Digital Fundus Image Stitching for Generation of Retinal Mosaicto Help Ophthalmologists

M. Usman Akbar, M. Usman Akram, Adeel M. Syed, and Joddat Fatima

***Abstract*—These days medical Imaging is a very important field and is helping doctors in various ways to diagnose diseases and treat them. Fundus image is used by doctors has a limited field of view and to overcome this problem a new technique is presented in this paper to stitch the images together to generate a mosaic image. Proposed methodology uses Weber local descriptor for feature extraction and then a new technique is presented for seamless stitching. A locally gathered database is used and results are compared with other state of the art techniques. Significance of proposed system can be seen from the results generated.**

***Keywords*—Fundus Image, Mosaic, Weber Local Descriptor, Homography, Seamless Blending, Image Stitching.**

I. Introduction

In the start of 19th century Digital Image Processing was considered part of print industry only but with the evolution of the computer world it took many domains to their peak. In early 1970's DIP was introduced in the field of medicine and became a promising field for detecting and providing assistance in automated medical diagnoses. Fundus camera is used to acquire Fundus images of the interior portion of the eye. It becomes very efficient and effective by using automated techniques for the doctors to diagnose various diseases. With the development of new algorithms and advancement of technology it has become easier to diagnose the diseases like Glaucoma and DR (Diabetic Retinopathy) with an improvement in the accuracy of diagnosis.

The images that are captured using the digital fundus camera have a limited field of view and normally range with in the limit of 30o to 45o restricting the visualization of the image to a limited area resulting in problems for the detection of disease. Therefor more than one images are captured so that the doctor can diagnose or even if it is further used for automated segmentation of different diseases. Fig. 1 shows a fundus image taken from a normal fundus camera and it can be observed that not all the region of retina is present in that region. In the fields of medical imaging, computer vision, remote sensing and pattern recognition image aligning and then stitching them together so that a mosaic can be generated different algorithms are used [1]. For this purpose two methodologies are followed. In first method a suitable searching technique is used for the searching of aligning point along with the creation of error metrics, but the performance of this method is not optimal [2]. In second method feature set extraction and then feature matching is used to match the images [3]. Due to fast androbustness, this method has advantage on the first method and is widely used.

Digital Fundus camera is used to acquire the retinal images. It is a camera attached to a low power microscope so that image of retina can be extracted. Fundus camera that is used to capture the image of retina is shown in Fig. 2(a) [4]. It is not possible for the camera to cover the complete field of view, as shown in Fig. 2(b).Therefor Image stitching technique is used to combine several images that are overlapping and taken from different angle with different field of view. This broadens the scope of the image, just like a camera is used to take a panorama image.

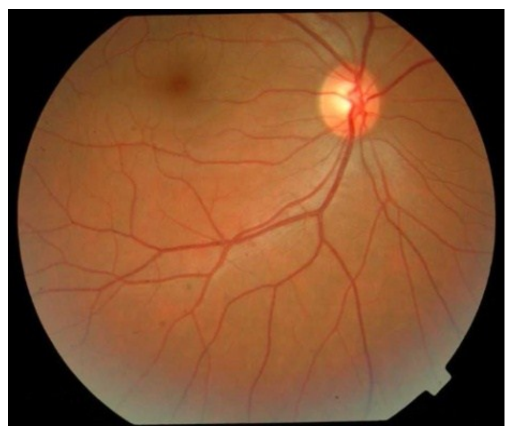

Fig.1. Restriction of the field of view in digital Fundus Image of the patient

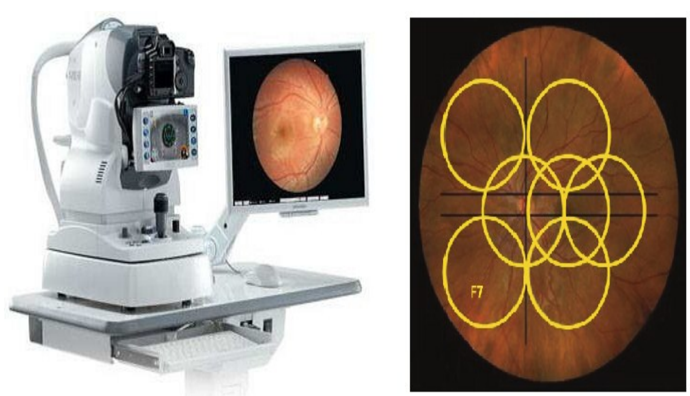

Fig.2.Fundus Camera and restricted field of view.

The main motivation is to align the images in such a way that they are attached seamlessly. The main task in generation of mosaic image is the feature extraction. Once features extracted then features matching and image alignment is done, and then after post processing we get a perfect mosaic image with greater field of view. The most common descriptor used for feature extraction are sparse descriptor which are SIFT [5], ASIFT [6] and SURF [7]. These descriptors are capable of detecting interest points located sparsely in an image

II. RELATED WORK

Different researchers have been using different techniques to generate the mosaic images some are discussed under

1. Li Fang Wei et al. presented a technique for retinal image mosaic [8]. PCA-SIFT (Principle Component Analysis-Scale Invariant Feature Transformation) was used in their proposed method for feature detection. After the selection of the feature set quadratic transformation model is used to simulate the anatomy of the eye. K-nearest neighbors was used for feature matching. Then random sample consensus (RANSAC) was used so to detect the inliers and then remove the outliers present in the image and in the end seamless stitching was achieved by using weighted mean.
2. LI Jupeng et al. proposed an algorithm that is able to generate automated mosaic of the curved human colour retinal images [9]. m-SIFT was used for feature detection as when flat image is used SIFT is unable to find features. Then for feature matching second nearest neighboring strategy was used. Once features are selected then the inliers of both images are selected. After that to align the images the bilinear wrapping technique was used and at the end to generate panoramic images multi blending technique was used. The drawback of this paper was that it was limited to stitching of two images.
3. Philippe C. Cattinet al. proposed an algorithm for generation of retinal mosaickingimaging using local features [10]. A 128 dimensional descriptor along with a hessian matrix based detector for the interest point detection was used. k-nearest neighbours was used for feature matching, while to match the inliers across the images RANSAC was used while some outliers were also filtered out by using RANSAC. After this anchor image selection and mapping estimation is completed and in the last they proposed to blend their images using laplacian pyramid.
4. Tae Eun Choe et al. proposes a method to generate optimal global mosaic images from retinal images [11]. In their proposed method they have used Y features that are used for feature extraction and for feature matching windows that are enclosing y-features were used. Once features are selected then for inliers matching RANSAC was used, and to find the optimal registration global pair-wise registration was used. For the selection of reference frame the frame that had the lowest selection error was selected. Once reference frame was selected all other images were blended onto the reference frame depending on the shortest path of each image.

III. PROPOSED METHODOLOGY

In this paper we have used WLD (Weber Local Descriptor) for the extraction of features. Which are further used for feature matching among the images that are to be stitched together and an image with larger field of view that fundus image can be generated. WLD is a dense descriptor and searches pixel by pixel for the extraction of features from the image being processed. After the extraction of features inliers are found through RANSAC. As the proposed system is capable of generating a mosaic image of more than two images therefore a similarity matrix is generated to find out that which two images will be stitched together. Then to align the images holography is calculated and images are blended together. Now if there is another images to be stitched together to the previously generated mosaic image, similarity matrix using RANSAC is generated again so the cycle continues. Flow chart of the proposed system can be seen in Fig. 3.Proposed algorithm is explained in detail below.

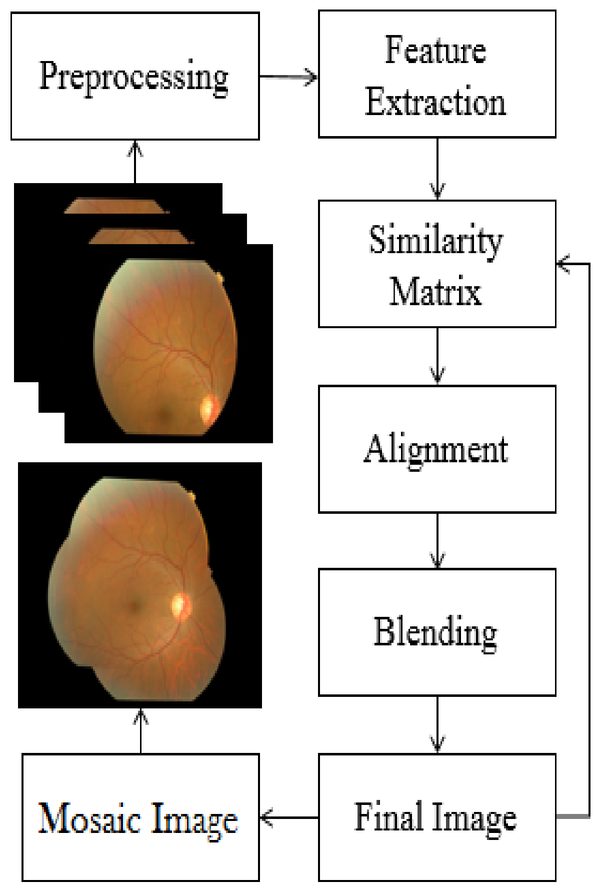

Fig.1. Flow Chart of proposed methodology

A. Weber Local Descriptor

There are two types of descriptors sparse descriptors and dense descriptors. Dense descriptors such as Local Binary descriptor and Weber local descriptors extract the feature using pixel y pixel data of the image. While sparse descriptors such as SIFT, SURF and ASIFT use corners and edges that are sparsely located on the given image. We have used WLD, it is generated from Weber's law.

Weber's law states that ratio of background intensity to increment threshold is constant, it is illustrated by equation 1.

$$\frac{\Delta I}{\mathrm{I}} = \boldsymbol{\kappa} \quad (1)$$

In the above equation Incremental threshold in the above equation is represented by ΔI, initial stimulus intensity is represented by I, while κ is the constant that remains the same if thereis a change in intensity on the left side of the equation. WLD contains two main components, Differential excitation (ξ) and Orientation (Θ).

i. Differential Excitation

The ratio between two terms is known as differential equation. First term is the intensity of current pixel and the second term is the difference calculated between average intensity of the current intensity and neighbors.

Differential excitation is computed by weber's law and uses the ratio change in initial intensity of current pixel at that particular moment and the change in the current pixel can be calculated by calculating the difference between the current pixel and their neighbors. In equation form it can be written as following

$$V_s^{00} = \sum_{i=0}^{p-1}(\Delta x_i) = \sum_{i=0}^{p-1}(x_i - x_c) \quad (2)$$

In this equation V_s^{00} represents the change in intensity, the number of neighbors' is represented by p, xc is the representation of the current pixel and Xi (i=0, 1,…..p-1) represents the i-th neighbor of xc. Differential excitation is the ratio of two terms as discussed earlier which is represented through eq. 3.

$$G_{ratio}(x_c) = {v_s^{00}}\Big/{v_s^{01}} \quad (3)$$

In the above equation Vs01 is the representation of intensity of the current pixel. There is a problem with the values that they may be too large or they can fall to a very small value depending on the input image, therefor to lemmatize the values arctangent function is applied and the values are set in between an upper and a lower limit. There for once the arctangent function is applied the equation can be written as shown in eq. 4.

$$\xi(x_c) = arctan\left[\frac{v_s^{00}}{v_s^{01}}\right] = arctan\left[\sum_{i=0}^{p-1}\left(\frac{x_i - x_c}{x_c}\right)\right] \quad (4)$$

This process can be seen in Fig. 4.

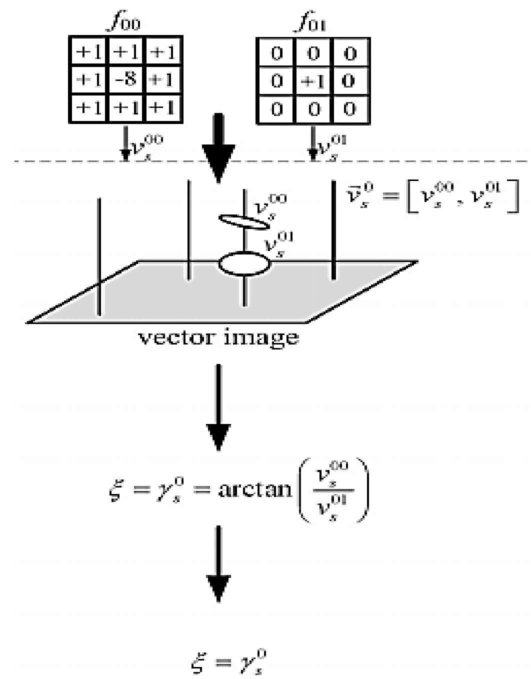

Fig.2. Differential Excitation [13]

Weber local descriptor is robust to change in contrast and can be seen from equation 4 and by the change in the contrast of each pixel initial intensity of current pixel is then multiplied and is divided by the intensity of the current pixel.

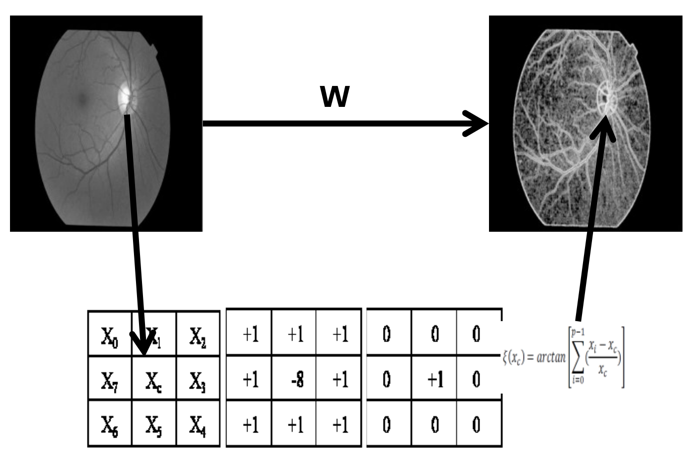

Fig.3. Single pixel used for Computing of Differential Excitation in an image

Fig. 5 shows how the calculation of single pixel in a digital Fundus Image is done. Difference of neighbor pixel with the current pixel can be seen through separate two filters respectively. Based on the distance from the current pixel neighbor pixels are selected, for example if we use the value of R=1 than pixels that are present at distance 1 will be selected as new neighbors', similarly if we take the value of R=2 then pixels at distance 2 will be selected as new neighbors'. We have used R=1 with selection of 8.

ii. Orientation

The second component is the orientation it can also be called gradient orientation. It can be computed by using the following eq.5.

$$\theta(x_c) = \gamma_s^1 = \arctan\left(\frac{v_s^{11}}{v_s^{10}}\right) \quad (5)$$

Output of filters f_s^{10} and f_s^{11} in the above equation which is the difference of vertical and horizontal neighbors' respectively is represented byv_s^{10} and v_s^{11} where $\boldsymbol{v_s^{10}} = \boldsymbol{x_5 - x_1}$ and $\boldsymbol{v_s^{11}} = \boldsymbol{x_7 - x_3}$.

We have used only differential excitation in our proposed algorithm. WLD is robust to illumination changes and noise

B. RANSAC

Once the features are detected then the next step is match the features of the given images and find the inliers and outliers in the images. All possible inliers in the two digital fundus images can be seen in Fig. 6. As it is known that RANSAC is an iterative approach so there for there will be high number of outliers in the figure 4 which will be removed in the later stages. All those combinations that have maximum inliers are chosen in the second step of RANSAC. Refined inliers matching can be seen in Fig. 7.

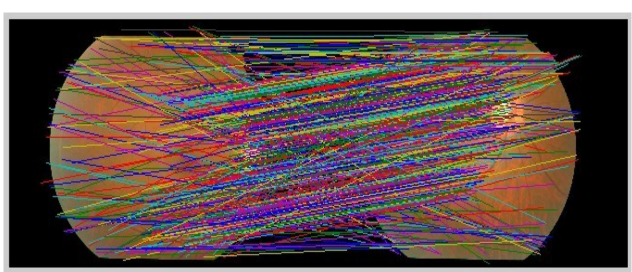

Fig.6. All Possible Inliers Matches between Fundus Images

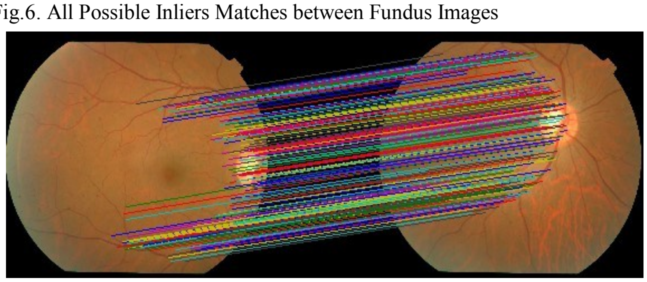

Fig. 7.Matching of Refined Inliers

C. Homography

It is a technique used for registration of images mostly [12]. In this technique one image used as constant and the other image is rotated and translated to bring it to optimal alignment for stitching. The result of the blended image can be seen in Fig. 8.

Fig. 8. Mosaic Image generated using two Images

D. Seamless Stitching

Now to achieve seamless stitching at the end a new algorithm is proposed. The steps for achieving the seamless stitching are explained below.

- The first step is to generate mask1 and mask2, for both input images the BG mask is generated (image1, image2).
- Then obtain fmask1, Gaussian low pass filter in freq. domain is used to filter mask1.
- Image1multiplied with fmask1.
- Image2 multiplied with inverted fmask2.
- Then after the summation of both multiplications result is generated.
- Result 2 is masked with the result and result1 is obtained.
- Image1 is masked with the mask generated from X-OR of mask1 and mask2, this was result2 is generated.
- And at the end summation of result1 and result2 will produce the final result.

Results of seamless stitching can be seen in Fig. 9, Fig 9-a shows the mosaic image without seamless stitching algorithm applied to it and Fig. 9-b represents the resulting image generated after applying the seamless stitching algorithm.

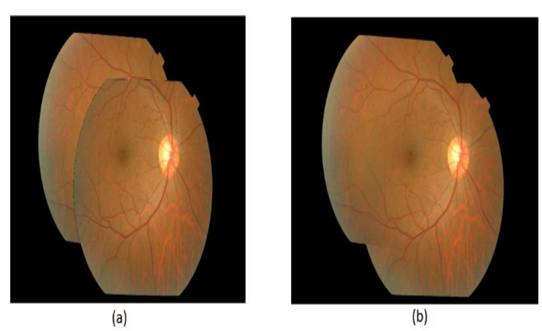

Fig. 9. Before and after seamless stitching

IV. RESULTS

A locally gathered database has been used for the evaluation of the proposed system. The dataset was collected with collaboration of AFIO (Armed Forces institute of Ophthalmology), Rawalpindi Pakistan. Topcon non mydriatic camera 200 cameras NW was used to collect the images and the number of 56 images were collected. A number of fifteen images were selected for this purpose and each patient had a minimum number of three images and a maximum number of six images for every patient that was selected for the dataset. Visual inspection was done to check the validity of the system.

The first three fundus images of a same patient to generate a mosaic image and the generated mosaic image are shown in Fig. 10. As discussed earlier first the mosaic image of first two images will be generated and then using that mosaic image as base the final mosaic image will be.

The proposed algorithm is capable of producing a mosaic image of 6 Fundus images using the same iterative process as discussed earlier. The six fundus image used to generate a mosaic image and the generated mosaic image is shown in Fig. 11.

The performance of the proposed system was evaluated with the comparison of other state of the art techniques present. For the comparison purpose the evaluation criterion was based on Accuracy and time

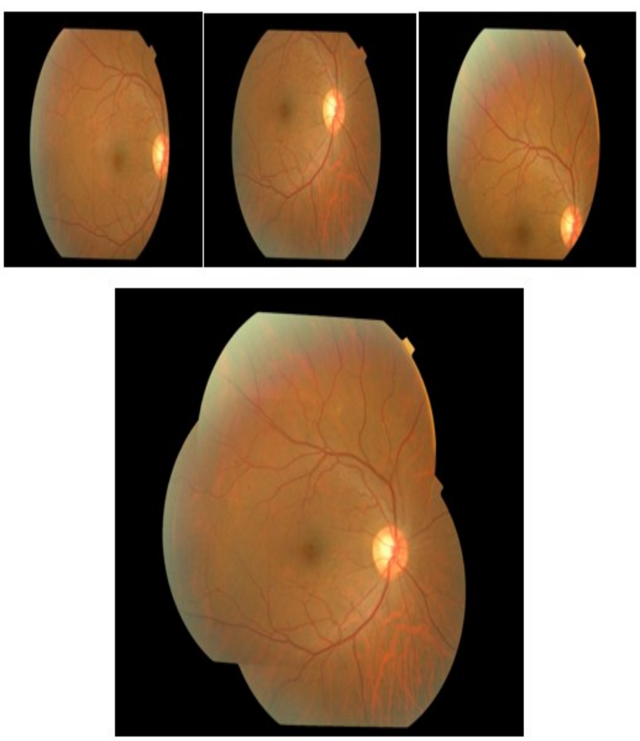

Fig. 10. Fundus Images and their Mosaic Image

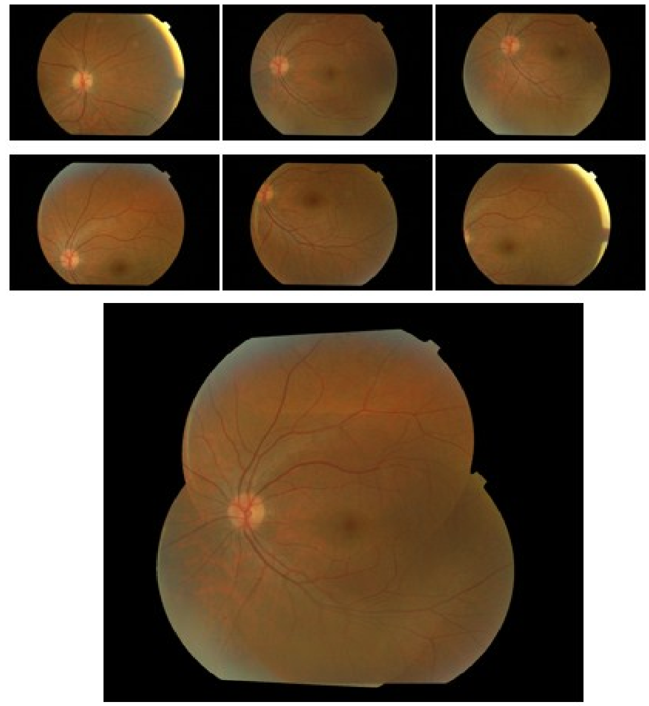

Fig.11. Fundus Images and the generated Mosaic Image

A. Accuracy

Accuracy of the result is measured on the basis of continuity of vessels perceived by human eye. Other state of the art techniques resultwas compared with our proposed system and the results are shown in the table below.

It can be seen in Table 1 shows that the results produced by SIFT and SURF are not accurate while ASIFT and WLD are able to produce similar results. Similarly, when mosaic was generation for 6 images then ASIFT and WLD were able to produce the accurate results while the results produced from SURF were extremely poor while the results produced from SIFT were inconsistent.

B. Timing Table

Various techniques average time taken to produce the results is given in the Table 2. The system used for the processing was Haier Core i3 fourth Gen with 1.70 GHz processor and 8 GB RAM. It can be seen form table 2 that SIFT and ASIFT were slower when compared with WLD technique.

Table 1 Comparison of Proposed System with other State of the Art Techniques

Techniques Used	Total Patients	Correct	Accuracy %
SURF	15	10	66.6
SIFT	15	13	86.6
ASIFT	15	15	100
WLD	15	15	100

TABLE 2.Average execution time

Techniques Used	2- Images	3- Images
SURF	13.78	68.13
SIFT	75.14	116.96
ASIFT	133.12	183.28
WLD	63.13	95.24

V. Conclusion

The proposed methodology was used to generate a mosaic image and then to remove the boundary in between the stitched images a seamless stitching algorithm was proposed and in the end the results were compared with other state of the art techniques and were observed that the methodology worked upon was the optimal for the generation of Mosaic images.

VI. Future work

In future this mosaic image can be further used in automated diagnosis of retinal diseases and will be very helpful in generating accurate results due to its larger field of view

References

[1] Yuliang Wang, Jianxin Shen, Wenhe Liao and Lin Zho: Automatic Fundus Images Mosaic Based on SIFT Feature, (2010)
[2] Richard Szeliski: Image Alignment and Stitching: A Tutorial, (2005)
[3] Nice, France, Matthew Brown and David G. Lowe: "Recognizing panoramas," International Conference on Computer Vision (ICCV 2003), (2003)
[4] Early Treatment Diabetic Retinopathy Study research group, “Early photocoagulation for diabetic retinopathy. ETDRS Rep. 9,” Ophthalmology, vol. 98, no. 5 suppl, pp. 766–785, (1991)
[5] David Lowe: "Object recognition from local scale-invariant features". Proceedings of the International Conference on Computer Vision, (1999)
[6] Jean-Michel Morel and Guoshen Yu: ASIFT: A new framework for fully Affine Invariant image comparison, (2011)
[7] M. A. Fischler and R. C. Bolles: Random sample consensus: A paradigm for model-fitting with applications to image analysis and automated cartography, Communications of the ACM 24. 381–395 (1981)
[8] Li Fang Wei, LinLin Huang, Lin Pan and Lun Yu: The Retinal Image Mosaic Based on Invariant Feature and Hierarchical Transformation Models, 2nd International conference on Image And Signal Processing, (2009)
[9] LI Jupeng, CHEN Houjin, YAO Chang and ZHANG Xinyuan: A Robust Feature Based Method for Mosaic of the curved Human Color Retinal Images, International Conference on Bio Medical Engineering and Informatics, (2008)
[10] Philippe C. Cattin, Hobert Bay, Luc Van Gool and Gabor Szekely: Retina Mosaicing Using Local Features, Medical Image Computing and Computer-Assisted Intervention, (2006)
[11] Tae Eun Choe, Isaac Cohen, Munwai Lee and Gérard Medioni: Optimal Global Mosaic Generation from Retinal Images, Proceedings of the 18th International Conference on Pattern Recognition, ICPR, (2006)
[12] D. Kriegman, “Homography Estimation”, Winter 2007
[13] Chen, Jie, Shiguang Shan, Chu He, Guoying Zhao, Matti Pietikäinen, Xilin Chen, and Wen Gao. "WLD: A robust local image descriptor." Pattern Analysis and Machine Intelligence, IEEE Transactions on 32, no. 9 (2010): 1705-1720.

Authors

Joddat Fatima, joddat.fatima@gmail.com
Bahria University Islamabad, Pakistan

M. Usman Akram, usmakram@gmail.com
National University of Sciences and Technology, Pakistan

Adeel M. Syed, adeelmuzaffar@gmail.com
Bahria University Islamabad, Pakistan

M. Usman Akbar, muhamadusman@outlook.com
Bahria University Islamabad, Pakistan