

RESEARCH

Open Access

An efficient approach for robust multimodal retinal image registration based on UR-SIFT features and PIIFD descriptors

Zeinab Ghassabi^{1*}, Jamshid Shanbehzadeh², Amin Sedaghat³ and Emad Fatemizadeh⁴

Abstract

Existing algorithms based on scale invariant feature transform (SIFT) and Harris corners such as edge-driven dual-bootstrap iterative closest point and Harris-partial intensity invariant feature descriptor (PIIFD) respectively have been shown to be robust in registering multimodal retinal images. However, they fail to register color retinal images with other modalities in the presence of large content or scale changes. Moreover, the approaches need preprocessing operations such as image resizing to do well. This restricts the application of image registration for further analysis such as change detection and image fusion. Motivated by the need for efficient registration of multimodal retinal image pairs, this paper introduces a novel integrated approach which exploits features of uniform robust scale invariant feature transform (UR-SIFT) and PIIFD. The approach is robust against low content contrast of color images and large content, appearance, and scale changes between color and other retinal image modalities like the fluorescein angiography. Due to low efficiency of standard SIFT detector for multimodal images, the UR-SIFT algorithm extracts high stable and distinctive features in the full distribution of location and scale in images. Then, feature points are adequate and repeatable. Moreover, the PIIFD descriptor is symmetric to contrast, which makes it suitable for robust multimodal image registration. After the UR-SIFT feature extraction and the PIIFD descriptor generation in images, an initial cross-matching process is performed and followed by a mismatch elimination algorithm. Our dataset consists of 120 pairs of multimodal retinal images. Experiment results show the outperformance of the UR-SIFT-PIIFD over the Harris-PIIFD and similar algorithms in terms of efficiency and positional accuracy.

Keywords: Scale invariance, Feature distinctiveness, Uniform spatial and scale distribution, Multimodal image registration, Retinal images

1. Introduction

Digital retinal images are widely used in the diagnosis and treatment of eye disorders such as glaucoma, diabetic retinopathy, and age-related macular degeneration [1,2]. In most cases, this requires image registration (IR) which is the process of geometrically aligning images of the same scene. There are three groups of IR applications: image mosaicking, temporal registration, and multimodal registration [3,4]. The first one aligns images of different viewpoints of one scene to generate a wider view of it. The second one aligns images of the same

scene obtained at different times to detect the changes. The third one aligns images obtained by different sensors to receive more complete information about the subject and integrate their data. For example, superimposing the unhealthy angiographic image on the red-free (RF) image improves the pathological and visibility of structural information of the second image for retina surgery [5]. Registration of the fluorescein angiographic (FA) and color images is helpful for accurate determination of lesion types and evaluation of various blood vessel abnormalities [1,6].

A variety of image registration techniques are proposed for retinal image registration (RIR) [7]. Tsai et al. [8] introduced several RIR softwares and commercial packages on the Web. Multimodal RIR methods are classified as area-

* Correspondence: z.r.ghassabi@gmail.com

¹Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran 4933-14155, Iran

Full list of author information is available at the end of the article

based (AB) and feature-based (FB). Hybrid approaches were also used [9,10].

AB methods choose a similarity metric and maximize it to find parameters of transformation function. Most common similarity measures in the context of multimodal RIR are mutual information (MI) [11,12], entropy correlation coefficient (ECC) [13], and phase correlation [14]. However, MI performance degrades when faced with a large amount of changes in the texture of retinal image and changes in scale [15]. There may be many local maxima in the MI function, which cause problems with optimization methods and lead to misalignment. MI is also weak in registering image pairs with too small overlaps. Therefore, ECC, a normalized measure of MI, was used on the vascular tree to register small overlapping images [13]. Its dependency on vessels restricts the efficiency of registration techniques for low-quality images. Phase-based methods are robust to lighting variations [16]. However, they fail to register images with high translation and content changes [11]. AB methods are computationally intensive because of using the entire content of images.

FB methods involve finding salient features like points, vessels, and regions in image pairs to compute the correspondences and parameters of the transformation function. FB methods utilize anatomical features like vessel bifurcations [9,17,18], crossovers, endpoints of vessels, and other point features like corners, points of high curvature, and Lowe's keypoints to register images. Bifurcations are invariant to intensity variations. However, their localization may be inaccurate [19]. Even if algorithms extract accurate bifurcation points, their extraction is difficult in low-quality and unhealthy images, and thus, they are not repeatable and not well distributed for a robust and accurate registration [20]. Corners and Lowe's keypoints are independent of vasculature and easier to detect than bifurcations in low-quality images [20]. Although corner points may be sufficient and uniformly

distributed in some retinal image modalities, they are not highly distinctive, repeatable, and scale invariant in the presence of color, and the FA image pairs with high content or scale changes [21] (Figure 1). Scale changes in multimodal retinal images are because of different resolutions of images and changes in the distance between the camera and the head.

Recently, Lowe's scale invariant feature transform (SIFT) algorithm [22] has been widely used for FB image registration. SIFT uses scale-space theory [23] to find important keypoints of image robust to scale changes. However, it is unable to identify adequate, stable, repeatable, and uniformly distributed features in multimodal retinal images [20,24]. Therefore, the edge-driven dual-bootstrap iterative closest point algorithm (ED-DB-ICP) [25] incorporated Lowe's keypoints with centerline points or edge points to register the RF images and FA sequences. Although estimating transformation parameters and determining correspondences are done at the same time using only one initial correct match [26], the approach is inadequate for registration of some multimodal retinal image pairs with scale changes, and when there are homogeneous non-vascular regions in high-resolution color images in contrast to texture regions of high-resolution FA images. An improved version of SIFT, i.e., mSIFT uses color information for extracting sufficient feature points [23]. However, it is inapplicable for multimodal images.

Regardless of feature point type, FB matching methods automatically find correspondences between two sets of features. Finding correspondences comprises methods using spatial relations after prematching, and methods using robust descriptors of features followed mismatch elimination. The former category used simple feature descriptor such as MI around feature points [13] or angle-based descriptor around bifurcations [26-29] to compute initial matches. Then, they used robust alignment algorithm to choose the best subset of matched

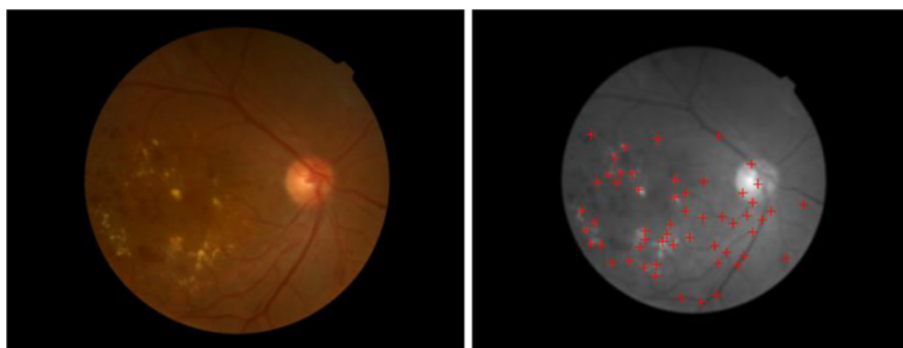


Figure 1 Distribution of features for the Harris corner detector in a color image. Corner points (i.e., red plus signs) are detected on non-vascular or pathological areas. This degrades feature matching performance due to non-repeatable locations in the second image in spite of vessels which are depicted in all modalities.

points against incorrect matches and multiple matches. The methods involved graph matching [30,31], self-organizing maps [32], expectation maximization [33,34], Hough transform [28], and relaxation labeling [6]. The methods require sufficient feature points to perform efficiently. Moreover, using invariant descriptors based on anatomical feature point information in low-quality or unhealthy retinal images limits sufficient detectable points. In addition, angle-based descriptors based on orientation and width information may be similar, and the points will not be distinctive.

Most image registration methods are based on local feature descriptor matching techniques. They consist of feature point detection, building local descriptors around the detected feature points and matching these descriptors. A review of local descriptors for feature matching can be found in [35]. SIFT is a popular local descriptor [22,36-38]. After extracting Lowe's keypoints in SIFT, a main orientation is assigned to the points. Descriptors are relative to this orientation and thus achieve invariance to image rotation. SIFT uses a nearest neighbor distance ratio strategy to match feature vectors.

There are two problems related to SIFT-based multimodal RIR in the sections of feature detector and descriptor. Recently, some improvements of feature descriptor part like restricted SIFT [39] and partial intensity invariant feature descriptor (PIIFD) [20] achieved satisfactory results. However, SIFT detector suffers from the quantity, quality, and distribution of extracted points particularly in the registration of multimodal images like angiographic and color images wherein there are

low-contrast color images or scale, content, and appearance changes between multimodal images. This paper improves the multimodal RIR by extracting well-distributed repeatable, reliable, and precisely aligned point pairs. The organization of this paper is as follows: The first part describes problems related to multimodal SIFT-based RIR. The second section presents the proposed integrated matching algorithm which makes robust multimodal RIR. Then, experimental results illustrate the effects of the proposed improvement.

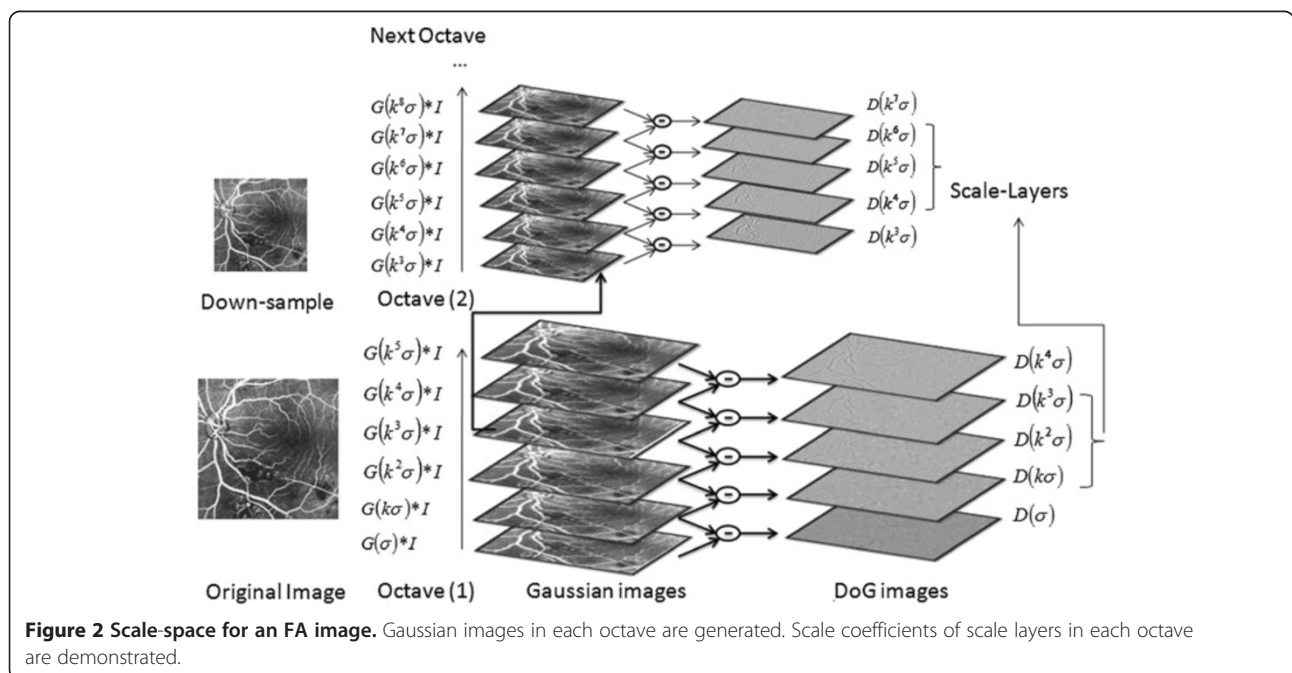
2. Problems related to the SIFT algorithm when applied to multimodal RIR

This part describes three problems of the SIFT-based retinal image registration and introduces related proposed solutions in the literature. These problems related to the feature detector, feature descriptor, and elimination of outliers are due to non-linear intensity changes among multimodal retinal images and the existence of repetitive patterns in the retinal images.

2.1 Problems related to the feature detection part

There are two steps to implement the standard SIFT detector algorithm:

1. *Detecting extrema using scale-space theory.* Scale-space theory provides difference of Gaussian (DoG) images for the SIFT to find scale invariant image pixels in all locations and scales (Figure 2). Scale-space $L(x, y, \sigma)$ of image I is given by a convolution of the Gaussian kernel $G(x, y, \sigma)$ with image I . The



scale-space pyramid is constructed by continuously convolving the variable scale Gaussian $G(x, y, k\sigma)$ with I which results in the smoothed image $L(x, y, \sigma)$. The DoG image produces scale layers, and it is defined as follows:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (1)$$

Scale-space pyramids are produced in several octaves. Lowe suggested $k = 2^{1/LN}$ [20], where LN is the number of scale layers in each octave and set to three by Lowe. Image size determines the number of octaves ON. Equation 2 presents the computation of the scale coefficient of the l , the scale layer in the o th octave in the scale-space pyramid.

$$\begin{aligned} SC_{ol} &= \sigma_0 2^{(o-1) + \frac{1}{LN}} = \sigma_0 k^{LN(o-1)+l}, & (2) \\ o &= 1, 2, \dots, ON; l = 1, 2, \dots, LN; k = 2^{1/LN}, \\ \sigma_0 &= 1.6. \end{aligned}$$

The extrema in the DoG pyramid will be found by comparing each pixel to its eight neighbors at the same scale layer, nine neighbors in the scale layer above, and nine neighbors in the scale layer below (i.e., 26 in total).

2. *Discarding unstable candidate points and refining the location of points with subpixel accuracy.* Lowe fitted a 3-D quadratic surface to find the subpixel location and scale of each extremum point $X = [x, y, \sigma]^T$ as follows:

$$\hat{X} = -\frac{\partial^2 D^{-1} \partial D}{\partial X^2 \partial X}. \quad (3)$$

In Equation 3, D is the derivative of DoG at candidate points. This method also permits us to reject unstable candidate points due to low difference with the neighboring pixels. Therefore, poor contrast points (i.e., $|D(\hat{X})| < T_c$, where $T_c = 0.03$) are rejected. Then, the candidate points which are lying at the edges are also eliminated using principle curvature analysis. The principle curvatures can be calculated from the Hessian matrix, and a threshold $T_r = 10$ is considered in the eigenvalues of the Hessian matrix to reject edge points above the threshold. Details about SIFT can be found in [22].

Although the SIFT detector produces distinctive and repeatable scale invariance features in computer vision applications, it fails to extract sufficient point features for high-order transformations in multimodal retinal images [20,40]. The different nature of multimodal images causes high sensitivity in the SIFT parameters, particularly T_c , which controls the number of extracted point features. Most color retinal images, particularly pathological retinal images, have low contrast between

structural information and background pixels. However, the visibility of structural information is high in some FA sequences due to imaging after the injection of a dye. As illustrated in Figure 3 for unique T_c value, Lowe's keypoint detector fails to detect or is unable to identify well spatially distributed and adequate distinctive points for color images. In some areas, SIFT produces a lot of insignificant features. In other regions, the detected features are too sparse. Although reducing the threshold (i.e., 0) increases the number of features, they are not useful as they mainly consist of non-stable and non-uniform distributed points. Therefore, the non-uniform distribution of feature points results in non-adequate repeatable points to find correspondences. In angiographic images, SIFT detects a number of redundant features which, if all are used for matching, can lead to high computational complexity. Redundancy of features severely hampers matching performance when they are not distinctive and results in the same descriptors for close points and thus mismatches. Therefore, if many correspondences are incorrect, they will produce an incorrect transformation function. Moreover, when matched feature points are near to each other, small location errors may lead to large scale and rotation estimation errors [20].

There is no unique T_c for different pairs of retinal images. Moreover, the number of detected points in each image is highly sensitive. To the best of our knowledge, no attempt was made to determine T_c in standard SIFT to control the detection of points for RIR. Some techniques increase the number of detected feature points in images when the SIFT detector fails to extract distinctive and adequate features. Tsai et al. [25] incorporated vessel centerline points or edge points to register low-quality retinal image pairs. The approach requires enhancement methods to extract sufficient points. Jupeng et al. [24] used color information to increase the number of detected Lowe's keypoints in color retinal images. Although it increased matched feature points using color information in detection and distinctive description, it is inapplicable to multimodal images. Chen et al. [20] utilized Harris corner points. They claimed that corners are sufficient, uniformly distributed, computationally efficient, and easy to implement. However, the Harris corner detector has no control over the spatial distribution of the extracted points, and its repeatability rate decreases when there are scale changes beyond 1.5 between images or in the presence of pathological color images. Zheng et al. [41] detected Salient feature regions (SFR) in structural information, like vessels based on a saliency metric including adaptive variance and gradient field entropy. Therefore, it is robust to background changes and pathologies in the retina in contrast to point detectors and anatomical feature extractors. The

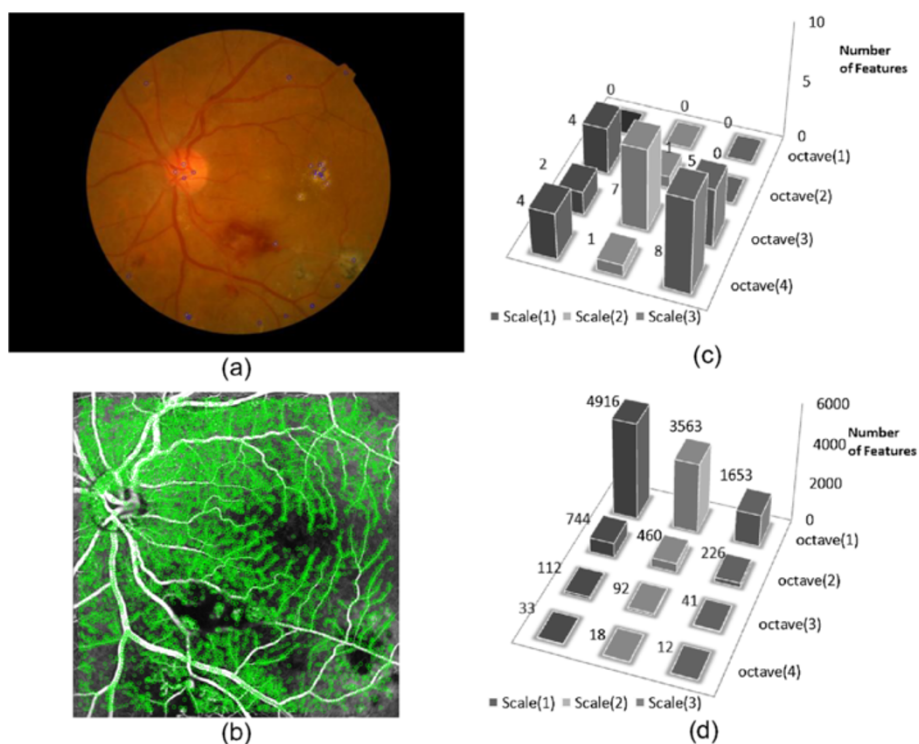


Figure 3 Non-uniform distribution of features. Original SIFT point detector in different scale layers of octaves for color and corresponding FA images for $T_c = 0.03$. It is required to set T_c manually for various retinal image pairs to achieve repeatable points. (a, b) Image-space distribution. The original SIFT fails to extract adequate features or distinctive ones in color images while detects a lot of redundant features in FA images (boundary points are masked). (c, d) Scale-space distribution. There are inadequate features in some scales which decrease the scale invariance image matching.

authors claimed that their method detects repeatable regions with least overlapping areas and thus result in a distinctive description of regions in contrast to the description of points when their surrounding areas may be overlapped. However, the saliency metric for extracting SFR does not perform on color images as well as the RF or auto-fluorescence (AF) images. This is due to the similarity of some abnormalities to vessels, and the low contrast between vessels and background of color retinal images which results in a similarity of color of vessels to color of background. Regardless of image modality, the above approaches often tend to produce misregistration results in the case of the FA (AF) with color fundus images of our dataset which have scale and content changes.

Standard SIFT does not have control over extraction of sufficient scale invariant feature points. However, reliable scale invariant point matching, particularly for multimodal images (e.g., color and the FA images), requires an adequate number of extracted features in each scale. The non-uniform distribution of points degrades the efficient number of the scale invariance features, too. Figure 3 shows a non-uniform distribution

of the standard SIFT features for multimodal retinal image pairs in different octaves and scale layers.

2.2 Problems related to feature description part

The application of the SIFT descriptor faces incorrect matches in multimodal RIR and angiographic sequences with non-linear intensity differences. The reason is that SIFT uses a gradient orientation histogram to calculate the main orientation and descriptors [22]. However, in multimodal images, the gradient orientation of corresponding points may be in opposite directions [20]. Some papers have addressed this drawback by either extracting the descriptor from the structural information (i.e., edges) of images or gradient mirroring [39].

Structural information such as blood vessels is depicted in every modality, and it remains the same for corresponding regions in multimodal images. Yuping et al. [40] and Tsai et al. [25] utilized SIFT on edge responses by only using gradient edge magnitude. This method detects several SIFT features but it requires preprocessing, including noise removal and contrast enhancement. The ED-DB-ICP addressed non-linear intensity changes by incorporating the SIFT on gradient

magnitudes and shape context (SC) to increase the distinctiveness of keypoints especially in flat areas. SC is invariant to rotation and scale and robust to linear and non-linear changes due to using edge points [42]. Rank order-matched keypoints were used in the ICP algorithm to find the parameters of the transformations. In situations where there are no reliable edges or structural content changes in the image pairs, SC fails to describe efficiently.

The SFR [41] used histograms of gradients and geometric information of gradients to describe SFRs. Contrast reversal was handled by rotating coarse-matched regions continuously to find the best rotation angle where the similarity of regions was high. Next, a clustering technique was performed to exclude fine-matched SFRs. The authors reported a better performance of their method in terms of speed in comparison to the ED-DB-ICP, but it was not invariant to non-linear intensity changes of the FA sequences or color and the FA image pairs.

Gradient mirroring [39] is also invariant to contrast reversals in multimodal images. It combines opposite gradient directions in the same bin for multimodal image pairs in the SIFT descriptor. However, it is less distinctive due to the reduced dimension of the SIFT which leads to low performance of matching. In order to achieve a more distinctive feature descriptor in multimodal images of brain registration, symmetric SIFT describes Lowe's keypoints [43]. The same idea in the PIIFD [20] describes surrounding fixed-size regions of corner points for multimodal RIR. The PIIFD is a linear combination of constrained gradient orientations from 0 to π , and its rotated version is used to address problems of opposite directions in corresponding points in multimodal images.

2.3 Problems related to SIFT in existence of repetitive patterns

Although SIFT encompasses a significant recall rate in computer vision applications, it faces wrong matches in retinal images due to the presence of repetitive patterns. Several methods, including random sample consensus (RANSAC) [44], Hough transform [28,45], and graph transform matching (GTM) [30] refine initial correspondences based on SIFT. The authors of GTM claimed that it outperforms RANSAC in discarding incorrect matches. However, the number of correct matched pairs depends on the value of K neighbors in the adjacency graph reconstruction. Moreover, when wrong matches have the same neighbors, they are excluded as correct matched pairs. Methods such as the ratio of rigid distances of any two correct matches [20] are rarely perfect in discarding false matches in retinal images where it is also possible to have mismatches with the same distance ratio.

3. Proposed matching algorithm

The proposed algorithm comprises the uniform robust scale invariant transform (UR-SIFT) feature extraction [46] to achieve higher robustness and uniformly distributed point features in retinal images. The algorithm continues with a section for finding correspondences, including computing the PIIFD descriptor, cross-matching, and outlier rejection parts. Finally, it computes the parameters of transformation using correctly matched point pairs. Figure 4 shows the flowchart of the registration framework.

3.1 UR-SIFT feature extraction

UR-SIFT extracts stable and distinctive point features that are uniformly distributed in both image and scale spaces. This results in the success of the feature-matching process. Details of the above properties are presented below:

- Stable features are repeatable and invariant for all deformations. Selecting feature points which are much higher than the contrast threshold T_c yields stable features. Here, after sorting all of the candidate points based on their $|D(X)|$ value, the range between the minimum and maximum values is partitioned into ten sections, and the candidate points that fall in the first section are discarded. Therefore, poor contrast extrema are removed automatically and without using the predefined T_c value.
- Distinctive features have a unique description. The entropy of the local region around the candidate point in the relevant Gaussian image specifies the feature distinctiveness.
- The scale-space distribution problem is addressed by assigning an appropriate number of point features to each scale layer of each scale-space octave.
- Uniform image-space distribution of extracted SIFT features is addressed using regular grid cells. In each grid cell, mean entropy of points, the number of available candidate points and mean contrast of points determine the number of required features in that cell.

Figure 5 shows the image-space and scale-space distributions of the UR-SIFT feature extraction on a color fundus photograph and corresponding angiographic images of Figure 3.

Algorithm: The UR-SIFT feature extraction algorithm

- 1: The number of extracted points N is initialized.
- 2: A scale-space pyramid is produced, and the following operations are performed for each scale layer.
 - 2.1: Extrema are detected according to the original SIFT, and poor contrast points are discarded based on the above explanation.

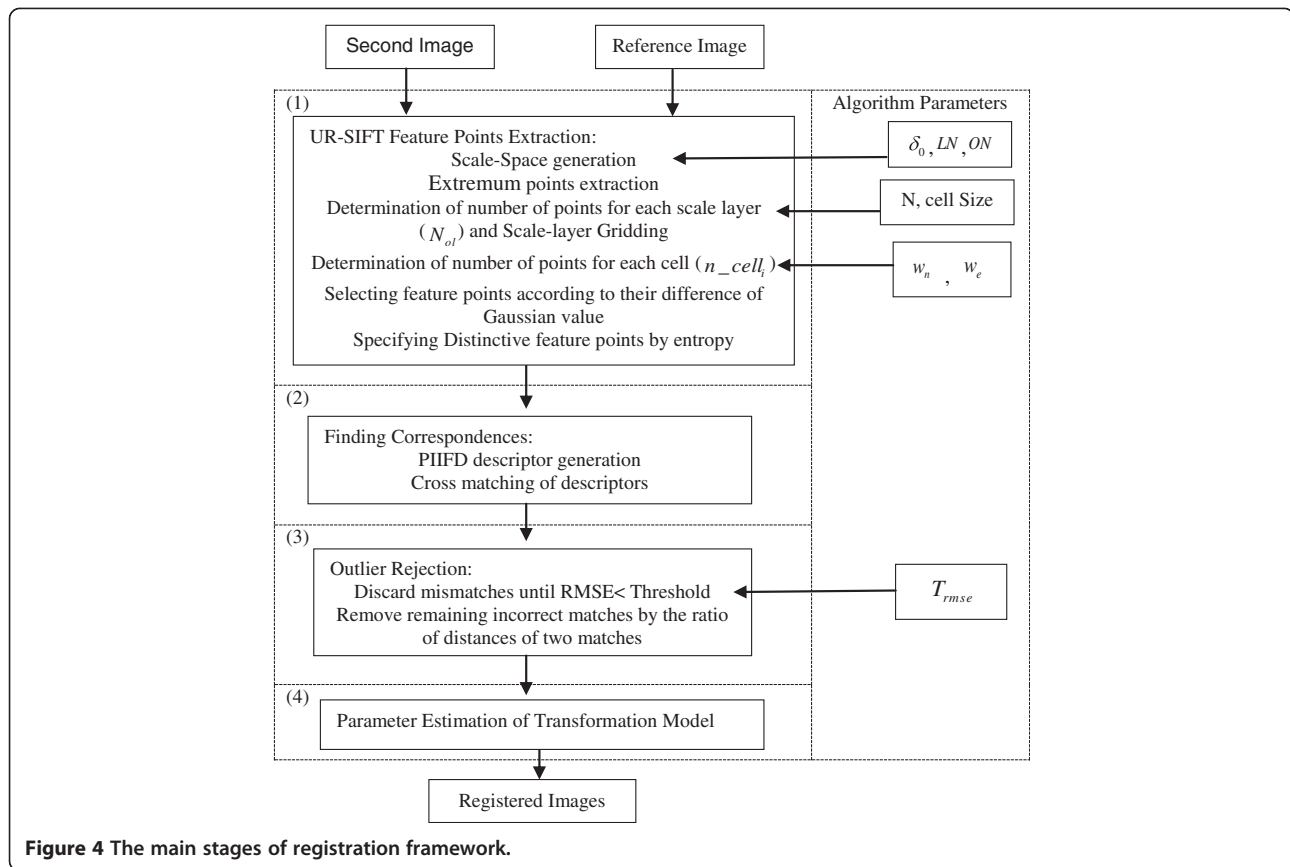


Figure 4 The main stages of registration framework.

2.2: The number of required features for each scale layer N_{ol} is given by $N_{ol} = N \cdot F_{ol}$, where F_{ol} is determined reverse to scale coefficient SC_{ol} as follows:

$$F_{ol} = \frac{SC_{11}}{SC_{ol}} f_0 \text{ and } f_0 = \frac{k^{ON \cdot LN - 1}}{\sum_{n=1}^{ON \cdot LN} k^{n-1}}. \quad (4)$$

2.3: The relevant Gaussian image of each scale-layer is partitioned into a grid, and the number of features for the i th cell of the grid is calculated as follows:

$$n_{cell_i} = N_{ol} \left[\frac{W_E E_i}{\sum_i E_i} + \frac{W_n n_i}{\sum_i n_i} + \frac{(1 - W_E - W_n) MC_i}{\sum_i MC_i} \right], \quad (5)$$

where $E_i = -\sum_j P_j \log_2 P_j$ is the entropy of the cell where P_j is probability of the j th pixel in the cell. n_i is the number of available features in the

i th cell, and MC_i is their mean contrast. W_E and W_n are entropy and feature number weight factors, respectively and are determined empirically.

2.4: For each grid cell, $3 \times n_{cell_i}$ of high-contrast features are considered. Their locations and scales with subpixel accuracy are refined. Points lying on edge points are removed, and the entropy of the surrounding area of the remaining points is computed. n_{cell_i} of features with the highest entropy value are selected in each cell. In retinal images, points with the highest entropy are those with surrounding salient structural information such as vessels. This characteristic of the UR-SIFT enables our proposed method to perform better in the presence of background pathologies in unhealthy multimodal retinal images, while vessels appear in all modalities.

For more details about the UR-SIFT, please see [46].

3.2 Finding correspondences

After extracting the UR-SIFT points, orientations are assigned to each extracted feature, and then the PIIFD

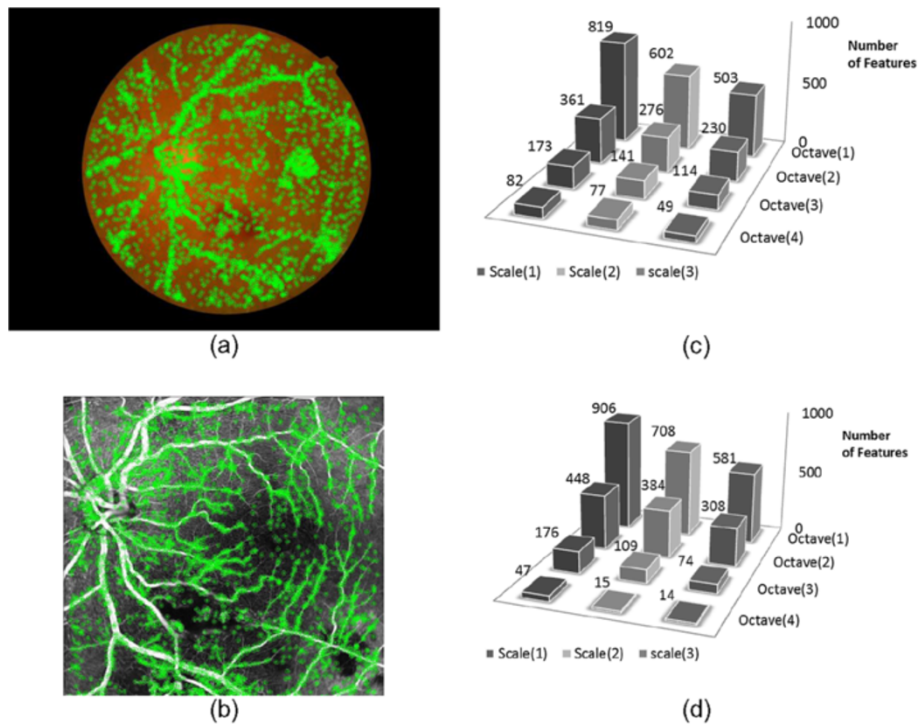


Figure 5 Distribution of features in the UR-SIFT for images of Figure 3. (a, b) Image-space distribution (boundary points are masked) and (c, d) scale-space distribution. The UR-SIFT detects areas with densely distributed blood vessels which are robust to background changes of the retina.

descriptors, which are the best descriptor for multimodal RIR [20], are generated in the neighborhood around the points in a relevant Gaussian image. The size of their neighborhood is selected by the scale of the Gaussian images where the points are detected. In the case of retinal images, computing descriptors in Gaussian images of scale space with various kernel sizes emphasizes structural features like vessels with varying widths. Small blood vessels can be described at low scale, and a connective edge map of wide blood vessels in larger scales results in smooth PIFD histograms. The minimum Euclidean distance between feature descriptors from the reference image to the second image and reverse direction is used as the matching criterion. However, there are still some mismatches which will be removed in the next section.

3.3 Outlier rejection

Although, we used high-quality and distinctive feature points to register retinal images, there are still mismatches due to the high number of repeated patterns in the retinal images. We found inliers among matched point pairs by checking them in a global transformation function between the reference and the second image. We can choose similarity, affine, polynomial transformation in the process. In our experiments on dataset, affine gives the

best results. Affine transformation is applied to all matched pairs. In an iterative process, each matched pair is checked in the global model, and the pairs which have the highest geometric error between the reference and the second image are removed one by one until achieving the root mean square error (RMSE) $< T_{RMSE}$ [6].

Let (x_r, y_r) and (x_s, y_s) be two corresponding feature points in the reference and second images. Parameters of transformation function A are computed by $(x_r, y_r) = A(x_s, y_s)$. All points of the second image are transformed into the coordinates of the points in the reference image, and the root of the average of the set of differences between the reference points and transformed points (x_r, y_r) is considered as RMSE:

$$RMSE = \sqrt{\frac{\sum_{k=1}^n (x_r^k - x_t^k)^2 + (y_r^k - y_t^k)^2}{n}} \quad (6)$$

After parameter estimation, the RMSE is computed, and if it is larger than the threshold value, the point with the highest error is removed. The process continues until the RMSE exceeds the threshold. Finally, corresponding points which have their error three times larger than RMSE are removed.

Although the remaining compatible matched pairs of this method are correct and adequate to find the parameters of transformation, they may lead to the elimination of some correct matches in the presence of a small threshold value for some challenging image pairs. On the other hand, for such image pairs, methods including the ratio of rigid distances of any two correct matches and the comparison of main orientations of matched pairs [20] produce some incorrect matches or may fail. To avoid drawbacks, a combination of two methods is considered. First, the threshold is considered a big value. All correspondences are checked in affine transformation, and the most probable incorrect matches are discarded up to the $RMSE < 6$ pixels is achieved. Second, any remaining features will be eliminated if the distance between them and other features in the first image is not equal to the distance of corresponding features in the second image.

3.4 Parameter estimation of transformation function

Various types of transformation functions like similarity, affine and quadratic models have been applied in the literature to register retinal images. Model types of some registration frameworks [20,41] are dependent on the number of matched points. Chen et al. in [20] first used an affine model to transform the second image into a reference image coordinate, and then they repeated the matching process again on the transformed image. This

resulted in an increased number of matched points. However, incorrect matched points in the first step led to an incorrect transformed image, unsuccessful next process, and then failure of registration. For image pairs with a wide range of overlaps, second-order polynomial transformation had less alignment average error than the affine model due to more flexible mapping in the presence of non-linear retinal curvature [13,18]. In most cases of our dataset, there is an approximately 40% overlap between image pairs while the color images cover the fundus completely and the FA images zoom on the pathologic region. There are some cases which both color and FA images cover the fundus completely, too. Moreover, the number of matches by the proposed algorithm is sufficient and the matches are distributed well in the surface of the retina. Hence, proposed method is able to constrain the high-order model needed for accurate registration. Given the above explanation, a second-order polynomial is applied. The parameters of the transformation function are estimated by the least square method [47]. Figure 6 shows the feature matching and registration results of images in Figure 3.

4. Results and discussion

We evaluated the proposed approach on some pairs of multimodal retinal images of 120 patients where state-of-the-art approaches failed to register due to the above

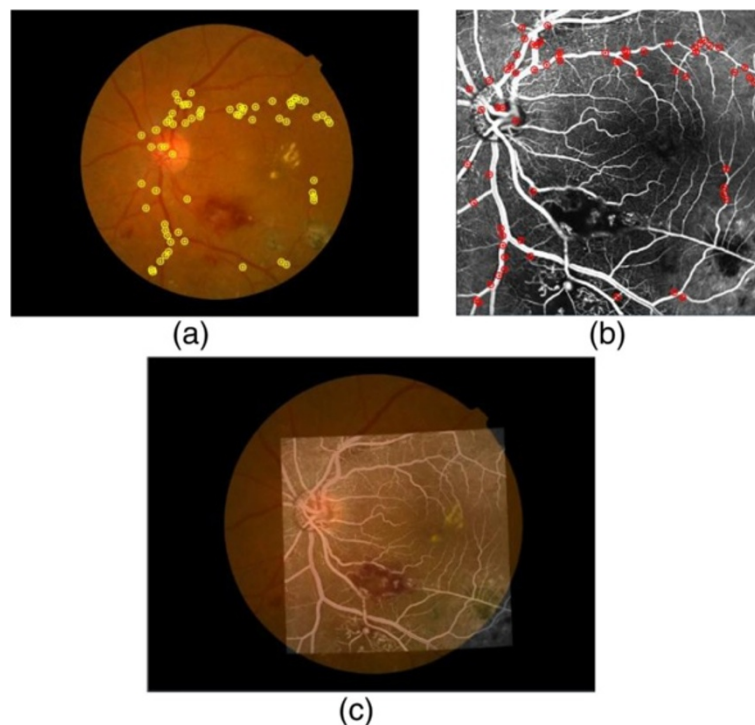


Figure 6 Feature matching and registration results. (a, b) Results of the proposed matching algorithm for images in Figure 3 with 4,000 feature points and $T_{RMSE} = 6$ and (c) the registration result, $RMSE = 1.8$.

explanation. However, the approach accuracy is reported in a comparison of the Harris-PIIFD on some multimodal retinal image pairs. The proposed algorithm was implemented in the MATLAB. ED-DB-ICP and Harris-PIIFD algorithms were downloaded as a binary executable program in C++ and the MATLAB source code program, respectively. All parameter settings of the UR-SIFT are presented in Section 4.4. The experiments were performed on a personal computer with Intel® Core™ (Intel, Santa Clara, CA, USA) i5 2.53 GHz and 4 GB of RAM.

4.1 Input image pairs

There is no available benchmark for the RIR dataset and ground truth to evaluate the registration results of different techniques. Our dataset consists of two datasets from different retinal image modalities such as color fundus images and the corresponding RF, AF images, and FA sequences to evaluate the proposed method. Each modality exhibits special retinal abnormalities or structures. The first dataset was collected from the Internet, and the second was provided by the Shahid Labbafi Nedjad Hospital and the Negah Ophthalmology Imaging Center of Iran. In the first dataset, there are some multimodal image pairs from previous related papers [20] and retinal online atlases [48]. They consist of 20 image pairs of color (RF) and corresponding FA images. The images have a resolution in the range from 300×250 to 650×500 pixels. A wide variety of diseases of the macula, vessels, and optic disc appear in the images. The second dataset comprises two modes. In the first mode, the camera views 30° of retina area, while in the second mode, 50° . An ophthalmologist requires an eye photographer to capture 30° field of view of the retinal images which include pathological regions of FA, RF, and AF. Zooming in pathological areas makes 30° images to depict the information from the retina with better resolution. The second dataset includes 60 image pairs of the first mode color, and corresponding FA sequences and AF images are captured by a Topcon 50 DX fundus camera (TRC-50IA, Tokyo, Japan) and HRA2 camera (Heidelberg Engineering, Heidelberg, Germany), respectively. Color images have a resolution of $2,048 \times 1,336$ pixels and cover the fundus completely. FA sequences and AF images are 768×822 pixels and are zoomed in the pathological region. A sample of such a dataset is shown in Figure 3. There are also 40 cases in which color and the FA image pairs cover the fundus completely. Color images are captured by a Topcon TRC NW200. The corresponding FA sequences are captured by a HRA2 camera (Heidelberg Engineering, Heidelberg, Germany). Color images have a resolution of $2,048 \times 1,536$ pixels. FA sequences have 768×822 pixels. A wide variety of symptoms of retinal diseases, such as hemorrhage, atrophy, drusen, exudates, tumors, vascular

diseases, and cataracts are considered in both datasets. FA sequences were obtained a few seconds after the injection of dye, and special phases of the FA images which exhibited pathologies were selected by an ophthalmologist for registration with color images. There are also five-color scanning laser ophthalmoscopy (SLO) image pairs. Color images are $3,888 \times 768$ pixels, and SLO images are 768×768 pixels. The registration of SLO and color images would help an analysis of the optic disc in the early detection of glaucoma [12].

The minimum overlapping area in the datasets is around 40% for multimodal images. In the small overlapping areas, if the amount of common structures between image pairs is sufficient, the registration will be successful.

4.2 Evaluation criteria

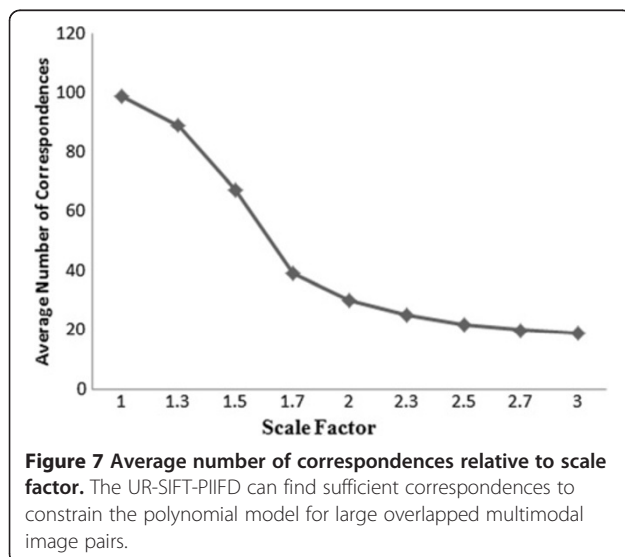
An evaluation of the overall performance of RIR can be carried out subjectively and objectively. An observer identifies subjectively the success or failure of the registration result using overlaid image pairs. However, visual evaluation needs an expert observer and is not adequate for comparing the performance of different techniques on a large dataset. Therefore, objective measurement is necessary to evaluate the registration result. The performance of image registration methods can be measured by the centerline error measure [18,26]. For each point on the vessel centerline, the distance to the nearest centerline point is found in the second image, and the median of these distances is considered a registration error. However, centerlines can be hard to detect for low-quality or unhealthy images. In practice, small difference will exist between coordinates of matched points. Therefore, accuracy of registration can be evaluated by the error in point placement as the RMSE between the transformed points and the points of reference image [5,24]. We select 20 matched points manually using MATLAB R2008a and generate ground truth to evaluate the proposed approach. Points have to be distributed uniformly with an accurate localization.

Successful registration is determined with regard to the RMSE value. A RMSE below 5 pixels is acceptable for clinical purposes [5]. The success rate [26] is the ratio of the number of image pairs with successful registration to the number of all image pairs. The success rate of the Harris-PIIFD is low for the registration of color images with other modalities. We compare the average RMSE for images that our proposed algorithm and the Harris-PIIFD register them successfully. For accurate and successful registration, we consider the $RMSE < 5$ pixels. The median and maximal distances between the reference points and transformed points of the second images are reported, too. A significant error (i.e., mean distance > 10 pixels over all matched points for each image pair) results in a registration failure.

4.3 Special test results

The significant innovation of this paper is the registration of color and the FA retinal image pairs which have considerable intensity, scale, and structural changes. This part evaluates the robustness of the proposed technique to scale insensitivity. Moreover, it considers registration performance on different phases of a complete FA sequence.

1. *Scale invariance test.* The largest scaling factor is 2.5 in the test dataset. Scale differences for multimodal image pairs are due to using different instruments to capture multimodal images, the sitting distance of the patient to the imaging instrument, and the required resolutions of each modality to show the structures of retinal images. UR-SIFT is able to find matched points in image pairs with a scaling factor of 5 [46]. However, most color and corresponding FA images with a 30° view have a scaling factor of about 1.5. Color and FA images which cover the fundus completely have a scaling factor of about 2.5. We automatically rescaled 30 selected image pairs with a scaling factor from 1 to 3 and applied the proposed algorithm on all the image pairs. The average number of matches across a range of scale changes can be seen in Figure 7. This experiment indicates that the UR-SIFT-PIIFD can find adequate correspondences to compute the parameters of function transformation. For all image pairs the registration results were successful. In the Harris-PIIFD, the author utilized Harris corners to provide locations for computing the PIIFD, whereas the SIFT features fail to detect sufficient salient points in multimodal retinal images. Our experiments on test dataset showed the low

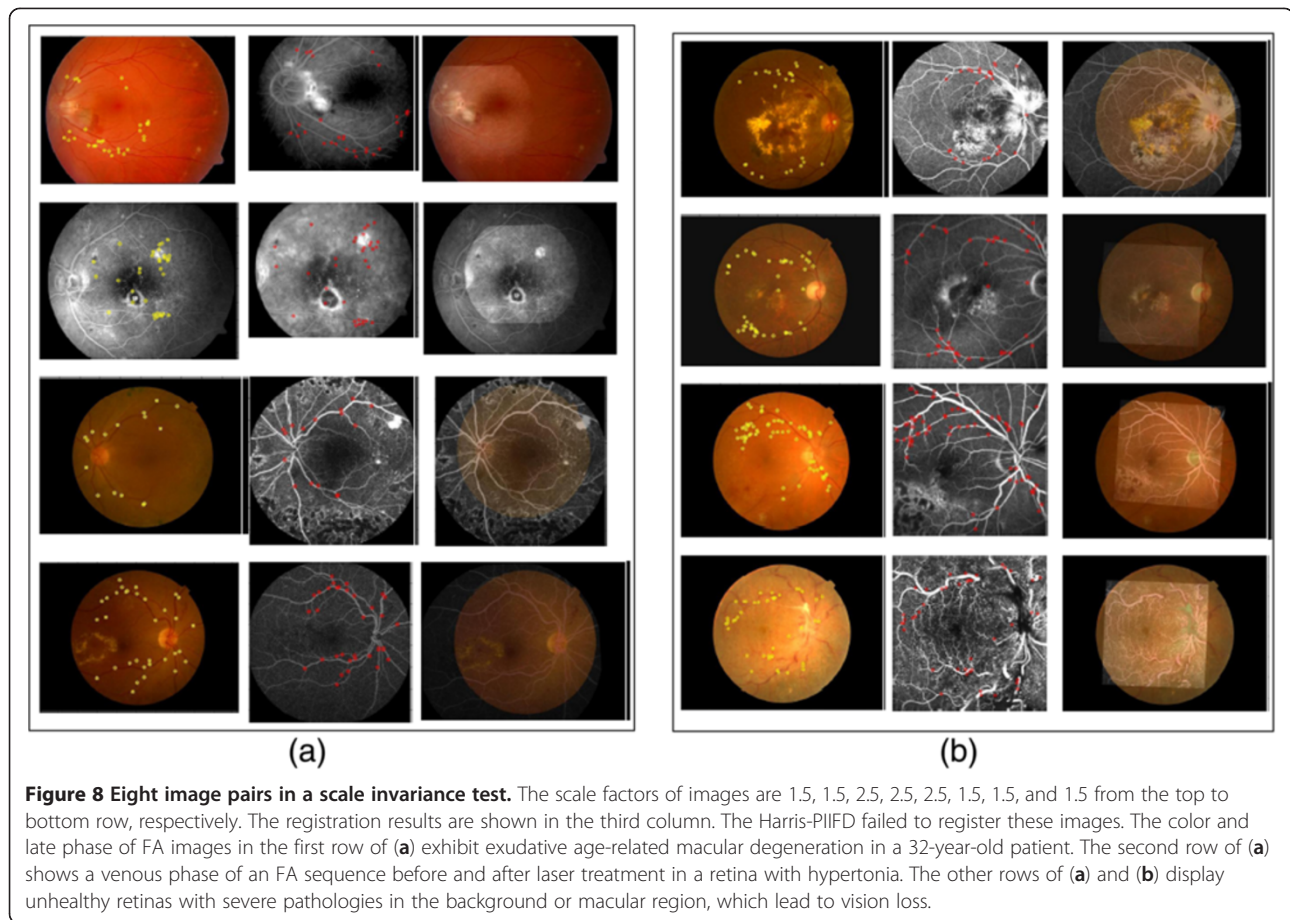


success rate of the Harris-PIIFD for registration of 30° and 50° images. It is applicable for retinal images which do not have much scale change. However, the Harris-PIIFD requires image resizing and smoothing to work properly for all multimodal images with and without scale changes. Selecting the SIFT features using the UR-SIFT method with controllability outperforms Harris corners in images both with and without scale changes which consist of low-content contrast retinal images. In image pairs with small changes of scale, the proposed algorithm increases the number of corresponding points with uniform distribution which lead to low RMSE errors. Figure 8 shows eight examples of registration results of our proposed method.

2. *The complete FA sequence test.* To evaluate the performance of the UR-SIFT-PIIFD for the registration of color images and the relevant complete FA sequence, we used five special patients. The Harris-PIIFD responds to the registration of the FA sequences with color images when there are color images with normal contrast or there are no large structural changes. Therefore, to register any FA phase to the color image using existing approaches, it is necessary to find an FA image whose registration with a color image is successful. This demands prior knowledge to find an intermediate FA phase whose registration to the color image and any FA image is successful. Unlike the Harris-PIIFD, the UR-SIFT-PIIFD algorithm provides a fully automatic registration system for registration of any phase of the FA sequence with a color image. The UR-SIFT-PIIFD fully registered four out of five patients. However, the Harris-PIIFD only registered two FA phases in the whole FA sequences.

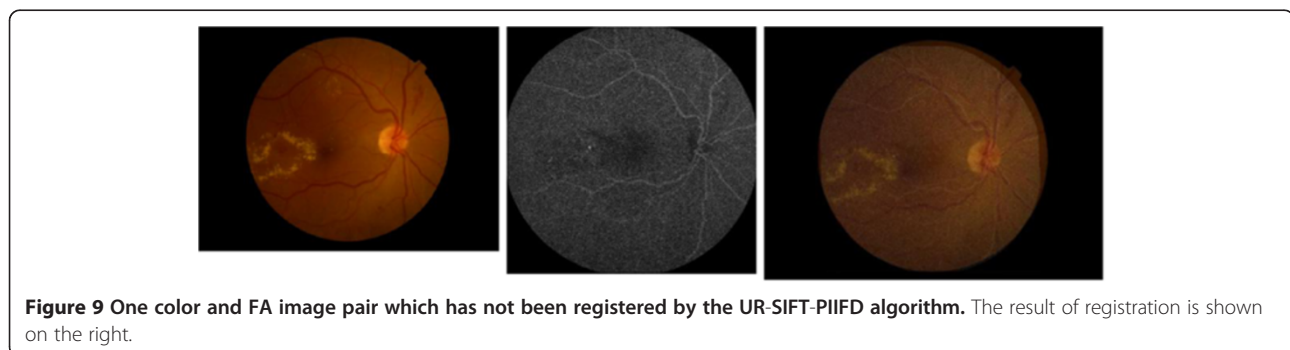
4.4 Parameters

The optimum values of parameters are determined in an experimental analysis of 120 multimodal image pairs with various content, scale, and appearance changes. The scale space has four octaves, each one of which includes three scale layers, and the optimum value of δ_0 is 1.6 as standard SIFT [22]. These values lead to maximum repeatable and increased corresponding feature points [22]. The value assigned to δ_0 was changed from 1 to 2 in 0.1 steps, and the maximum number of true correspondences was achieved when $\delta_0 = 1.6$. It is obvious that the Gaussian retinal images of octaves that are higher than fourth octave are highly blurred and non-informative to be described.



The number of extracted feature points in the UR-SIFT was determined based on 0.4% of image size. The 2,500 feature points for the first dataset and 4,000 feature points for the second dataset are adequate to ensure proper working of the proposed method. However, in some cases of challenging image pairs, 4,500 feature points result in an expected working of the proposed method. The size of each cell in the grid is 200×200 pixels. The optimum values of W_E and W_n parameters are

determined empirically. The values of the parameters were changed in a range from 0 to 1 with a step length of 0.1. The maximum values of the average number of correct correspondences with well distribution were achieved by using $W_E = 0.2$ and $W_n = 0.5$. This means that point features are preserved between color and angiographic images on the vessels where the DoG images have maximum responses (i.e., high contrast), and entropy is also appropriate for measuring structural saliency. There-



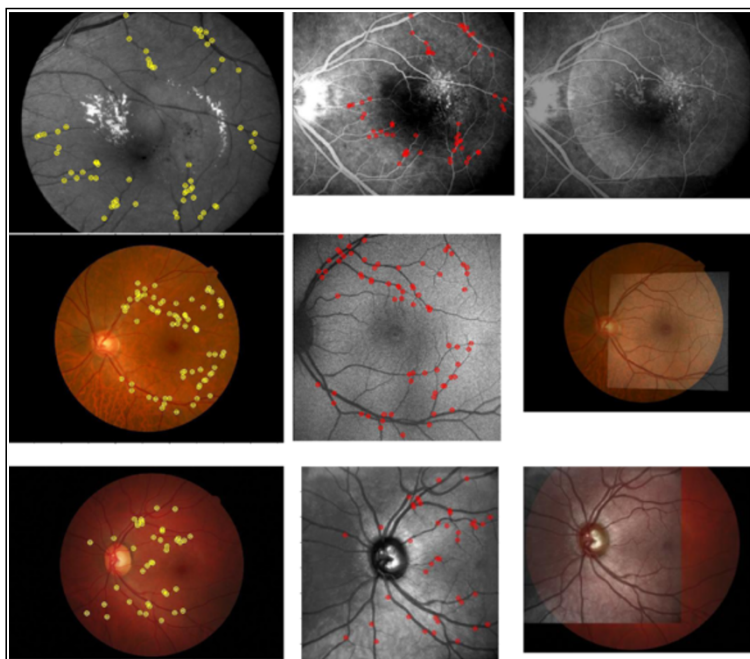


Figure 10 Results for multimodal image pairs: RF-FA, AF-color, and color-SLO. The first row exhibits an RF and venous phase of its FA sequence in a retina with dilated blood vessels. The scaling factors of the image pairs are 1.5, 1.5, and 2.5, respectively. The Harris-PIIFD failed to register these images.

fore, a feature subset selection combining entropy and mean contrast outperforms the individual saliency measure acting alone in terms of the overall success rate.

The parameters to compute the PIIFD descriptors were determined as suggested by Chen et al. [20]. The window for determining the dominant orientation of the feature points is 13×13 pixels.

4.5 Overall performance of registration and comparative test results

Six out of 60 30° image pair registrations and 9 out of 40 50° image pair registrations failed in our approach (Figure 9). However, this works out as 52 for 30° images and 29 for 50° image pairs in the Harris-PIIFD algorithms. It is worth pointing out that the Harris-PIIFD resized the original images and then filtered the resized images to work properly. Therefore, the resolution of the registration result is low. Consequently, losing some pixels leads to an imprecise clinical analysis. The proposed algorithm fails when there are very low-quality image pairs and much content changes between the color and different frames of the angiographic retinal images. Figure 9 shows the registration result of an image pair. The angiographic retinal image is quite noisy and has low content due to the lens problems of the patient's eyes. Existing algorithms are unable to register these images. In the feature matching process, dissimilar PIIFD descriptors may be produced for the corresponding

extracted feature points due to the appearance of pathology in only one of the modalities. Hence, the presence of lesions in the neighborhood of feature points in one of the modalities will adversely affect the results of the PIIFD descriptor which is based on gradient orientation information [49]. It is also possible to have similar descriptors for non-corresponding extracted points in retinal images due to the repetitive patterns or low content of poor-quality images. These result in few correct correspondences (e.g., 3) even when informative selected features exist. In this situation, given the recent success of the generalized dual-bootstrap iterative closest point (GDB-ICP) [26] which is capable of successfully aligning a pair of images starting from just one correct

Table 1 Overall performance of registration for the second dataset

	ED-DB-ICP		Harris-PIIFD		UR-SIFT-PIIFD	
	30°	50°	30°	50°	30°	50°
Field of view	30°	50°	30°	50°	30°	50°
Success rate (%) of color-FA registration	0	0	13.33	27.5	90	77.5
Accuracy (pixel)						
Average RMSE					2.9	2.7
Average minimum distance					0.7	0.5
Average maximum distance					8	6.2
Average median distance					3.03	2.3

The ED-DB-ICP failed for all image pairs.

correspondence [50], the UR-SIFT-PIIFD can provide the initial matches for the GDB-ICP algorithm and may register images successfully.

We also obtain satisfactory results of the proposed integrated approach in other modalities (Figure 10). The Harris-PIIFD fails to register these images due to scale changes and lack of enough feature points of color images. The ED-DB-ICP failed to register all the image pairs of the second dataset. Because magnitude (i.e., the edge-driven image of color images) is not informative, the feature point extractor algorithm fails. The SIFT failed to register the first two image pairs. However, correspondences in the third image pairs have no uniform distribution such as matching points of the UR-SIFT-PIIFD. Hence, non-uniform distribution of points cannot constrain the polynomial function transformation for accurate and successful registration.

Table 1 indicates the overall performance of the proposed method on the second dataset in comparison with the Harris-PIIFD. The success rate and average RMSE values of the Harris-PIIFD for the first dataset were 35% and 2.1, respectively, while they were 60% and 1.8, respectively, for the UR-SIFT-PIIFD. The average RMSE value of the whole image pairs of the second dataset, where the registration was successful, was 2.83.

The average number of initial correspondences using the proposed algorithm was 64 for the first dataset, 58 for the 30° images, and 25 for the 50° images of the second dataset. The average number of final correspondences after outlier discarding was 61 for the first dataset and 54 for the 30° and 19 for the 50° images of the second dataset. The average number of correct matched points for the image pairs, where both the Harris-PIIFD and the proposed algorithm succeeded, was 18 and 79, respectively. The number of true correspondences increased in comparison with the Harris-PIIFD in the case of the image pairs without large-scale variations. These resulted in increasing registration accuracy particularly when the correspondences were uniformly distributed [4]. Moreover, the number of matched points was adequate to compute quadratic model parameters.

5. Conclusions

In this paper, we have improved the registration of retinal image pairs particularly that of multimodal angiographic and color images with content and scale changes, as well as in the presence of low-quality color images and noisy background of the FA images. Although the Harris-PIIFD demonstrated a high success rate with multimodal image pairs, its weakness was in the selection of scale invariance and distinctive, stable points to register the mentioned image pairs. The base of improvement was that of using high-quality UR-SIFT

features in the uniform distribution of both the scale and image spaces to compute the PIIFD descriptor. In contrast to the Harris-PIIFD, the UR-SIFT-PIIFD does not resize the original images. The experimental results on a variety of multimodal retinal image pairs, which the Harris-PIIFD and the ED-DB-ICP failed to register, showed the advantages of the improvement. The main direction for future research and improvements in retinal image registration is to provide a more powerful point detector to deal with very noisy and low-quality FA and blurred color images, particularly for the 50° images in some eye diseases like cataracts. Moreover, they are the requirements to generate a more powerful local feature descriptor for multimodal images with high changes of content and a more efficient outlier rejection strategy with a local model fit to deal with the spherical geometry of the eyeball. The UR-SIFT-PIIFD can provide initial matches in the registration of multimodal 200° color and angiographic image pairs. In our future research, we will extend the UR-SIFT to apply to 3-D medical image registration, too.

Abbreviations

AB: Area-based; AF: Auto-fluorescence; ECC: Entropy correlation coefficient; ED-DB-ICP: Edge-driven dual bootstrap iterative closest point; FA: Fluorescein angiography; FB: Feature-based; GDB-ICP: Generalized dual-bootstrap iterative closest point; GTM: Graph transform matching; IR: Image registration; MI: Mutual information; PIIFD: Partial intensity invariant feature descriptor; RANSAC: Random sample consensus; RF: Red-free; RIR: Retinal image registration; RMSE: Root mean square error; SFR: Salient feature regions; SIFT: Scale invariant feature transform; SLO: Scanning laser ophthalmoscopy; UR-SIFT: Uniform robust scale invariant transform.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

The authors would like to thank Dr. Jian Chen and Dr. Chia-Ling Tsai for providing source codes for the Harris-PIIFD algorithm and an executable program of the ED-DB-ICP algorithm. The Authors would like to thank Dr. Jan Odstrcilik for providing SLO fundus retinal image pairs, and the Iranian Negah Ophthalmology Center and the Shahid Labbafi Nedjad Educational Hospital for providing the second dataset of the retinal image pairs. The Authors are also thankful to Mrs. Zohreh Salimi for performing Fluorescein Angiography in shahid Labbafi Nedjad imaging center.

Author details

¹Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran 4933-14155, Iran. ²Department of Computer Engineering, Kharazmi University, Tehran 37551-31979, Iran. ³Faculty of Geodesy and Geomatics Engineering, K.N. Toosi University of Technology, Tehran 19967-15433, Iran. ⁴Biomedical Signal and Image Processing Laboratory, School of Electrical Engineering, Sharif University of Technology, Tehran 14588-89694, Iran.

Received: 31 May 2012 Accepted: 3 April 2013

Published: 28 April 2013

References

1. S Dithmar, FG Holz, *Fluorescence Angiography in Ophthalmology* (Springer Medizin Verlag, Hiedelberg, 2008)
2. National Eye Institute, Eye health information. (2008). <http://www.nei.nih.gov/health/>. Accessed 27 May 2012

3. B Zitova, J Flusser, Image registration methods: a survey. *Image. Vis. Comput.* **21**, 977–1000 (2003). <http://www.sciencedirect.com/science/article/pii/S0262885603001379>.
4. AA Goshtasby, *2-D and 3-D Image Registration for Medical, Remote Sensing, and Industrial Applications* (Wiley, Hoboken, 2005)
5. GK Matsopoulos, MPA Asvestas, NA Mouravliansky, KK Delibasis, Multimodal registration of retinal images using self organizing maps. *IEEE Trans. on Med. Imaging* **12**, 1557–1563 (2004). doi:10.1109/TMI.2004.836547
6. F Laliberte, L Gagnon, Y Sheng, Registration and fusion of retinal images—an evaluation study. *IEEE Trans. Med. Imag.* **22**, 661–673 (2003). doi:10.1109/TMI.2003.812263
7. Z Ghassabi, J Shanbehzadeh, A study of retinal image registration and mosaicking, in *Proceedings of International Conference on Information and Computer Applications* (Dubai, 18–20 March 2011)
8. CL Tsai, B Madore, M Leotta, M Sofka, G Yang, A Majerovics, HL Tanenbaum, CV Stewart, B Roysam, Automated retinal image analysis over the internet. *IEEE Trans. Inform. Tech. Biomed* **12**, 480–487 (2008). doi:10.1109/TITB.2007.908790
9. C Tae, Eun, I Cohen, 2-D registration and 3-D shape inference of the retinal fundus from fluorescein images. *Med. Image Anal.* **12**, 1174–1190 (2008). doi:10.1016/j.media.2007.10.002
10. GK Matsopoulos, NA Mouravliansky, KK Delibasis, KS Nikita, Automatic retinal image registration scheme using global optimization techniques. *IEEE Trans. Inform. Tech. Biomed* **3**, 47–60 (1999). doi:10.1109/4233.748975
11. N Ritter, R Owens, J Cooper, R Eikelboom, P van Saarloos, Registration of stereo and temporal images of the retina. *IEEE Trans. Med. Imag.* **18**, 404–418 (1999). doi:10.1109/42.774168
12. L Kubacka, J Jan, Registration of bimodal retinal images - improving modifications, in *Proceeding of 26th IEEE Symposium on Engineering in Medicine and Biology Society* (San Francisco, 2004), pp. 1–5
13. T Chanwimaluang, G Fan, SR Fransen, Hybrid retinal image registration. *IEEE Trans. Info. Tech. Biomed.* **10**, 129–142 (2006). doi:10.1109/TITB.2005.856859
14. W Wei, C Houjin, L Jupeng, Y Jiangbo, A registration method of fundus images based on edge detection and phase-correlation, in *Proceeding of the First IEEE Symposium on Innovative Computing, Information Control* (Beijing, 30 August–1 September 2006)
15. F Zana, JC Klein, A registration algorithm of eye fundus images using a Bayesian Hough transform, in *Proceedings of the Seventh International Conference of Image Processing and Its Applications* (Manchester, July 1999), pp. 13–15
16. LG Brown, A survey of image registration techniques. *ACM Comput. Surv* **24**, 325–376 (1992). doi:10.1145/146370.146374
17. J Yu, C Hung, BN Liou, Fast algorithm for digital retinal image alignment, in *Proceedings of IEEE International Conference on Engineering in Medicine and Biology Society* (Seattle, November 1989), pp. 9–12
18. A Can, CV Stewart, B Roysam, HL Tanenbaum, A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. *IEEE Trans. Pattern Anal. and Mach. Intell.* **24**, 347–364 (2002). doi:10.1109/34.990136
19. C-L Tsai, CV Stewart, HL Tanenbaum, B Roysam, Model-based method for improving the accuracy and repeatability of estimating vascular bifurcations and crossovers from retinal fundus images. *IEEE Trans. Info. Tech. in Biomed* **8**, 122–130 (2004). doi:10.1109/TITB.2004.826733
20. J Chen, J Tian, N Lee, J Zheng, T Smith, F Laine, A partial intensity invariant feature descriptor for multimodal retinal image registration. *IEEE Trans. Biomed. Eng* **57**, 1707–1718 (2010). doi:10.1109/TBME.2010.2042169
21. C Schmid, R Mohr, C Bauckage, Evaluation of interest point detectors. *Int. J. Comput. Vis.* **37**, 151–172 (2000). doi:10.1023/A:1008199403446
22. DG Lowe, Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2**, 91–110 (2004). doi:10.1023/B:VISI.0000029664.99615.94
23. AP Witkin, Scale-space filtering, in *Proceedings of the 8th International Conference on Artificial Intelligence* (Karlsruhe, August 1983), pp. 8–12
24. L Jupeng, C Houjin, Y Chang, Z Xinyuan, A robust feature-based method for mosaic of the curved human color retinal images, in *Proceedings of Biomedical Engineering and Informatics* (Sanya, Hainan, May 2008), pp. 27–30
25. CL Tsai, CY Li, G Yang, KS Lin, The edge driven dual-bootstrap iterative closest point algorithm for registration of multimodal fluorescein angiogram sequences *IEEE Trans. Med. Imag* **29**, 636–649 (2010). doi:10.1109/TMI.2009.2030324.
26. CV Stewart, CL Tsai, B Roysam, The dual-bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE Trans. Med. Imag.* **22**, 1379–1394 (2003). doi:10.1109/TMI.2003.819276
27. C Li, Z Xiao-Long, Retinal image registration using bifurcation structures, in *Proceedings of the 18th IEEE International Conference of Image Processing* (Brussels, 2011), pp. 11–14
28. F Zana, JC Klein, A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform. *IEEE Trans. Med. Imag.* **18**, 419–428 (1999). doi:10.1109/42.774169
29. L Gang, CV Stewart, B Roysam, K Fritzsche, Y Gehua, HL Tanenbaum, Predictive scheduling algorithms for real-time feature extraction and spatial referencing: application to retinal image sequences. *IEEE Trans. Biomed. Eng.* **51**, 115–125 (2004). doi:10.1109/TBME.2003.820332
30. W Aguilar, Y Frauel, F Escolano, ME Martinez-Perez, A Espinosa-Romero, MA Lozano, A robust graph transformation matching for non-rigid registration. *J. Image and Vis. Comput.* **27**, 897–910 (2008). doi:10.1016/j.imavis.2008.05.004
31. WE Hart, MH Goldbaum, Registering retinal images using automatically selected control point pairs, in *Proceedings of IEEE International Conference of Image Processing* (Austin, November 1994), pp. 13–16
32. VE Markaki, PA Asvestas, GK Matsopoulos, Application of Kohonen network for automatic point correspondence in 2D medical images. *Comput. Biol. Med.* **39**, 630–645 (2009)
33. N Ryan, C Heneghan, P de Chazal, Registration of digital retinal images using landmark correspondence by expectation maximization. *Image Vis. Comput.* **22**, 883–898 (2004). doi:10.1016/j.imavis.2004.04.004
34. E de Ves, RC Wilson, ER Hancock, Matching blood vessel patterns with the generalised EM algorithm, in *Proceedings of the 14th International Conference on Pattern Recognition* (Brisbane, August 1998), pp. 16–20
35. K Mikolajczyk, C Schmid, A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**, 1615–1630 (2005). doi:10.1109/TPAMI.2005.188
36. D Lowe, *SIFT demo program, Version 4*, 2005. <http://www.cs.ubc.ca/spider/lowe/keypoints/>. Accessed 27 May 2012
37. W LiFang Wei, H LinLin, P Lin, Y Lun, The retinal image registration based on scale invariant feature, in *Proceedings of IEEE 3rd International Conference on Biomedical Engineering and Informatics* (Yantai, October 2010), pp. 16–18
38. W LiFang Wei, H LinLin, P Lin, Y Lun, The retinal image mosaic based on invariant feature and hierarchical transformation models, in *Proceedings of IEEE 2nd International Conference on Image and Signal Processing* (Tianjin, October 2009), pp. 17–19
39. A Kelman, M Sofka, CV Stewart, Keypoint descriptors for matching across multiple image modalities and nonlinear intensity variations, in *Proceedings of IEEE Symposium on Computer Vision and Pattern Recognition* (Minneapolis, June 2007), pp. 17–22
40. L Yuping, G Medioni, Retinal image registration from 2D to 3D, in *Proceedings of IEEE Symposium on Computer Vision and Pattern Recognition* (Anchorage, 23–28 June 2008)
41. J Zheng, J Tian, K Deng, X Dai, X Zhang, M Xu, Salient feature region: a new method for retinal image registration. *IEEE Trans. Inform. Tech. Biomed.* **15**, 221–232 (2011). doi:10.1109/TITB.2010.2091145
42. CL Tsai, S-T Huang, K-S Lin, S-J Chen, Robust pairwise registration for images of indocyanine-green angiographic sequences, in *Proceedings of IEEE 10th Symposium on Pervasive Systems, Algorithms and Networks* (Kaohsiung, 14–16 December 2009)
43. J Chen, J Tian, Real-time multi-modal rigid registration based on a novel symmetric-SIFT descriptor. *Progr. Nat. Sci.* **19**, 643–651 (2009). doi:10.1016/j.pnsc.2008.06.029
44. MA Fischler, RC Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Comm. ACM*, 381–395 (1981). doi:10.1145/358669.358692
45. J Park, JM Keller, PD Gader, RA Schuchard, Hough based registration of retinal images, in *Proceedings of IEEE International Conference on Systems, Man and Cybernetics* (San Diego, 11–14 October 1998)
46. A Sedaghat, M Mokhtarzadee, H Ebadi, Uniform robust scale-invariant feature matching for optical remote sensing images. *IEEE trans. Geosci. Rem. Sens.* **49**, 4516–4527 (2011). doi:10.1109/TGRS.2011.2144607
47. C Stewart, Robust parameter estimation in computer vision. *SIAM Rev.* **41**, 513–537 (1999). doi:10.1137/S0036144598345802

48. ZEISS, *Retinal Atlas*. http://www.zeiss.de/czj/de/op/en/zeiss/index_frames.html, Accessed 12 May 2012
49. YB Bathina, NV Kartheek, J Sivaswamy, Robust matching of multimodal retinal images using a Radon transform based local descriptor, in *Proceedings of the 1st ACM International health informatics Symposium (IHI)* (Arlington, 11–12 November 2010)
50. A Kelman, M Sofka, CV Stewart, Keypoint descriptors for matching across multiple image modalities and non-linear intensity variations, in *Proceedings of the IEEE Computer Society Workshop on Image Registration and Fusion (in conjunction with IEEE CVPR)* (Minneapolis, 18–23 June 2007)

doi:10.1186/1687-5281-2013-25

Cite this article as: Ghassabi et al.: An efficient approach for robust multimodal retinal image registration based on UR-SIFT features and PIIFD descriptors. *EURASIP Journal on Image and Video Processing* 2013 2013:25.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
