



Creation of Retinal Mosaics for Diabetic Retinopathy Screening: A Comparative Study

Tânia Melo¹(✉) , Ana Maria Mendonça^{1,2} , and Aurélio Campilho^{1,2} 

¹ INESC TEC - Institute for Systems and Computer Engineering,
Technology and Science, Porto, Portugal
`tania.f.melo@inesctec.pt`

² Faculty of Engineering of the University of Porto, Porto, Portugal
`{amendon,campilho}@fe.up.pt`

Abstract. The creation of retinal mosaics from sets of fundus photographs can significantly reduce the time spent on the diabetic retinopathy (DR) screening, because through mosaic analysis the ophthalmologists can examine several portions of the eye at a single glance and, consequently, detect and grade DR more easily. Like most of the methods described in the literature, this methodology includes two main steps: image registration and image blending. In the registration step, relevant keypoints are detected on all images, the transformation matrices are estimated based on the correspondences between those keypoints and the images are reprojected into the same coordinate system. However, the main contributions of this work are in the blending step. In order to combine the overlapping images, a color compensation is applied to those images and a distance-based map of weights is computed for each one. The methodology is applied to two different datasets and the mosaics obtained for one of them are visually compared with the results of two state-of-the-art methods. The mosaics obtained with our method present good quality and they can be used for DR grading.

Keywords: Diabetic retinopathy screening · Retinal mosaicking
Image registration · Image blending · Qualitative evaluation

1 Introduction

Diabetic retinopathy (DR) is an eye disease that affects the vision of a high number of diabetic people worldwide. It is characterized by the presence of several retinal lesions, being a leading cause of blindness [7]. In order to prevent the development of DR, regular eye examinations need to be done in the diabetic patients.

However, in a screening program, there is a large number of retinal images to be manually examined by the medical doctors, which is labor intensive, time consuming and expensive [6]. Furthermore, in these screening programs, there

is usually more than one image per eye, since the analysis of a single image may not be enough to make a good eye evaluation. Thus, the creation of retinal mosaics from images of the same eye can reduce substantially the time spent by the medical doctors on DR screening. Through the analysis of mosaics, the ophthalmologists can do a complete eye examination at a single glance instead of examining each portion of the eye in a different image [2].

Therefore, a methodology for retinal mosaic creation is herein proposed. The main novelty of this methodology is related to the sequence of operations performed for combining the overlapping images.

Section 2 presents a brief introduction about image mosaicking and a summary of the most relevant state-of-the-art methods in this research area. A description of the proposed methodology is presented in Sect. 3. The results obtained are shown and discussed in Sect. 4 and the main conclusions are stated in Sect. 5.

2 Image Mosaicking Methods

Image mosaicking is the process of obtaining a wider field-of-view of a 3D scene through the alignment of a set of overlapping images which represent part of that scene [4]. Due to its wide range of applications, this process has become an attractive research area and, thus, a high number of image mosaicking algorithms have been developed over the last years. A large number of those methods are focused on the creation of retinal mosaics.

The image mosaicking methods comprise two main steps: image registration and image blending. The first step consists in the identification of common regions in the images and their alignment into the same coordinate system based on the estimated geometric transformations; the second step aims to minimize the discontinuities in the transition regions. As mentioned in [4], the image mosaicking methods can be classified according to the algorithms used in these two steps.

Regarding to the registration step, the methods can be classified as global, local or hybrid methods. While global methods align the images based on the similarity of the pixel intensities, local methods detect relevant keypoints in the images, find the correspondences between those points and use them in order to estimate the geometric transformation that registers the images. Hybrid methods combine both strategies [5]. Legg et al. [8] developed a global method which computes the mutual information between “windows” of the images to be registered. Since the mutual information is obtained by computing the entropy from the probability distribution of the data, the authors also propose the use of an adaptive probability distribution estimation in order to increase the accuracy of the registration procedure. In a different way, Hernandez-Matas et al. [5] apply the Scale Invariant Feature Transform (SIFT) [9] to detect and describe some relevant keypoints in the images and then estimate the geometric transformation based on the correspondences established between the points detected. In order to discard the incorrect correspondences (outliers) and obtain more accurate estimations of the transformation matrix, the authors employ the Random

Sample Consensus (RANSAC) method [3]. Stewart et al. [11] also developed a local method, but instead of applying a feature detector to obtain the salient points, the authors use vascular landmarks (branching and cross-over points) for registering the images. Adal et al. [1] proposed a hybrid model that aligns the images based on a vasculature-weighted mean squared difference similarity metric. With this metric the authors combine both global and local cues. In general, local or feature-based methods present best performance in the registration of image pairs with a small overlap in comparison to global or area-based methods.

Regarding to the transformation models used for warping the images, some of the state-of-the-art methods apply 2D linear transformations. For instance, Matsopoulos et al. [10] detect the vessel centerlines, extract the vessel bifurcation points and use a novel implementation of self organizing maps for establishing the correspondences between those points. Then, the parameters of the affine transform are estimated based on those correspondences. However, since linear transformations do not take into account the curvature of the eye, several authors apply non-linear transformations to obtain more accurate results. For instance, Adal et al. [1] use a deformation model with increasing complexity in order to estimate the 12 parameters of a global quadratic model. Since the non-linear transformations do not take necessarily into account the shape and size of the eye, Hernandez-Matas et al. [5] apply a 3D model of the eye to estimate the transformation that relates the views from which the images to be registered were acquired.

Although the methods used for registering retinal images are slightly different from those used for aligning other types of images (because the scenes which are captured have different shapes and different landmarks), the methods applied for blending images of different 3D scenes can be similar. As mentioned in [4], these methods can be divided into three main groups: feathering-based, pyramid-based and gradient-based methods [4]. Feathering-based methods blend the images through a weighted combination of the overlapping images, pyramid-based methods convert the input images into band-pass pyramids to obtain the final mosaics and the gradient-based methods set the gradients across seams to zero for smoothing out the color differences [4].

3 Methodology

Like most of the state-of-the-art methods, the methodology herein proposed for retinal mosaicking comprises two main steps: retinal image registration and blending.

Although the methods applied for registering the retinal images have already been used by other authors for this purpose, the operations performed in the blending step in order to smooth the transitions between the overlapping images are novel in this application field.

In the following subsections a detailed description of these methods is done.

3.1 Registration of Retinal Images

In order to accurately align two retinal images, a sequence of operations is performed (Fig. 1).

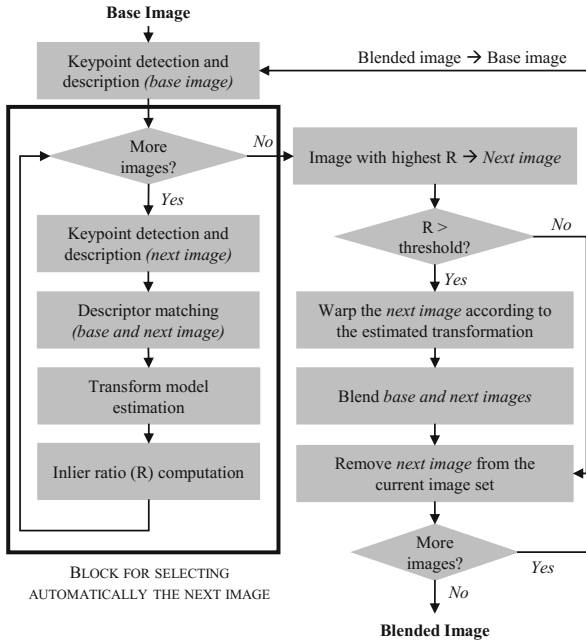


Fig. 1. General scheme of the methodology implemented

First, several keypoints are detected in both images and described based on the Scale Invariant Feature Transform (SIFT) [9]. The advantage of using SIFT is that the detected keypoints are invariant to image scaling and rotation, they are distributed over the entire image and the descriptor is distinctive, which allows a keypoint to find its correct match with high probability [5, 9].

Then, for each keypoint detected in the next image (image which is being registered), the corresponding keypoint (closest neighbor) in the base image is found through the computation of the Euclidean distance between the keypoint descriptors. Since the keypoints which are located outside the overlapping region have not a corresponding keypoint in the base image, a large number of incorrect matches are expected. In order to exclude these matches, a threshold is applied to the ratio between the distances to the closest and the second closest neighbors of each keypoint [9].

Since some incorrect matches still remain after this filtering, the RANSAC method is used for finding the consistent matches and estimating the parameters of the geometric transformation. Taking into account the curvature of the eye, a non-linear quadratic transformation is applied. Then, the inlier ratio is computed

by dividing the number of matches used by the RANSAC method for estimating the twelve parameters of the polynomial transformation by the total number of matches.

Knowing that the main goal of this work is to create retinal mosaics from sets of two or more images, it is necessary to define the order by which the images are registered. In order to do that automatically, the steps described above are repeated for all pairs that can be formed by the base image and the images which have not yet been attached. The next image to be registered is associated with the image pair with the highest inlier ratio (see Fig. 1).

In order to avoid the incorrect alignment of images that do not overlap, the next image is only reprojected into the same coordinate system of the base image if the inlier ratio is greater than 0.1. Otherwise, that image is discarded and it does not appear in the final mosaic.

3.2 Blending of the Overlapping Images

After reprojecting the overlapping images into the same coordinate system, the pixels of the overlapping regions have to be combined in order to smooth the transitions between the images and obtain mosaics with high quality.

Considering that the overlapping images can present different illumination patterns and corresponding points can have different intensities in the two images, an illumination correction is performed (Fig. 2).

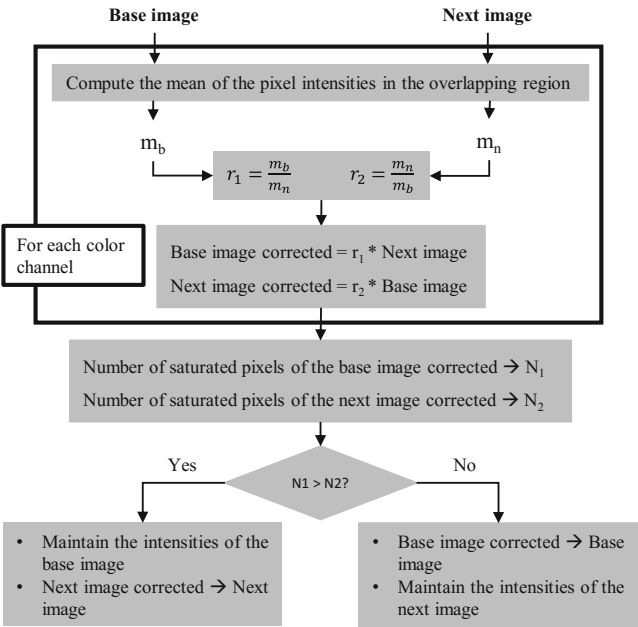


Fig. 2. Overall scheme of the illumination correction process

For that, the mean of the pixel intensities in the overlapping region is computed for each image separately. Then, two parameters (r_1 and r_2) are computed based on those mean values (see Fig. 2) and multiplied by the registered images in order to obtain the intensity compensated images. The image which is definitively corrected after this process is the one which presents the lowest number of saturated pixels. The pixel intensities of the other image remain constant.

After the illumination correction process, the two overlapping images are combined for obtaining the final mosaic (Fig. 3).

In order to identify the region where the two images intersect, their masks have to be obtained. For that, thresholding and morphological operations are applied. The mask of the overlapping region results from the intersection of the masks of the registered images.

A map of weights is also computed for each image (Fig. 3E and F) based on the Euclidean distances between the pixels of the overlapping region and the closest point of that image which is outside the intersection zone (see Eqs. 1 and 2). In the end, the pixel intensities of the composite image are a weighted sum of the pixel intensities of the base and next images (Eq. 3).

$$w_1(x, y) = \begin{cases} 0 & \text{if } m_1(x, y) = 0 \\ \frac{d_2}{d_1 + d_2} & \text{if } m_1(x, y) = 1 \wedge R(x, y) = 1 \\ 1 & \text{otherwise} \end{cases} \quad (1)$$

$$w_2(x, y) = \begin{cases} 0 & \text{if } m_2(x, y) = 0 \\ \frac{d_1}{d_1 + d_2} & \text{if } m_2(x, y) = 1 \wedge R(x, y) = 1 \\ 1 & \text{otherwise} \end{cases} \quad (2)$$

$$I(x, y) = w_1(x, y) * I_1(x, y) + w_2(x, y) * I_2(x, y) \quad (3)$$

where I_1 and I_2 are the base and next images aligned into the same coordinate system, m_1 and m_2 are their masks, R is the mask of the overlapping region and d_1 and d_2 are the Euclidean distances between the point (x, y) and the closest point in m_1 and m_2 , respectively, which is outside the intersection zone.

4 Results

4.1 Datasets

In order to evaluate the performance of the methodology developed for retinal mosaicking, two datasets are used. One of them is the dataset stored in the Rotterdam Ophthalmic Data Repository. This dataset includes 1120 fundus images of size 2000×1312 pixels and a 45° field-of-view. The images were acquired from 70 patients during two different visits. For each patient, 4 images per eye and visit are available. In this work, only intravisit mosaics are created. Therefore, images which are used by Adal et al. [2] for creating other types of mosaics (intervisit mosaics) are discarded.

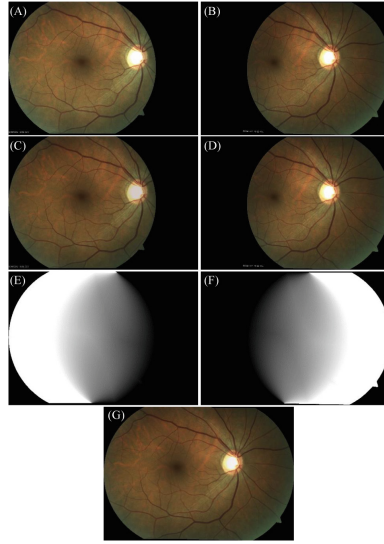


Fig. 3. Image blending: (A-B) Input images reprojected into the same coordinate system; (C-D) Images A and B, respectively, after color compensation (only the intensities of the image A were changed during the process); (E-F) Weight maps associated with C and D, respectively; (G) Blended image

Since the final intravital mosaics obtained from these images through the methods proposed in [2,11] are also publicly available, they are used in this work for comparison purposes. However, these mosaics were obtained from the images of this dataset after performing the normalization referred in [2]. In order to obtain similar mosaics and be possible the comparison with the other methods, the normalized images provided with the dataset are used as input images in our method.

The performance of the methodology herein presented is also evaluated for a private dataset provided by Hospital S. João (Porto). In this dataset, there are images from the two eyes of 21 patients. For each eye, 3 images with a 45° field-of-view and 7 images with a 30° field-of-view are available. However, for this dataset there are not other mosaics to compare with the results of our method.

4.2 Discussion of the Results

In this work, the quality of the final mosaics obtained from the images captured in the Rotterdam Eye Hospital is visually evaluated by comparison with the mosaics obtained by the methods proposed in [2,11].

Although most image sets give rise to high quality mosaics, there is a small number of mosaics which are not composed by the four images or have at least one image misplaced. Figures 4 and 5 show an example of a good and a bad result, respectively.

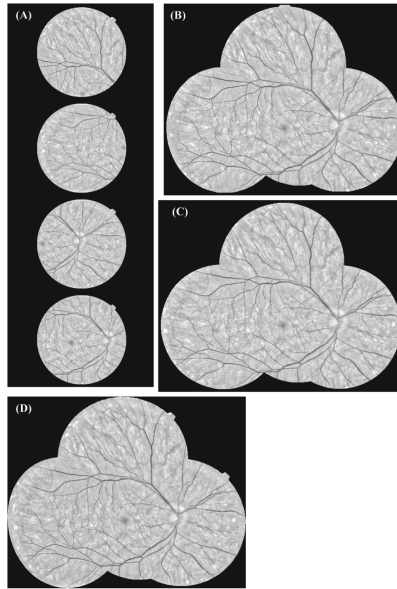


Fig. 4. Retinal mosaicking - example 1: (A) Original images; (B-C) Mosaics obtained with the methods described in [11] and [2], respectively; (D) Mosaic obtained with the method herein proposed

Through the analysis of Fig. 4, it is possible to observe that the mosaic obtained with our method from the images shown in Fig. 4A is very similar to the results obtained with the other state-of-the-art methods. Taking into account that the other two mosaics were classified as “acceptable” by an experienced grader [2], we can also infer that the mosaic created with our method also has an “acceptable” quality and it can be used for DR grading.

In the other hand, the low quality of the original images presented in Fig. 5A, as well as the low degree of overlap between them, interferes with the number of keypoints detected in the overlapping regions. Because of this, the RANSAC method is not able to detect a sufficient number of consistent matches and the method considers that the images do not overlap. In this way, the images are not registered and an incomplete mosaic is obtained. However, the mosaics obtained by the other methods (Fig. 5B and C) for the same set of images also have at least one image misplaced.

Although the registration step fails in special cases, the transitions between overlapping images are imperceptible in all mosaics. This allows to conclude that the operations proposed in this work for image blending produces very good results.

In order to show that this method can be applied to different datasets, the methodology herein proposed is also applied to the dataset provided by Hospital S. João (see an example in Fig. 3). For this set of images, the final mosaics also present a good quality, which proves the robustness of our method.

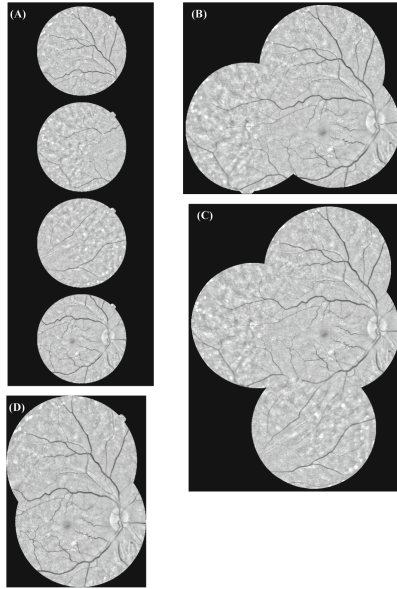


Fig. 5. Retinal mosaicking - example 2: (A) Original images; (B-C) Mosaics obtained with the methods described in [11] and [2], respectively; (D) Mosaic obtained with the method herein proposed

5 Conclusions

In this work, it is possible to conclude that the mosaics obtained with our method are very similar to the mosaics obtained using the methods proposed in [2, 11]. Although most mosaics present high quality and can be used for DR grading, there is a small number of mosaics that appear incomplete or have at least one image misplaced. This happens mainly when the original images present low quality or the overlapping degree between them is too small that few keypoints are detected in the overlapping region. In this way, the use of a more robust keypoint detector can improve the performance of the method herein proposed, specially in the cases where the overlapping degree between the images is small.

Although the registration step fails in these special cases, all mosaics present imperceptible transitions between the overlapping images, which demonstrates the good performance of the operations introduced in the blending step.

Although the scores given by two experienced graders to the mosaics obtained using the methods described in [2, 11] have been used as reference for evaluating the performance of the methodology herein proposed, as future work we propose the use of a quantitative measure for evaluating more accurately the quality of the final mosaics.

Acknowledgments. This work is financed by the ERDF - European Regional Development Fund through the Operational Programme for Competitiveness and Interna-

tionalisation - COMPETE 2020 Programme, and by National Funds through the FCT - Fundação para a Ciência e a Tecnologia within project CMUP-ERI/TIC/0028/2014.

References

1. Adal, K.M., Ensing, R.M., Couvert, R., van Etten, P., Martinez, J.P., Vermeer, K.A., van Vliet, L.J.: A hierarchical coarse-to-fine approach for fundus image registration. In: Ourselin, S., Modat, M. (eds.) WBIR 2014. LNCS, vol. 8545, pp. 93–102. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08554-8_10
2. Adal, K.M., van Etten, P.G., Martinez, J.P., van Vliet, L.J., Vermeer, K.A.: Accuracy assessment of intra- and intervisit fundus image registration for diabetic retinopathy screening. *Invest. Ophthalmol. Vis. Sci.* **56**(3), 1805–1812 (2015). <https://doi.org/10.1167/iovs.14-15949>
3. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **24**(6), 381–395 (1981). <https://doi.org/10.1145/358669.358692>
4. Ghosh, D., Kaabouch, N.: A survey on image mosaicing techniques. *J. Vis. Commun. Image Represent.* **34**, 1–11 (2016). <https://doi.org/10.1016/j.jvcir.2015.10.014>
5. Hernandez-Matas, C., Zabulis, X., Triantafyllou, A., Anyfanti, P., Argyros, A.A.: Retinal image registration under the assumption of a spherical eye. *Comput. Med. Imaging Graph.* **55**, 95–105 (2017). <https://doi.org/10.1016/j.compmedimag.2016.06.006>
6. Jelinek, H., Cree, M.: *Automated Image Detection of Retinal Pathology*. CRC Press, Boca Raton (2009)
7. Lee, R., Wong, T.Y., Sabanayagam, C.: Epidemiology of diabetic retinopathy, diabetic macular edema and related vision loss. *Eye Vis.* **2**(1), 17 (2015). <https://doi.org/10.1186/s40662-015-0026-2>
8. Legg, P.A., Rosin, P.L., Marshall, D., Morgan, J.E.: Improving accuracy and efficiency of mutual information for multi-modal retinal image registration using adaptive probability density estimation. *Comput. Med. Imaging Graph.* **37**(7), 597–606 (2013). <https://doi.org/10.1016/j.compmedimag.2013.08.004>
9. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**(2), 91–110 (2004). <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
10. Matsopoulos, G.K., Asvestas, P.A., Mouravliansky, N.A., Delibasis, K.K.: Multi-modal registration of retinal images using self organizing maps. *IEEE Trans. Med. Imaging* **23**(12), 1557–1563 (2004). <https://doi.org/10.1109/TMI.2004.836547>
11. Stewart, C.V., Tsai, C.L., Roysam, B.: The dual-bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE Trans. Med. Imaging* **22**(11), 1379–1394 (2003). <https://doi.org/10.1109/TMI.2003.819276>