Geometry-Based Optic Disk Tracking in Retinal Fundus Videos

Anja Kürten[1], Thomas Köhler[1,2], Attila Budai[1,2], Ralf-Peter Tornow[3], Georg Michelson[2,3], Joachim Hornegger[1,2]

[1]Pattern Recognition Lab, FAU Erlangen-Nürnberg
[2]Erlangen Graduate School in Advanced Optical Technologies (SAOT)
[3]Department of Ophthalmology, FAU Erlangen-Nürnberg
anja-kuerten@web.de

Abstract. Fundus video cameras enable the acquisition of image sequences to analyze fast temporal changes on the human retina in a non-invasive manner. In this work, we propose a tracking-by-detection scheme for the optic disk to capture the human eye motion on-line during examination. Our approach exploits the elliptical shape of the optic disk. Therefore, we employ the fast radial symmetry transform for an efficient estimation of the disk center point in successive frames. Large eye movements due to saccades, motion of the head or blinking are detected automatically by a correlation analysis to guide the tracking procedure. In our experiments on real video data acquired by a low-cost video camera, the proposed method yields a hit rate of 98% with a normalized median accuracy of 4% of the disk diameter. The achieved frame rate of 28 frames per second enables a real-time application of our approach.

1 Introduction

Fundus photography is a widely used modality to diagnose diseases such as diabetic retinopathy or glaucoma by detection of anomalies in single 2-D images. Novel fundus video cameras enable the observation of fast temporal changes on the retina for new applications such as the measurement of the time course of the fundus reflection to analyze the cardiac cycle [1]. During examination, the human eye undergoes natural as well as guided movements. For applications such as patient guidance or the fixation of a certain position on the fundus by a stimulus, eye motion must be captured on-line as feedback in a control system. Tracking algorithms determine the position of an object in successive frames and can be used to capture eye motion. Therefore, the optic disk (OD) visible as bright elliptical spot, represents a robust feature for tracking.

While there are numerous approaches for OD localization, OD tracking remains challenging and has been less considered in literature. Difficulties arise due to the decreased image quality in terms of resolution and illumination conditions of common fundus video cameras compared to conventional fundus cameras [2]. Furthermore, tracking has to deal with large eye movements due to saccades, head motion or blinking during examination. In an early work, Koozekanani et

al. [3] have proposed a method to track the OD using a tiered detection scheme containing Hough transform, eigenimage analysis and geometric analysis. However, this approach is computationally demanding which makes a real-time application cumbersome. In terms of object detection methods, Loy and Zelinsky [4] have introduced the fast radial symmetry transform (FRST) as a point of interest detector to highlight points of high radial symmetry. Compared to the circular Hough transform, this method has a low computational complexity and thus it is well suited for real-time applications. Recently, the FRST has been proposed by Budai et al. [5] for OD localization in high-resolution fundus photographs.

In this paper, we propose a tracking algorithm that exploits the geometry of the OD to capture its radius and position in successive frames of a fundus image sequence. Therefore, our method employs the FRST in a tracking-by-detection scheme. Large eye motion due to saccades or head movements are detected automatically by a correlation analysis to guide the tracking procedure. The proposed method enables real-time OD tracking during examination.

2 Materials and Methods

2.1 Fast Radial Symmetry Transform

For every pixel $\boldsymbol{p}$ in the input image $\mathbf{I}$, the position of the center point candidate $\boldsymbol{c}(\boldsymbol{p})$ is given by $\boldsymbol{c}(\boldsymbol{p}) = \boldsymbol{p} + \left\lceil \frac{\nabla \mathbf{I}(\boldsymbol{p})}{\|\nabla \mathbf{I}(\boldsymbol{p})\|_2} \cdot r \right\rfloor$, where $\nabla \mathbf{I}(\boldsymbol{p})$ is the gradient of the input image and r is the radius. An orientation projection image $\mathbf{O}_r$ and a magnitude projection image $\mathbf{M}_r$ of the same size as $\mathbf{I}$ are initialized with zeros. Then, the values are incremented for every center point candidate according to:

$$\mathbf{O}_r(\boldsymbol{c}(\boldsymbol{p})) := \begin{cases} \mathbf{O}_r(\boldsymbol{c}(\boldsymbol{p})) + 1 & \text{if } \gamma \geq \nabla \mathbf{I}(\boldsymbol{p}) \geq \beta \\ \mathbf{O}_r(\boldsymbol{c}(\boldsymbol{p})) & \text{otherwise} \end{cases} \tag{1}$$

$$\mathbf{M}_r(\boldsymbol{c}(\boldsymbol{p})) := \begin{cases} \mathbf{M}_r(\boldsymbol{c}(\boldsymbol{p})) + \|\nabla \mathbf{I}(\boldsymbol{p})\|_2 & \text{if } \gamma \geq \nabla \mathbf{I}(\boldsymbol{p}) \geq \beta \\ \mathbf{M}_r(\boldsymbol{c}(\boldsymbol{p})) & \text{otherwise} \end{cases} \tag{2}$$

where $\mathbf{O}_r(\boldsymbol{c}(\boldsymbol{p}))$ and $\mathbf{M}_r(\boldsymbol{c}(\boldsymbol{p}))$ denote the orientation projection and magnitude projection image for center point candidate $\boldsymbol{c}(\boldsymbol{p})$, β is a gradient threshold parameter to exclude small gradients caused by noise and γ is a gradient threshold parameter to exclude high gradients that are likely to be caused by blood vessels [5]. The projection images are combined to a single map $\mathbf{F}_r(\boldsymbol{p})$ according to:

$$\mathbf{F}_r(\boldsymbol{p}) = \frac{\mathbf{M}_r(\boldsymbol{p})}{k_r} \left(\frac{|\tilde{\mathbf{O}}_r(\boldsymbol{p})|}{k_r} \right)^{\alpha}, \text{ where } \tilde{\mathbf{O}}_r(\boldsymbol{p}) = \begin{cases} \mathbf{O_r}(\boldsymbol{p}) & \text{if } \mathbf{O_r}(\boldsymbol{p}) < k_r \\ k_r & \text{otherwise,} \end{cases} \tag{3}$$

where α is the radial strictness parameter and k_r is a normalization factor. Finally, the radial symmetry contribution map $\mathbf{S}_r$ is computed as $\mathbf{S}_r = \mathbf{F}_r * \mathbf{A}_r$, where $\mathbf{A}_r$ is a two-dimensional Gaussian with standard deviation $\sigma = 0.25n$. Local maximums in $\mathbf{S}_r$ indicate bright regions of high radial symmetry and are detected for OD localization.

2.2 Tracking Algorithm

The proposed algorithm is divided into two steps: (i) During the initialization, the OD is localized and its radius is estimated. (ii) For tracking, the position of the OD is updated for every frame. In a preprocessing step for each frame, we apply a Gaussian filter for noise reduction and a dilation with subsequent erosion to suppress blood vessels that influence the detection of circular objects.

Radius Estimation. The radial symmetry contribution map $\mathbf{S}_r(\mathbf{I})$ is calculated for a set of radii $R = \{r_1, r_2, ...\}$ in image $\mathbf{I}$. We search for the maximum in all maps $\mathbf{S}_r(\mathbf{I})$ and estimate the OD radius $\hat{r}$ as:

$$\hat{r}(\mathbf{I}) = \arg\max_{r \in R}(\max(\mathbf{S}_r(\mathbf{I}))), \tag{4}$$

where $\max(\mathbf{S}_r(\mathbf{I}))$ is the maximum in the radial symmetry contribution map $\mathbf{S}_r(\mathbf{I})$. As the radius estimation in a single image may be affected by noise, we estimate the radius in a number of successive frames during initialization and compute the median radius $\tilde{r}$ over all estimates $\hat{r}$. In order to improve efficiency and since the OD size is fixed over time, this radius is used for tracking based on the FRST in all subsequent frames.

Position Update. The OD position is detected by computing the FRST map $\mathbf{S}_{\tilde{r}}(\mathbf{I})$ for the current frame based on the radius estimate $\tilde{r}$. Then, the center point is updated by detecting the position of the maximum value $\max(\mathbf{S}_{\tilde{r}}(\mathbf{I}))$. The efficiency of the position update is improved by computing $\mathbf{S}_{\tilde{r}}(\mathbf{I})$ only in a quadratic window centered on the OD position of the previous frame with an edge length of l times the OD diameter. In case of large eye movements due to saccades, the OD leaves the window and its new position cannot be detected relative to the previous frame. Therefore, we compute the normalized cross-correlation ρ between optic disk windows in successive frames. For $\rho < \rho_0$, we detect a saccade and update the OD position using the FRST in the whole image. Additionally, frames affected by blinking are excluded from tracking. We compute the Bhattacharyya coefficient [6] b between the histogram of the first image and the histogram of the current frame to exclude frames with $b < b_0$.

2.3 Experiments

We evaluated our algorithm on a set of fundus image sequences acquired with a low-cost video camera that was constructed by Dr. Ralf-Peter Tornow. The camera uses LED illumination and a CCD camera (640 x 480 pixels) to acquire video sequences with up to 50 frames per second (fps) and a field of view of 15° without dilating the pupil. We captured image sequences from the left eye of six healthy subjects under two scenarios: (i) The subject fixated on a bright spot, so that the fixation position was stable. (ii) The subject fixated on different points in order to produce saccades. After three seconds of fixating on the starting point,

	Hit rate [%]		Mean error [ODD]		Median error [ODD]	
	A	B	A	B	A	B
TbD (proposed)	98.0	98.9	0.13 ± 0.21	0.12 ± 0.18	0.03	0.04
MST	86.9	88.9	0.35 ± 0.43	0.32 ± 0.41	0.20	0.19
Expert A				0.03 ± 0.02		0.03
Expert B			0.03 ± 0.02		0.03	

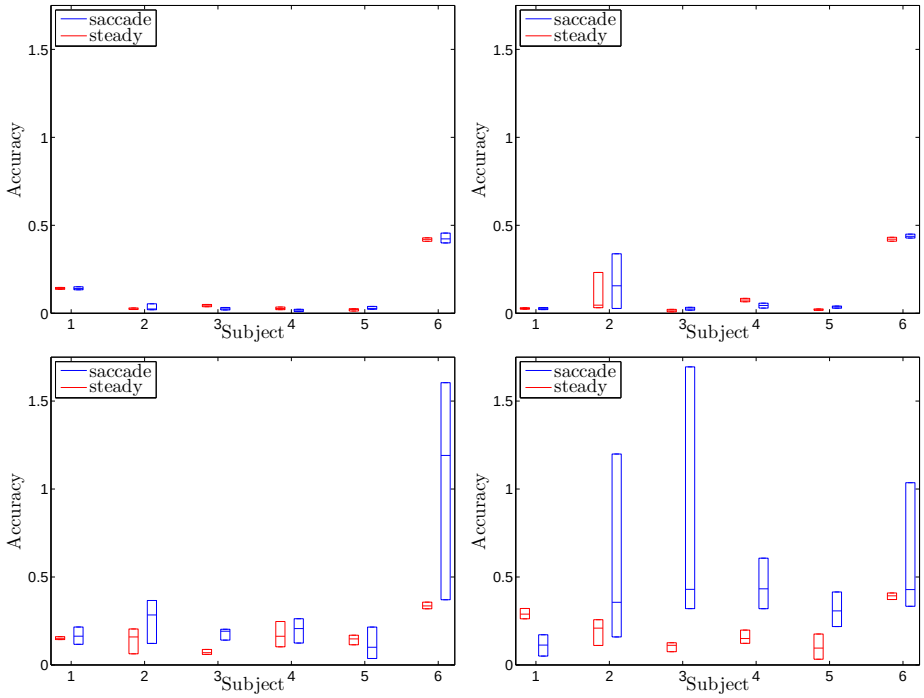
Table 1. Hit rate and accuracy measured in OD diameter (ODD) of the proposed tracking-by-detection algorithm (TbD) and mean shift tracking (MST). We also evaluated the inter-observer variance between expert A and expert B.

the subject alternated between looking to the right, fixating on the starting point and looking to the left. Each sequence has a duration of 15 seconds and was taken with two different frame rates: 25 and 50 fps. Two ophthalmic imaging experts created ground truth data by labeling the OD center in every 15^{th} frame and measuring the vertical OD diameter in the first frame. The data from one subject was used for parameter adjustment and was excluded from the results. The range of the radii was set to $R = \{90, 95, \ldots, 120\}$. Six frames were used for initialization. The edge length of the OD window was chosen as 1.25 times the OD diameter. The saccade threshold was set to $\rho_0 = 0.25$ and the blink threshold was set to $b_0 = 0.7$. For the FRST, the radial strictness parameter was selected to be $\alpha = 2$ and the gradient thresholds were set to $\beta = 10$ and $\gamma = 80$.

Our method was compared to intensity-based mean shift tracking (MST) [6]. All algorithms were implemented in C++ and runtimes were evaluated on an Intel i3 CPU with 2.4 GHz and 4 GB RAM. Robustness was assessed by computing the hit rate for each algorithm. A hit is detected if the estimated OD center is inside the labeled OD. The accuracy of tracking was evaluated by computing the euclidean distance between the detected and the labeled center.

3 Results

The accuracy achieved by mean shift tracking and the proposed method itemized for the examined subjects is shown as boxplot in Fig. 1. The measuring unit is OD diameter (ODD), which is the euclidean error of the center point detection normalized with the corresponding ground truth OD diameter. For both algorithms, tracking was more reliable for video data acquired with a lower frame rate (25 fps) and thus improved conditions in terms of illumination and contrast. Mean shift tracking showed higher errors for saccade sequences. Tab. 1 summarizes the hit rates as well as mean, standard deviation and median of the accuracy over all test sequences. Our algorithm achieved a mean hit rate of 98.0%, a mean accuracy of 0.13 ± 0.21 ODD and a median accuracy of 0.04 ODD which is in the range of the median inter-observer variance of 0.03 ODD. Mean shift tracking achieved a mean hit rate of 86.9% with a mean accuracy of 0.35 ± 0.43 ODD and a median accuracy of 0.20 ODD. A qualitative comparison be-



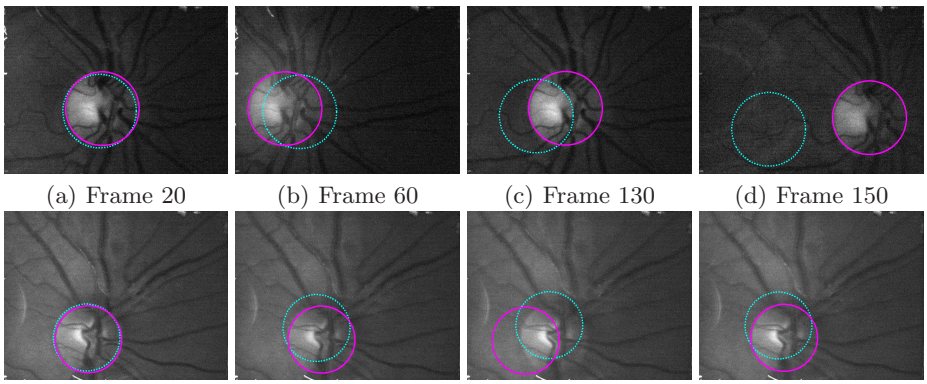
(a) 25 fps (b) 50 fps

Fig. 1. Boxplots of the tracking accuracy for the proposed method (top row) and mean shift tracking (bottom row) normalized by the OD diameter (ODD) for 25 fps (a) and 50 fps (b). We evaluated the euclidean distance between the estimated OD center point and the mean of two experts used as ground truth.

tween both methods for two different subjects is presented in Fig. 2. In particular, the proposed method was able to detect the OD even in case of saccades whereas mean shift resulted in erroneous center point detections. In terms of run-time, we achieved a mean frame rate of 28 fps. The mean shift approach achieved a frame rate of 200 fps. We used the same initialization to estimate the OD radius for all methods, which had a mean runtime of 1.9 seconds.

4 Discussion

We proposed a tracking-by-detection framework to capture the position and the radius of the OD in successive frames of retinal fundus image sequences. The FRST is employed for tracking and exploits the circular shape of the OD. Compared to intensity-based mean shift tracking, the proposed method shows improved robustness with respect to large eye movements due to saccades. Our proof-of-concept implementation enables a real-time application for 25 fps video data. A GPU implementation holds great potential to speed-up our algorithm.



(a) Frame 20 (b) Frame 60 (c) Frame 130 (d) Frame 150

(e) Frame 20 (f) Frame 90 (g) Frame 305 (h) Frame 360

Fig. 2. Tracking results for different frames obtained by mean shift tracking (dotted blue line) and the proposed method (magenta line).

In our approach, the OD detection and position update may be affected by blood vessels which have a shape and gradient similar to the OD boundary as shown in Fig. 2(g). In this case, the inclusion of vascular information as proposed by [3] could improve the robustness. Future work will focus on a hybrid approach to integrate color and texture information to the proposed algorithm in order to improve tracking reliability.

Acknowledgments. The authors gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German National Science Foundation (DFG) in the framework of the excellence initiative.

References

1. Tornow RP, Kopp O, Schultheiss B. Time Course of Fundus Reflection Changes According to the Cardiac Cycle. Invest Ophthalmol Vis Sci. 2003;44:Abstract 1296.
2. Köhler T, Hornegger J, Mayer M, Michelson G. Quality-Guided Denoising for Low-Cost Fundus Imaging. In: Proc BVM 2012. Springer; 2012. p. 292–297.
3. Koozekanani D, Boyer KL, Roberts C. Tracking the optic nervehead in oct video using dual eigenspaces and an adaptive vascular distribution model. IEEE Trans Med Imaging. 2003;22(12):1519–1536.
4. Loy G, Zelinsky A. Fast radial symmetry for detecting points of interest. IEEE Trans Pattern Anal Mach Intell. 2003;25(8):959–973.
5. Budai A, Aichert A, Vymazal B, Hornegger J, Michelson G. Optic disk localization using fast radial symmetry transform. In: IEEE 26th International Symposium on Computer-Based Medical Systems (CBMS); 2013. p. 59–64.
6. Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking. IEEE Trans Pattern Anal Mach Intell. 2003;25(5):564–577.