$

where $\bar{\chi}^r$ and χ''^m are defined previously and $\mathbf{\Theta}$ is a 2×6 matrix. In Eq. (8), only the elements of $\mathbf{\Theta}$ are unknowns. This transformation model is obtained in [4] for retinal imaging by assuming a quadratic surface for the retinal, a rigid transformation between two viewpoints, and a weak-perspective camera projection model. All involved parameters are combined in $\mathbf{\Theta}$.

When the transformation between the two landmark point sets is deformable, the TPS model [6] can be employed, as expressed by

$$\bar{\chi}^r = \chi^m + \chi'^m A^T + \mathbf{\Phi} \Delta^m \tag{9}$$

where A is a 2×3 matrix containing the parameters of an affine transformation, $\mathbf{\Phi}$ is a $N_m \times N_m$ symmetric matrix containing the information about the floating-landmark set's internal structural relationships and its elements can be precomputed as $\mathbf{\Phi}_{i,j} = \|\chi_i^m - \chi_j^m\|^2 \log \|\chi_i^m - \chi_j^m\|$, and Δ^m is an $N_m \times 2$ matrix for which each row denotes the non-affine deformation of the corresponding floating-landmark. In Eq. (9), A and Δ^m are unknowns and need to be estimated.

Note that Eq. (8) and Eq. (9) are all linear to the unknowns, which leads to a linear formulation which can be solved with the LP techniques.

3.2 Optimization

The optimization scheme for our point matching task can be formulated as a constrained minimization problem of an integrated objective function:

$$\mathcal{O} = \mathcal{O}_f + \lambda \mathcal{O}_c = tr\left((C + \lambda D)^T E\right) \tag{10}$$

where $\mathcal{O}_f$ and $\mathcal{O}_c$ are defined in Eq. (7) and Eq. (6), respectively, and λ (e.g. 0.5) is an adjusting parameter to control the "softness" of the estimated correspondence matrix E. For this optimization scheme, constraints are expressed in Eq. (2), Eq. (3), Eq. (8) or Eq. (9), unknowns are the correspondence matrix E and the transformation parameters: $\mathbf{\Theta}$ in Eq. (8), or A and Δ^m in Eq. (9).

One benefit of our optimization scheme is that it can be resolved with linear programming (LP) [11]. It is obvious because the objective function and the constraints involved in the optimization scheme are all linear to the unknowns if

removing $\mathcal{O}_c$ from Eq. (10) (i.e. setting $\lambda = 0$). With $\mathcal{O}_c$, the objective function is not linear any more due to the nonlinearity of $\mathcal{O}_c$ to E in Eq. (6). Fortunately, $\mathcal{O}_c$ can be linearized by computing D with Eq. (5) while fixing E in Eq. (4). Hence, our optimization scheme beomes an iterative process. At each iteration, the estimation of E in previous iteration is used to compute D. For the first iteration, E can be estimated by removing $\mathcal{O}_c$ from Eq. (10). *This means that our landmark matching technique does not require any manual initialization.*

Formulation as Linear Programming. We provide expressions in a canonical form for our LP based optimization scheme by taking the transformation in Eq. (8) as an example. Other models can be handled with a similar way.

We first define some additional symbols by treating all elements in the correspondence matrix E, matrix $\mathbf{\Theta}$ in Eq. (8), matching cost matrix C, and the to-centroid deviation matrix D as a single column and denoting them with vertical vectors $\mathbf{e}$, θ, $\mathbf{c}$, and $\mathbf{d}$, respectively.

We then create a binary permutation matrix P_e in size $N_m \times (N_m \cdot N_r)$, such that $P_e\mathbf{e} = E\mathbf{1}$. Then, we have the reformulation of Eq. (2):

$$P_e\mathbf{e} = \mathbf{1}. \tag{11}$$

We further divide Eq. (8) into two parts: x-mapping and y-mapping, and represent each of them as a linear function of $\mathbf{\Theta}$. The x-mapping is written as

$$\bar{\chi}_x^r = \chi_x^m + \chi'^m \mathbf{\Theta}_x^T \tag{12}$$

where $\bar{\chi}_x^r$ is the first column corresponding to the x-coordinates of $\bar{\chi}^r$, χ_x^m is the first column of χ^m, and $\mathbf{\Theta}_x$ is the first row of $\mathbf{\Theta}$ and in charge of x-mapping. It is easy to further create a permutation matrix P_θ^x in size of 6×12, such that $\mathbf{\Theta}_x^T = P_\theta^x \theta$. Therefore, Eq. (12) can be rewritten as

$$\chi'^m P_\theta^x \theta = \bar{\chi}_x^r - \chi_x^m. \tag{13}$$

With $Q_\theta^x = \chi'^m P_\theta^x$ and $b_\theta^x = \bar{\chi}_x^r - \chi_x^m$, Eq. (13) can be rewritten concisely as

$$Q_\theta^x \theta = b_\theta^x. \tag{14}$$

Similarly y-mapping is expressed as

$$Q_\theta^y \theta = b_\theta^y. \tag{15}$$

We are now ready to provide the formulation of linear programming:

$$\min \begin{bmatrix} \mathbf{c} + \lambda\mathbf{d} \\ \mathbf{0} \end{bmatrix}^T \begin{bmatrix} \mathbf{e} \\ \theta \end{bmatrix} \tag{16}$$

subject to

$$\begin{bmatrix} P_e & \mathbf{0} \\ \mathbf{0} & Q_\theta^x \\ \mathbf{0} & Q_\theta^y \end{bmatrix} \begin{bmatrix} \mathbf{e} \\ \theta \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ b_\theta^x \\ b_\theta^y \end{bmatrix} \tag{17}$$

and

$$-\mathbf{e} \leq 0. \tag{18}$$

In this formulation, $\mathbf{e}$ and θ are unknowns.

There are different ways to solve a LP problem and various solvers are available online. We chose the interior point solver of GLPK[1] for its efficiency.

Algorithm Overview. The optimization of our landmark matching approach is briefed as the following five steps:

1. Solving the LP problem specified by Eq. (16), Eq. (17), and Eq. (18) without considering the to-centroid deviations (i.e. set $\lambda = 0$ in Eq. (16)).
2. Updating the to-centroid deviation matrix using Eq. (5) with current estimation of the correspondence matrix.
3. Increasing λ by $\delta\lambda$ (e.g. 0.2).
4. Solving the LP problem specified by Eq. (16), Eq. (17), and Eq. (18) with considering the to-centroid deviation matrix.
5. Go to step 2 until the stopping criterion is satisfied.

The high accuracy of our algorithm comes from its at least five unique characteristics. First, step 1 conveniently provides automatic initialization. Second, steps 1 and 4 provide optimal solution with the LP techniques. Third, the correspondences and transformation model are estimated simultaneously in step 4 and the transformation model can help to guide the estimation of correspondences. Fourth, gradually increasing λ in step 4 approaches a binary correspondence matrix with less ambiguity in the assignment. Finally, the transformation model help to resist noise and outliers.

Our algorithm can produce an accurate estimation by repeating steps 2-5 for at most 3 times. Through our experiments, it can finish the optimization process within 1.5 seconds to match about 180 landmarks using a Dell PC with 2.39 GHz Intel Core 2 CPU. At the end, each floating-landmark is assigned the reference-landmark with the largest estimated correspondence likelihood value.

4 Landmark Detection and Feature Extraction

Landmark detection and feature extraction are also two important procedures of our retinal image registration approach. For landmark detection, the vascular-landmark detection technique in [4] is employed. Several steps are involved in this detection process: specification of seed points of vessel using 1D edge detection along a series of evenly spaced rows and columns, tracing out vascular structures with an exploratory procedure, and detection of landmarks by locating intersections of traces where three or more traces meet.

To extract features characterizing the local appearance of each landmark, we introduce a reinforced self-similarities (SS) descriptor. The SS descriptor was

[1] http://gnuwin32.sourceforge.net/packages/glpk.htm

recently proposed in [14] to measure similarity between two visual entities of images or videos. SS has been shown to bear high differentiating ability. Compared with SS, most other widely known descriptors (e.g. SIFT [3]) assume the existence of a common underlying visual property between matching images, e.g. intensities or responses of certain filters. SS relaxes this assumption and tries to capture the geometric layouts of local intensity-pattern's repetitions in nearby image locations. In other words, SS describes the local internal layouts of the self-similarities instead of the intensity pattern itself. This is very important for a robust descriptor. For example, landmarks of retinal images from different modalities can have very different local intensity patterns but the geometric layouts of pattern's repetitions in the vicinity of the landmark are very similar. As a result, SS can be used in both intra-modality and inter-modality retial image registrations. We omit the details of computing SS for brevity. They can be found in [14].

The proposed reinforced SS descriptor is based on the original SS descriptor but tailored to be invariant to large image rotations. Although the original SS features are robust to local affine deformation and radially increasing non-rigid deformation [14], they become fragile when a large rotation happens between the matching images. It is because both the order of the entries in the SS descriptor vector and the partitioned bins [14] can change with rotation. Invariance to large rotations is particularly demanding for the retinal images considering the large range of view points involved in the retinal imaging process [5]. This rotation invariance can be simply accomplished by computing Hu's seven moment invariants [8] of the outputs of the original SS descriptors. Therefore, the reinforced SS results in a feature vector with seven entries for each landmark point.

5 Results and Discussions

We performed a series of experiments to evaluate our retinal image registration algorithm both quantitatively and qualitatively. The transformation of each experimental image pair was formed synthetically by ourselves or introduced in the imaging process.

5.1 Synthetic Transformation

In order to quantitatively compare our method with state-of-the-art, we chose a digital fundus retinal image (as shown in Fig. 1 (a)), warped it with 50 synthetic affine transformations, and then ran our LP based method with the features in [4] (denoted by "LP"), our LP based method with the reinforced SS features introduced in this paper (denoted by "LP-SS"), the hierarchical method in [4] (denoted by "Hierarchical") and the hybrid method in [5] (denoted by "Hybrid"). These synthetic affine warps were created by randomly choosing a rotation angle in $[0\ 2\pi]$, translation in $[-8\ 8]$ pixels for both directions, scale in $[0.3\ 3.5]$, and shear in $[-1.8\ 1.8]$ for both directions. Note that for each pair of images, the ground-truth position in the warped image of each pixel in the original image is known. To measure the registration accuracy, we employed the mean square

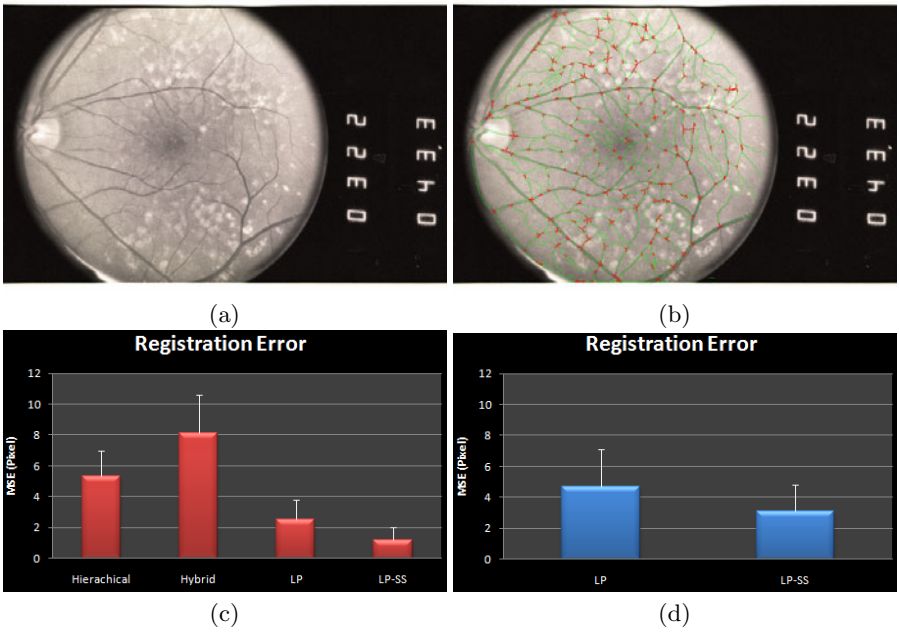

Fig. 1. (a): A digital fundus retinal image. (b): Bifurcation/crossover landmark points (red) and vessel's centerline pixels (green) detected by the technique in [4]. (c) and (d): Registration error bars over 50 synthetic affine transformations and 50 synthetic nonlinear deformations, respectively.

error (MSE) between the ground-truth position and the estimation over all pixels in the original image. As shown by the error bars in Fig. 1 (c), we can see that our LP based algorithms produced higher accuracies, showing the effectiveness of our optimization technique, and that the reinforced SS features outperformed the features in [4], showing the superior differentiating ability of our features.

We also tested our algorithm on 50 nonlinear deformation fields simulated with the TPS model. We first manually specified 15 bifurcation/crossover points in the image and treated them as the control points of the TPS model. We then then synthesized these fields by randomly setting the affine transformation with the same way as above and randomly displacing control points in a range $[-4\ 4]$ pixel for both directions. The registration errors of our LP algorithm with the features in [4] and the reinforced SS features are shown in Fig. 1 (d). We can see that both cases performed well on the nonlinearly deformed images while the reinforced SS features produced higher accuracies than the features in [4].

5.2 Real Transformation

We chose 19 pairs of retinal images and tested on them with our LP based algorithm using the reinforced SS features and the hierarchical approach in [4]. The centerline error described in [4] was used as the error measure for validation.

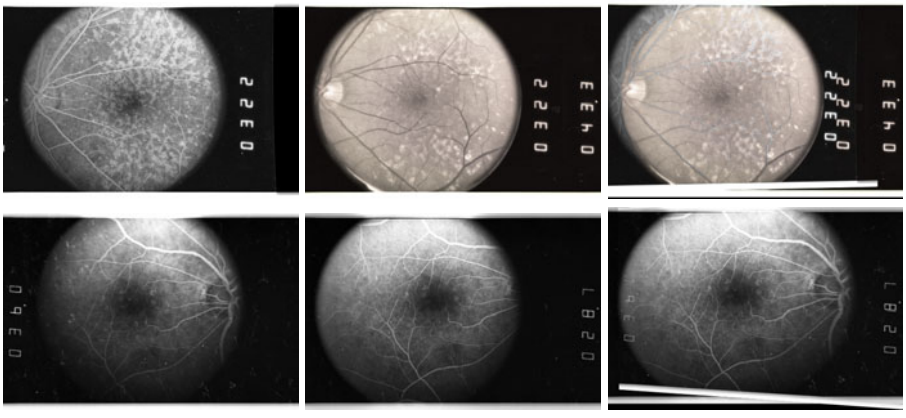

Image 1 Image 2 Mosaiced Image

Fig. 2. Two example pairwise image mosaics produced by our approach

The hierarchical approach in [4] resulted in a mean/std error value of 4.3/3.8, and in contrast, our method achieved a mean/std error value of 1.2/0.7, showing the obvious performance improvements of our method.

Two examples of pairwise image mosaic formed by our retinal registration algorithm are shown in Fig. 2. Disregarding the image modality difference (fluorescein angiogram image vs. digital fundus image) and image changes caused by drusen in the above pair and the sever bias field artifact in the below pair, the mosaicing accuracy of our method is perceptually very high.

6 Conclusion and Future Work

We proposed a new landmark matching based retinal image registration scheme which bears at least two superior properties. First of all, a better optimization can be achieved. It originates from the proposed novel landmark matching formulation which enables the linear programming techniques in solving the matching problem and enables a joint estimation of correspondences and transformation model. Moreover, a lower matching ambiguity can be obtained. It benefits from the proposed reinforced self-similarities (SS) descriptor. The features from this descriptor were shown to yield a higher differentiating ability in both intra-modality and inter-modality registrations through our experiments.

Our future work would include more validation experiments, extension of our method to the 3D case, and applications to other images in various modalities (e.g. MR brain images and CT lung images). The extension to a 3D case can simply be achieved by extending the reinforced SS descriptor to 3D and including a 3D transformation model.

Acknowledgement

The authors gratefully acknowledge NIH support of this work via grants P30 EY001583 and P30 NS045839.

References

1. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE TPAMI 24, 509–522 (2001)
2. Berg, A.C., Berg, T.L., Malik, J.: Shape matching and object recognition using low distortion correspondence. In: CVPR, pp. 26–33 (2005)
3. Brown, M., Lowe, D.G.: Automatic panoramic image stitching using invariant features. IJCV 74, 59–73 (2007)
4. Can, A., Stewart, C.V., Roysam, B., Tanenbaum, H.L.: A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. IEEE TPAMI 24, 347–364 (2002)
5. Chanwimaluang, T., Fan, G., Fransen, S.R.: Hybrid retinal image registration. IEEE Transactions on Information Technology in Biomedicine 10, 130–142 (2006)
6. Chui, H., Rangarajan, A.: A new point matching algorithm for non-rigid registration. Computer Vision and Image Understanding 89, 114–141 (2003)
7. Gholipour, A., Kehtarnavaz, N., Briggs, R., Devous, M., Gopinath, K.: Brain functional localization: a survey of image registration techniques. IEEE TMI 26, 427–451 (2007)
8. Hu, M.K.: Visual pattern recognition by moment invariants. IRE Transactions on Information Theory, 179–187 (1962)
9. Jiang, H., Yu, S.X.: Linear solution to scale and rotation invariant object matching. In: CVPR (2009)
10. Kolar, R., Kubecka, L., Jan, J.: Registration and fusion of the autofluorescent and infrared retinal images. IJBI 2008, 513478 (2008)
11. Matousek, J., Gartner, B.: Understanding and Using Linear Programming. Springer, Heidelberg (2007)
12. Matsopoulos, G.K., Mouravliansky, N.A., Delibasis, K.K., Nikita, K.S.: Automatic retinal image registration scheme using global optimization techniques. IEEE TMI 3, 47–60 (1999)
13. Rohr, K., Stiehl, H.S., Sprengel, R., Buzug, T.M., Weese, J., Kuhn, M.H.: Landmark-based elastic registration using approximating thin-plate splines. IEEE TMI 20, 526–534 (2001)
14. Shechtman, E., Irani, M.: Matching local self-similarities across images and videos. In: CVPR (June 2007)
15. Shen, D., Davatzikos, C.: Hammer: Hierarchical attribute matching mechanism for elastic registration. IEEE TMI 21, 1421–1439 (2002)
16. Stewart, C.V., ling Tsai, C., Roysam, B.: The dual-bootstrap iterative closest point algorithm with application to retinal image registration. IEEE TMI 22, 1379–1394 (2003)
17. Zana, F., Klein, J.C.: A multimodal registration algorithm of eye fundus images using vessels by hough transform. IEEE TMI 18, 419–428 (1999)
18. Zheng, Y., Doermann, D.: Robust point matching for nonrigid shapes by preserving local neighborhood structures. IEEE TPAMI 28, 643–649 (2006)