AUTOMATIC ADAPTIVE OPTICS RETINAL IMAGES MONTAGING

Eva Valterova
Doctoral Degree Programme (3.), FEEC BUT
E-mail: valterova@vutbr.cz

Supervised by: Radim Kolar
E-mail: kolarr@feec.vutbr.cz

Abstract: The adaptive optics images capture limited field of view. Nevertheless, the quantitative retina assessment requires analysis over extended areas captured by images on various retina positions. The fully automated method for image registration and montaging is presented. The method utilizes scale invariant feature transform (SIFT) for feature extraction from preprocessed images. The method is tested on 200 images of normal healthy patients. The montages of 20 eyes were created and evaluated by normalized mutual information metrics and the results showed high alignment accuracy.

Keywords: retina, adaptive optics, image registration, SIFT

1 INTRODUCTION

Adaptive optics (AO) technology offers high resolutions retinal imaging. The wavefront sensor and deformable mirror measure and correct optical aberrations present in the optic path of the human eye [1]. Thus the resolution near diffraction is reached and individual retinal cell types, such as ganglion cells, photoreceptors and retinal pigment epithelial cells are visible *in vivo* [2, 3]. The structure and function of retinal cell observation are possible to utilize as a relevant marker of pathological retinal changes or disease monitoring [1, 4, 5].

The AO imaging struggle with a limited field of view, typically of size 4° by 4°, which corresponds to approximately 1.1 mm^2. Nevertheless, larger retinal area acquisition is necessary for the comprehensive examination, thus multiple adjacent images must be acquired across overlapping retinal regions and montaged to cover a larger retina area. The montage requires accurate registration of individual adjacent images. Generally, the montage can be created in Photoshop (Adobe Systems, Mountain View, California) [6], ImageJ (National Institutes of Health, Bethesda, Maryland) [7] or other programs based on manual registration. However, this approach is highly time-consuming, requires an experienced evaluator, and is moreover complicated by image noise, blur, and contrast variability across images.

Several automated solutions for adaptive optics scanning laser ophthalmoscopy (AOSLO) image montaging have been published. The registration part of montaging is mostly based on image transformation estimation from detected specific features [8, 9, 10, 11]. *Li et al.* [9] used principal component analysis-scale invariant features (PCA-SIFT) to match horizontally AOSLO adjacent images. The outlier features gained by SIFT are eliminated by the random sample consensus (RANSAC) algorithm. The horizontally montaged images have inhomogeneous illumination, thus were registered in one wide-field montage by cross-correlation. Inspired by this approach, an automated algorithm for AOSLO images was developed by *Chen et al.* [10]. The registration accuracy was evaluated by two metrics - normalized cross-correlation (NCC) and normalized mutual information (NMI). Later, the same group [11] published an approach, where the SIFT features detection is replaced by finding the

center locations of the cones.

This paper describes an automated method for AO image registration into montage upon the method presented by *Chen et al.* [10] for AOSLO image montaging. The method is adapted for AO data, which in comparison with AOSLO images, have a larger field of view and therefore have lower resolution. Also, an AO image is acquired by a flood illumination camera, in contrast to an AOSLO image, which is raster scanned across the retina and thus its contrast varies more than the contrast of the AO image. The method results are evaluated by NMI metric.

2 MATERIALS AND METHODS

2.1 AO IMAGE ACQUISITION

The images were acquired by AO retinal camera with flood illumination (rtx1, Imagine Eyes, Orsay, France). The size of each retinal image is 1500×1500 pixels and field of view $4^\circ \times 4^\circ$. AO images were obtained from both eyes on 10 normal healthy subjects. From each of twenty eyes were acquired 10 images, which correspond to different retinal locations. One image is focused on the center of the fovea, afterward three images in the temporal direction from the fovea area and two images in each meridian (nasal, inferior, and superior) are acquired. The step between images was chosen to $2^\circ$for sufficient overlap of adjacent images (see an example of final montage in Figure 3 - b). Based on the acquisition approach, the preliminary position of each image is known, but is not sufficiently accurate for image montage, due to minor eye movements and unstable eye fixation (see Figure 2 - a).

2.2 PROBLEM FRAMEWORK AND PREPROCESSING

The AO image capturing the fovea area defines the origin of the coordinate system for each eye and all other images from an eye are registrated to that central image or to one of adjacent overlapping images, which is the closest to the central position and thus is already registered in the previous iteration.

The overlapping parts from reference x_{ref} and moving image x_{mov} are cropped and shown in Figure1. The images are characteristic by inhomogenous blur and also darker and lighter areas, which typically corresponds to vessel areas and adjacent light-reflective photoreceptor groups, respectively. The blur variability across images differs, as is shown in Figure1 - a,b, and thus complicates the registration. Hence the local blur effect in images is suppressed by an averaging filter with a window of size 15. In opposite, the darker and lighter areas are affected by blur barely and mutually correspond in images. Therefore the histogram equalization is used to highlight these areas in x_{ref} and x_{mov}.

In the next step the characteristic image features, similar for both images, are detected to find an alignment transformation between the x_{ref} and x_{mov}.

2.3 SCALE INVARIANT FEATURE TRANSFORMS (SIFT)

Correct and accurate transformation function determination is dependent on the detection of similar key locations as characteristic features in both images. The SIFT algorithm [12] is utilized for identifying these characteristic features, due to its advantage in robustness to illumination and local affine distortions. The features from x_{ref} and x_{mov} are matched to corresponding pairs and the Euclidean distance between two feature locations in each pair is computed. Obtained matching features are shown in Figure 1 - c,d as lines connecting the green points, which denotes the feature position. As is shown, numerous points are matched incorrectly. Thus the Euclidean distance of each match is rounded to tens and then the matches with Euclidean distance, which have the highest incidence, are considered as correct and used for computation of transformation matrix. The correct matches are denoted in Figure1 by red lines. In transformation is considered the influence of shift, rotation, and

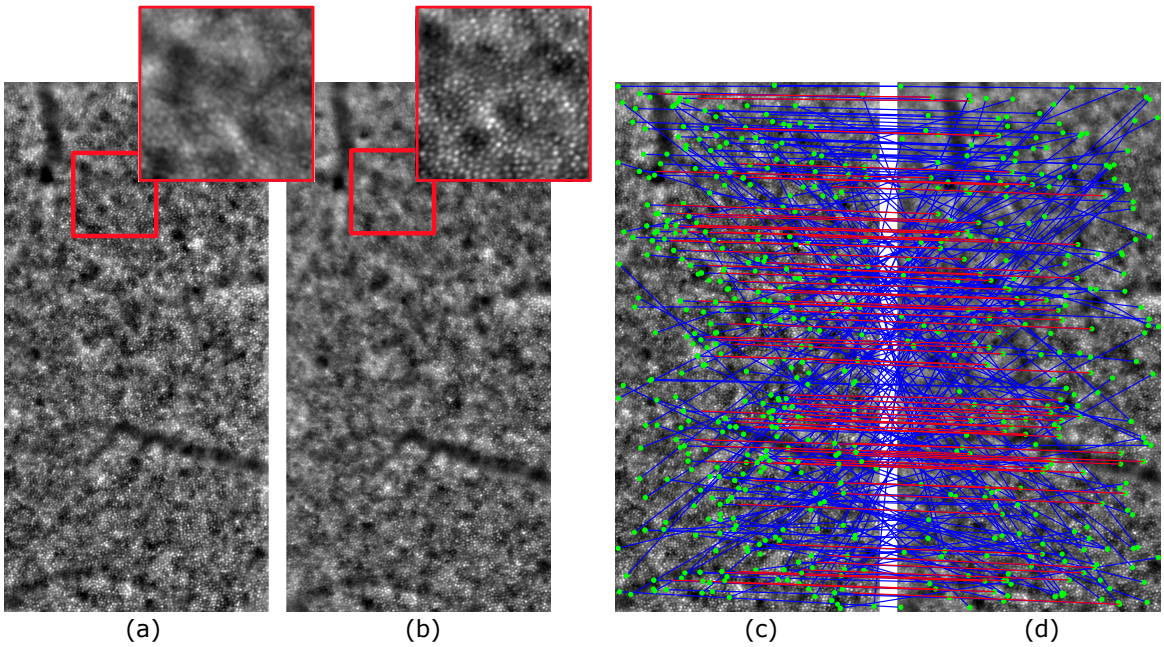


Figure 1: Overlapping parts from (a) reference imagex_{ref} and (b) moving image x_{mov} with zoom in areas, which highlight blur variability. Second set of images show detected features, marked by green points determining position and matched by blue lines in (c) x_{ref} and (d) x_{mov}. The red lines denotes correct matches used for transformation matrix computation

scaling. The moving image is transformed by a computed transformation matrix and thus is registered to the coordinate system.

2.4 ALIGNMENT ACCURACY

The registration accuracy assessment is based on image intensities in the overlapping regions. The normalized mutual information (NMI) [13] is used for image similarity evaluation. This metric is robust to systematic intensity variations, therefore is suitable for AO images. The measure is given by

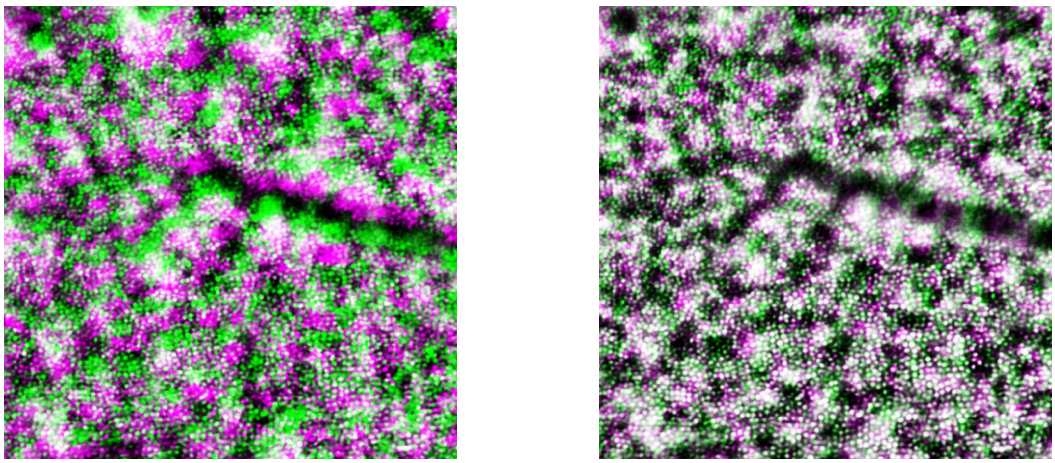
$$NMI = \frac{H(X)+H(Y)-H(X,Y)}{\sqrt{H(X)H(Y)}}, \tag{1}$$

where $H(X)$ and $H(Y)$ are computed from marginal and joint probability density of the possible image intensities in images X and Y. NMI ranges from 0 to 1, where 1 represents perfect alignment.

3 RESULTS AND DISCUSSION

Each of 9 of 10 images from an eye was registered (1 image capturing fovea is allocating origin of the coordinate system). The registration was performed by transformation matrix determined by matched SIFT points. The example of image matching is shown in Figure 2. The images shows x_{ref} and x_{mov} overlaid in different color bands. Gray regions in the composite image show where the two images have the same intensities. Magenta and green regions show where the intensities are different. Figure 2 - a shows the image pair before registration while Figure 2 -b shows the composite image after application of the proposed approach. In the registered image (Figure2 - b) photoreceptors from x_{ref} fit to photoreceptors in x_{mov} and the colored parts denoting the inter-image differences corresponds

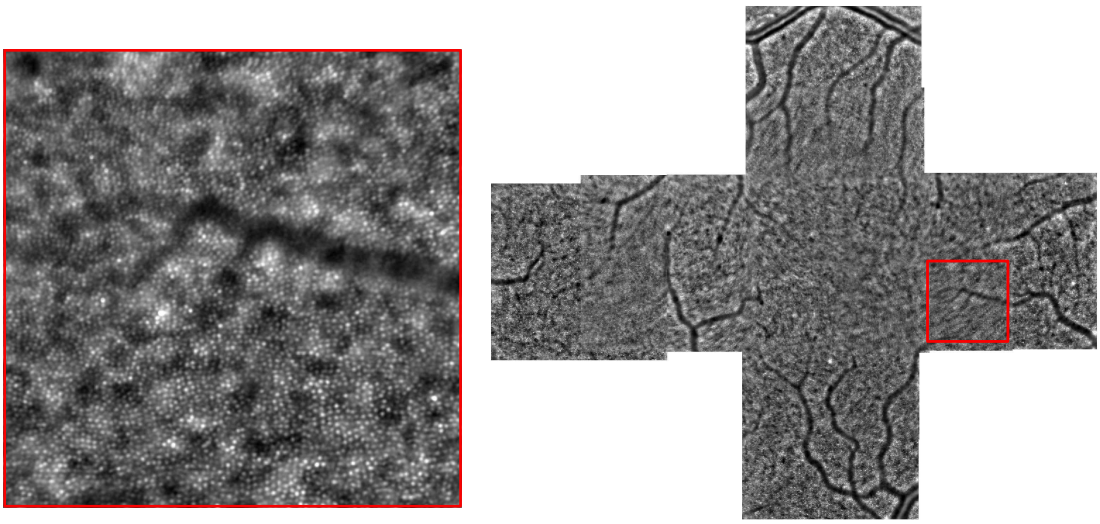
to intensity variability between images. The overlaid is averaged and shown in Figure 3 - a. The example of 10 image montage is shown in 3 - b.



(a) (b)

Figure 2: (a) Overlap of x_{ref} and x_{mov} on preliminary defined position without registration. (b) overlap of x_{ref} and x_{mov} after registration. Magneta and green regions denotes the different images in reference and moving image, respectively.

Chen et al. measured the registration accuracy via NMI. The measure on AOSLO data of normal healthy patients was in the range 0.062-0.075. Our computed NMI, evaluated over all overlapping regions, is 0.08 ± 0.02, which is in comparison with results in [10] exceedingly sufficient.



(a) (b)

Figure 3: (a) Averaged image from x_{ref} and x_{mov} marked in (b) resultant montage of 10 images from right eye.

4 CONCLUSION

An approach for fully automated AO image montaging is presented. It utilizes SIFT method applied on pre-processed images. This method is fast, but typically provides a lot of incorrect matches, which is its main drawback. Therefore, we proposed an efficient approach solving this problem by incorrectly matches elimination based on feature Euclidean distance highest occurrence. After the

incorrectly matched points elimination, the transformation matrix for each image is obtained and alignment is performed. From 200 images of ten patients were created 20 montage images. The efficiency over all overlapping regions is measured by NMI metric and is 0.08 ± 0.02.

REFERENCES

[1] J. S. Gill, M. Moosajee, and A. M. Dubis, "Cellular imaging of inherited retinal diseases using adaptive optics," *Eye*, vol. 33, no. 11, pp. 1683–1698, 2019.

[2] D. R. Williams, "Imaging single cells in the living retina," *Vision Research*, vol. 51, no. 13, pp. 1379–1396, 2011.

[3] M. Prasse, F. G. Rauscher, P. Wiedemann, A. Reichenbach, and M. Francke, "Optical properties of retinal tissue and the potential of adaptive optics to visualize retinal ganglion cells in vivo," *Cell and Tissue Research*, vol. 353, no. 2, pp. 269–278, 2013.

[4] M. Georgiou, A. Kalitzeos, E. J. Patterson, A. Dubra, J. Carroll, and M. Michaelides, "Adaptive optics imaging of inherited retinal diseases," *British Journal of Ophthalmology*, vol. 102, no. 8, pp. 1028–1035, 2018.

[5] M. Paques, S. Meimon, F. Rossant, D. Rosenbaum, S. Mrejen, F. Sennlaub, and K. Grieve, "Adaptive optics ophthalmoscopy," *Progress in Retinal and Eye Research*, vol. 66, no. 1, pp. 1–16, 2018.

[6] J. Carroll, M. Neitz, H. Hofer, J. Neitz, and D. R. Williams, "Functional photoreceptor loss revealed with adaptive optics," *Proceedings of the National Academy of Sciences*, vol. 101, no. 22, pp. 8461–8466, 2004.

[7] A. L. Chew, D. M. Sampson, I. Kashani, and F. K. Chen, "Agreement in cone density derived from gaze-directed single images versus wide-field montage using adaptive optics flood illumination ophthalmoscopy," *Translational Vision Science & Technology*, vol. 6, no. 6, pp. 1–13, 2017-11-01.

[8] B. Xue, S. S. Choi, N. Doble, and J. S. Werner, "Photoreceptor counting and montaging of en-face retinal images from an adaptive optics fundus camera," *Journal of the Optical Society of America A*, vol. 24, no. 5, pp. 1364–1372, 2007.

[9] H. Li, J. Lu, G. Shi, and Y. Zhang, "Automatic montage of retinal images in adaptive optics confocal scanning laser ophthalmoscope," *Optical Engineering*, vol. 51, no. 5, pp. 1–6, 2012-5-1.

[10] M. Chen, R. F. Cooper, G. K. Han, J. Gee, D. H. Brainard, and J. I. W. Morgan, "Multi-modal automatic montaging of adaptive optics retinal images," *Biomedical Optics Express*, vol. 7, no. 12, pp. 4899–4918, 2016.

[11] M. Chen, R. F. Cooper, J. C. Gee, D. H. Brainard, and J. I. W. Morgan, "Automatic longitudinal montaging of adaptive optics retinal images using constellation matching," *Biomedical Optics Express*, vol. 10, no. 12, pp. 6476–6496, 2019.

[12] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 2, 1999, pp. 1150–1157 vol.2.

[13] A. Strehl and J. Ghosh, "Cluster ensembles—a knowledge reuse framework for combining multiple partitions," *Journal of machine learning research*, vol. 3, no. Dec, pp. 583–617, 2002.