

# Journal of Biomedical Optics

SPIEDigitalLibrary.org/jbo

## **Speed quantification and tracking of moving objects in adaptive optics scanning laser ophthalmoscopy**

Johnny Tam  
Austin Roorda

# Speed quantification and tracking of moving objects in adaptive optics scanning laser ophthalmoscopy

Johnny Tam<sup>a</sup> and Austin Roorda<sup>a,b</sup>

<sup>a</sup>University of California, Berkeley and University of California, San Francisco, Joint Graduate Group in Bioengineering, Berkeley, California 94720

<sup>b</sup>University of California, Berkeley, School of Optometry, Berkeley, California 94720

**Abstract.** Microscopic features of the human retina can be resolved noninvasively using an adaptive optics scanning laser ophthalmoscope (AOSLO). We describe an improved method to track and quantify the speed of moving objects in AOSLO videos, which is necessary for characterizing the hemodynamics of retinal capillaries. During video acquisition, the objects of interest are in constant motion relative to the background tissue (object motion). The background tissue is in constant motion relative to the AOSLO, due to continuous eye motion during video recordings (eye motion). The location at which AOSLO acquires data is also in continuous motion, since the imaging source is swept in a raster scan across the retina (raster scanning). We show that it is important to take into consideration the combination of object motion, eye motion, and raster scanning for accurate quantification of object speeds. The proposed methods performed well on both experimental AOSLO videos as well as synthetic videos generated by a virtual AOSLO. These methods improve the accuracy of methods to investigate hemodynamics using AOSLO imaging. © 2011 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.3548880]

Keywords: spatiotemporal plots; tracking; raster scan; eye motion; motion contrast; adaptive optics.

Paper 10582R received Oct. 26, 2010; revised manuscript received Jan. 5, 2011; accepted for publication Jan. 6, 2011; published online Mar. 11, 2011.

## 1 Introduction

The eye is a window through which the vasculature can be directly observed. Recently, the adaptive optics scanning laser ophthalmoscope (AOSLO) has made it possible to directly acquire videos of leukocyte movement through the smallest capillaries in the human eye, without the use of injected dyes.<sup>1</sup> AOSLO is a research-grade instrument that can take images at higher resolution and contrast compared to what is clinically available and is based on adaptive optics technologies, which correct for optical aberrations of the eye.<sup>2</sup> The AOSLO design has been described in detail elsewhere<sup>3,4</sup> and has been used to study phenomenon such as the relationship between photoreceptor function and visual receptive fields<sup>5</sup> and interpretation of static vascular features.<sup>6</sup> However, there are key issues that must be addressed to make AOSLO an effective system for the study of hemodynamics.

Quantification of object speeds in AOSLO videos is important for hemodynamics, but is confounded by raster scanning, eye motion, and object motion. Since videos are acquired using a raster scanning system, different pixels within a video frame are acquired at different points in time. This affects the appearance of moving objects – the apparent object speed is dependent on the speed of the raster scan (Fig. 1). The magnitude of the error in measured speed due to raster scanning depends on the configuration of AOSLO imaging parameters, but can be as large as 37.8% (Sec. 2.6). There is also constant eye motion that occurs during acquisition of video frames.<sup>7,8</sup> Since the raster scan continuously scans in a fixed pattern, this results in unique distortions in each video frame.<sup>9,10</sup> Finally, the object itself is

also in constant motion, simultaneous to raster scanning and eye motion. The motions of the object, eye, and raster scan must be considered simultaneously for accurate quantification of object speeds in an AOSLO system.

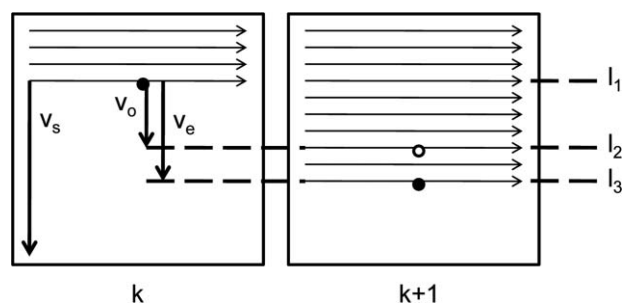
In this paper, we describe methods for tracking and accurate speed quantification of moving objects in AOSLO videos. We use spatiotemporal (ST) plot analysis and motion contrast enhancement to track moving objects and measure apparent object speeds. Apparent object speeds are then corrected using a slope modification method to correct for errors introduced by eye motion and raster scanning. The accuracy of the proposed methods is validated using synthetic data sets generated by a virtual AOSLO.

## 2 Materials and Methods

### 2.1 AOSLO Imaging

Videos were acquired on an AOSLO as previously described.<sup>1</sup> AOSLO videos can be acquired using different imaging configurations, depending on the application. We consider three different configurations of imaging parameters (Type 1, 2, and 3), as described in Table 1. Type 1 videos were acquired for the purpose of analyzing blood flow, Type 2 for analyzing photoreceptors, and Type 3 for analyzing both blood flow and photoreceptors. Type 1 videos were acquired using a green wavelength laser, which has been reported to be the optimal wavelength for obtaining good contrast in blood flow imaging.<sup>11</sup> Type 2 and 3 videos were acquired using a near-IR laser — a more desirable wavelength in terms of (i) risks due to laser damage

Address all correspondence to: Johnny Tam, University of California, Berkeley, School of Optometry, Room 485 Minor Hall, Berkeley, CA 94720-2020. Tel: 510 642 2825; Fax: 510 643 5109; E-mail: johnny@berkeley.edu.



**Fig. 1** Raster scan problem formulation showing speed overestimation for a downward moving object from frame  $k$  to  $k+1$ . The vertical speed of the scanner is  $v_s$ . In frame  $k$ , the object (solid black circle) is at line  $l_1$ . After exactly one frame, when the scanner has reached line  $l_1$  on frame  $k+1$ , the object is at line  $l_2$ . However, the scanner catches the object at line  $l_3$ . The object speed,  $v_o$ , is overestimated as  $v_e$ .

on the retina, (ii) compatibility with the AOSLO hardware, and (iii) overall subject comfort. Type 1 and 3 videos were acquired for longer durations in order to increase the number of leukocytes that could be counted. Examples of frames from Type 1, 2, and 3 videos are shown in Fig. 2.

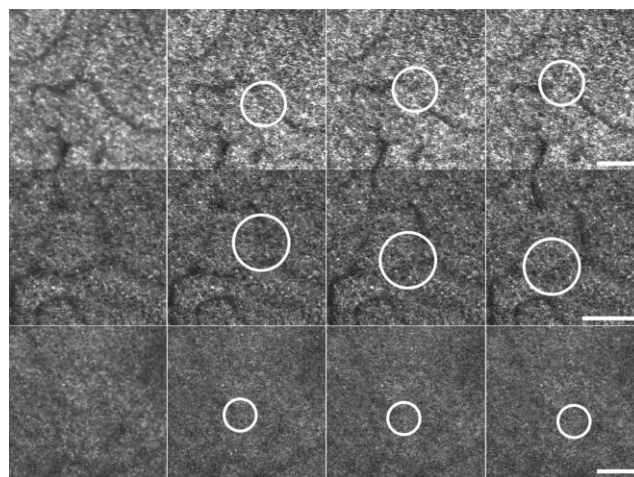
Since the Type 1 videos had the highest spatial contrast for the moving blood cells, we developed our methods using only Type 1 videos, and used Type 2 and 3 videos as well as synthetic videos generated by a virtual AOSLO for verification.

## 2.2 Preprocessing

Raw videos were preprocessed to correct for distortions due to raster scanning and eye motion, without considering object mo-

**Table 1** Imaging parameters and subject data for various types of AOSLO videos. Three representative configurations of imaging parameters used for AOSLO imaging, for three different subjects. Representative values are given for the scale factors, which vary by small amounts across different imaging sessions due to small variations in hardware alignment and eye morphology.

Specification	Type 1	Type 2	Type 3
Imaging wavelength [nm]	532	840	840
Frame rate [Hz]	30	30	60
Raw video frame size [pixels <sup>2</sup> ]	525×512	512×512	512×525
Field of view (approx.) [deg <sup>2</sup> ]	1.5×1.5	1.2×1.2	1.5×1.5
X scale factor [pixels/deg]	342	414	328
Y scale factor [pixels/deg]	342	409	330
Length of video [seconds]	40	2–10	40
Retinal scale factor [mm/deg]	0.28008	0.28697	0.28889
Subject age	37	26	24
Refractive error, sphere [D]	+1.0	−1.0	+0.5
Refractive error, cylinder [D]	−0.25	0.0	0.0



**Fig. 2** Examples of retinal images noninvasively acquired in human subjects using an AOSLO, for a Type 1 (top row), Type 2 (middle row), and Type 3 (bottom row) video. The first column shows the averaged image of all frames for each video. The remaining three columns are three consecutive frames, showing a leukocyte (circled) traveling through a parafoveal retinal capillary. The frames in these videos have been preprocessed. The foveal center is near the top right corner for all three videos. Circular dots are photoreceptors, and dark fuzzy lines usually correspond to capillaries. However, the locations of all capillaries are not obvious. Scale bar, 100  $\mu\text{m}$ .

tion. Preprocessing involves desinusoiding, stabilization, cropping, and frame deletion.

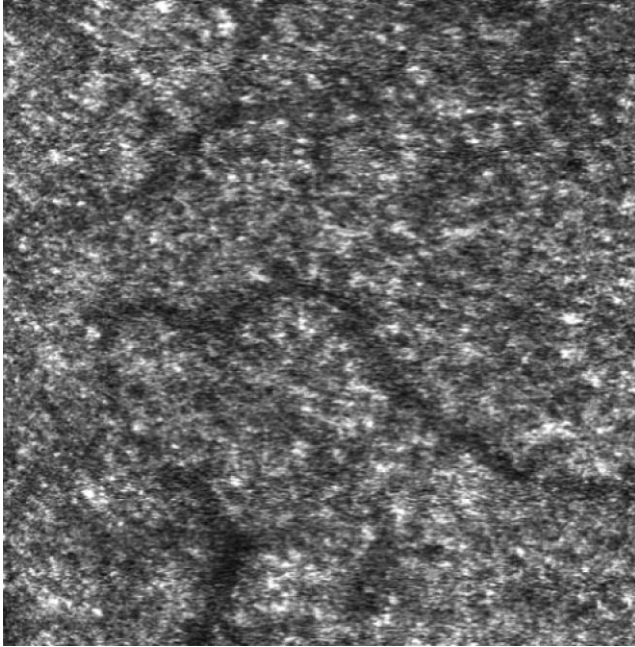
### 2.2.1 Desinusoiding and stabilization

To achieve high line density and high frame rates, the AOSLO employs a resonant scanner combined with a sensor that reads in data at a constant rate. The velocity of the scanner varies sinusoidally across each scan line, which results in a horizontal distortion in the raw videos. Desinusoiding corrects this distortion, which is characterized from videos of calibration grids. The velocity of the scanner is slowest at the left and right edges of the frame and fastest in the middle; thus, there are more pixels per retinal area toward the edges compared to the center. The redistribution of pixels can result in a desinusoiding artifact due to a change in the distribution of noise. We minimized this artifact using median and Gaussian filtering (Sec. 2.3).

Stabilization is the process that corrects for the distortions due to eye motion that occur during acquisition of each raster-scanned frame. Detailed procedures for desinusoiding and stabilization can be found elsewhere.<sup>9,10</sup> Briefly, the task involves splitting each frame in a video into a set of horizontal strips, each of which is registered using affine transformations to a desinusoided reference frame and reassembled using linear interpolation. The result is a desinusoided and stabilized video as well as a high-frequency eye motion trace. This trace is important for accurate calculation of the distance that a moving object has traveled (Sec. 2.5).

### 2.2.2 Cropping the video

Due to eye motion, there are regions of the retina that are not present in all video frames, particularly at the edges of each frame. To account for this, the desinusoiding and stabilization



**Video 1** Type 1 AOSLO video, slowed to 5 fps (real time frame rate is 30 fps). An example of an improper frame resulting in poor registration is at 0:03 s into the video. (MPEG, 3.73 MB). [URL: <http://dx.doi.org/10.1117/1.3548880.1>]

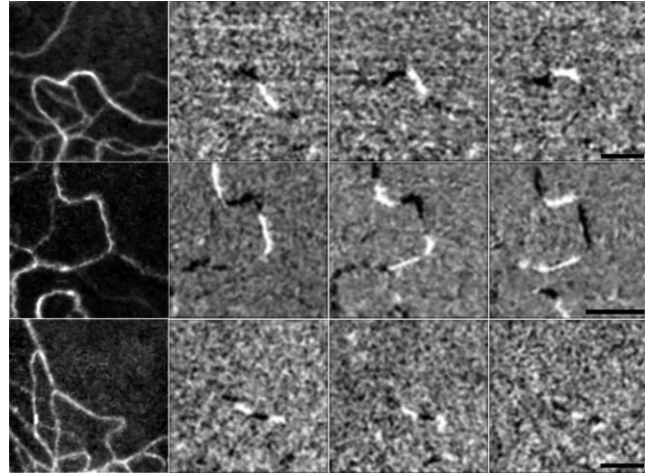
process introduces borders around each frame so that each registered frame will be of the same size. The thickness of each border changes according to the eye motion. We crop the video such that each frame contains only the portion of the video that was visible in the majority of all frames, thereby eliminating the black borders. The number of lines at the top edge that were removed due to cropping were stored in a lookup table and used in the calculation of object speeds (Sec. 2.5).

### 2.2.3 Frame deletion

In the processed videos, there were three types of improper frames that were identified for deletion. First, insufficient overlap between the image and the reference frame resulted in poor stabilization. This occurred when the eye wandered too far away from its fixation target. Second, blinks resulted in the image intensity dropping to zero throughout the blink. Third, large saccades, involuntary fast eye movements, caused intraframe shearing and distortion on single frames and prevented proper image stabilization (Video 1). To generate high quality images of photoreceptors and vessels, these “improper” frames were deleted. However, we did not delete any frames for the speed analysis since deletion of frames would increase the apparent speed of a moving object.

## 2.3 Visualization of Moving Objects and Vessels

Since spatial contrast is low, motion contrast enhancement is used to visualize moving objects and vessels (Fig. 3). Methods for motion contrast enhancement have been previously described.<sup>6,12–14</sup> We implement a method that works well with AOSLO videos,<sup>6</sup> using a multiframe division video and a standard deviation image. The multiframe division videos were used



**Fig. 3** Visualization of vessels and moving objects by applying motion contrast enhancement to the Type 1 (top row), Type 2 (middle row), and Type 3 (bottom row) AOSLO videos shown in Fig. 2. The first column shows the standard deviation image, which enhances the contrast of vessels. The next three columns show three consecutive frames of the multiframe division video, which enhances the contrast of moving objects. Scale bars, 100  $\mu\text{m}$ .

to visualize moving objects and the standard deviation image was used to visualize vessels. Median and Gaussian filtering were applied before and after calculation of the multiframe division video, respectively.

A preprocessed video has moving blood cells in front of a stationary background tissue, consisting of photoreceptors and vessels. Given two frames,  $I_{j(x,y)}$  and  $I_{j+1(x,y)}$ , the division image  $D_j(x,y) = \frac{I_j(x,y)}{I_{j+1(x,y)}}$  emphasizes the objects in motion as long as the intensity of background tissue remains relatively constant. Here,  $I_{j(x,y)}$  represents the intensities of frame  $j$  at position  $(x,y)$ . Division images are used instead of difference images to enable arithmetic averaging of multiple frames, which improves the signal to noise ratio, as opposed to using the arithmetic average of two consecutive difference images, which yields no improvement in signal to noise.<sup>15</sup> We defined a multiframe division video as  $M_j(x,y) = \frac{D_j(x,y)+D_{j+1(x,y)}}{2}$ .

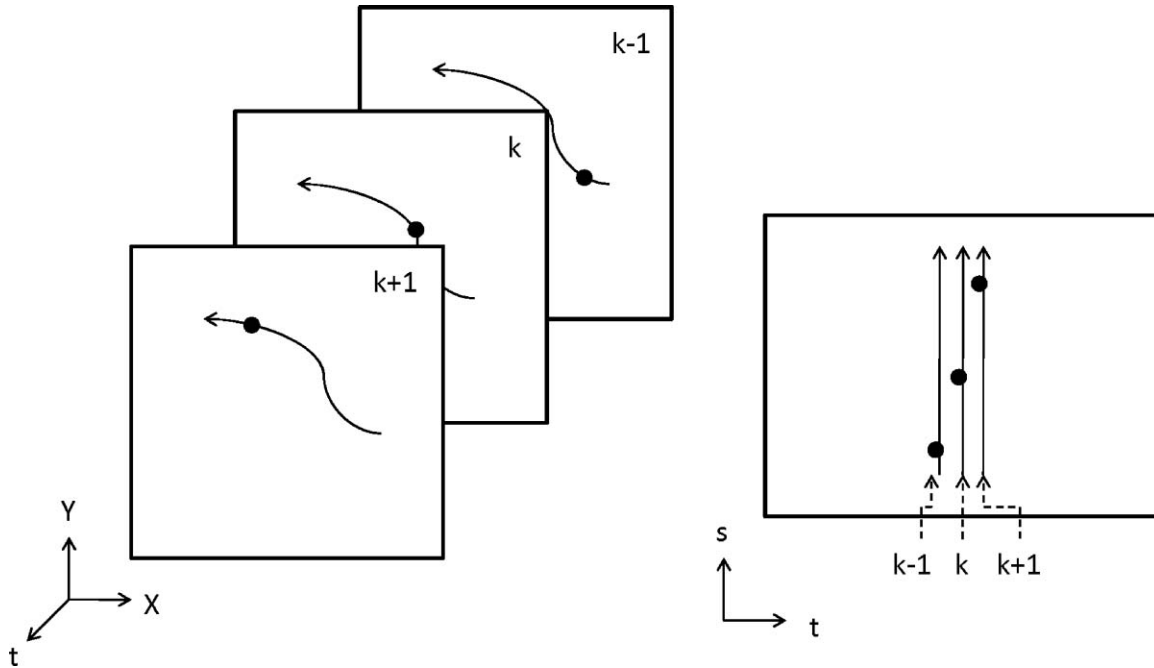
To visualize the perfused vessels, an image was calculated from the multiframe division video, using the geometric standard deviation. For a video with  $n$  frames, the geometric standard deviation image,  $S(x,y)$ , is defined as

$$S(x,y) = \exp \left[ \sqrt{\frac{\sum_{j=1}^n [\ln M_j(x,y) - \ln \bar{M}(x,y)]^2}{n-1}} \right], \quad (1)$$

$$\bar{M}(x,y) = \sqrt[n]{\prod_{j=1}^n M_j}. \quad (2)$$

## 2.4 Object Tracking

We used cell paths on ST plots for object tracking.<sup>15</sup> ST plots are a method to visualize hemodynamics and offer two major advantages. First, the ST plot representation is more compact, which assists in pattern identification. Variables such as density, frequency, and variations in speed both spatially and temporally



**Fig. 4** Conversion of  $X$ - $Y$ - $t$  coordinate system (left) into  $s$ - $t$  coordinate system (right). Three consecutive frames,  $(k - 1, k, k + 1)$ , are shown with a dark circle representing a single leukocyte traveling along a vessel centerline.

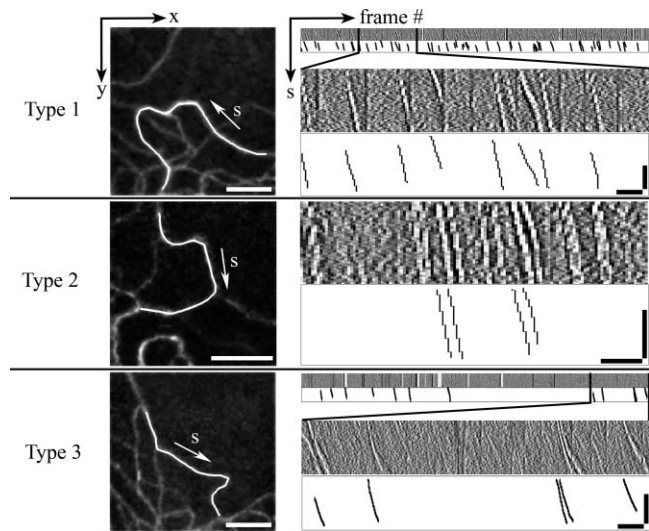
can easily be observed on ST plots, but not from direct examination of the original video. Second, the dimensional complexity of the problem is reduced from a 4-dimensional (3D+1T) problem to a 3-dimensional (2D+T) problem, which minimizes the computation cost. These plots have been used for many systems,<sup>12, 16-18</sup> including AOSLO systems.<sup>15, 19</sup>

ST plots were generated by converting an  $X$ - $Y$ - $T$  coordinate system into an  $s$ - $T$  coordinate system (Fig. 4). Consider an arbitrary vessel in a sequence of frames, with an object that moves along the trajectory of the vessel, given by  $f(x,y)$ . By plotting intensity values along the vessel, and discarding all other pixel values, a two-dimensional plot can be generated that shows the movement of individual objects traveling through a one-dimensional line, given by  $f(s)$ , where  $f(x,y) \leftrightarrow f(s)$  is naturally defined, with the first coordinate of  $f(x,y)$  mapping to the first element of  $f(s)$ , the second coordinate to the second element, and likewise for the remaining elements. Since we are exactly specifying the mapping from each pixel in  $X$ - $Y$ - $T$  space to  $s$ - $T$  space, the mapping is invertible.

For single-file flow through capillaries seen by AOSLO videos, there is no loss in speed information when switching from the  $X$ - $Y$ - $T$  representation to the  $s$ - $T$  representation.<sup>15</sup>

Motion contrast enhancement improves the ST plots by increasing the accuracy of vessel centerline extraction and by increasing the contrast of cell paths.<sup>15</sup> Using motion contrast enhanced ST plots, we manually extracted cell traces for the tracking and speed quantification. To identify traces, the user was presented with a graphical user interface showing a portion of a ST plot. The user identified traces by selecting points along that trace. For consistency, points were selected at the border between the dark and bright regions of the trace, on the leading edge (Fig. 5). After points were selected, interpolation was performed using piecewise splines constrained to the pixel resolution.

For tracking, the coordinates of each extracted trace were used to register the location of the blood cells in the video. ST coordinates  $(s, t)$  were converted back to video coordinates  $(x, y, t)$  using the invertible mapping defined during generation



**Fig. 5** Generation of motion contrast enhanced ST plots from AOSLO videos using the same videos shown in Fig. 2. The first column shows the vessel centerline that was selected for analysis. The second column shows ST plot analysis. On the ST plots, diagonal streaks represent moving objects, while vertical streaks represent blinks or saccades, where the video intensity drops for 1-3 frames. Extracted traces are directly shown beneath the corresponding ST plot. Type 1 video (top), all frames (1 to 1200), and frames 201 to 400 magnified from entire strip. Type 2 video (middle), all frames (1 to 126). Type 3 video (bottom), all frames (1 to 2387), and frames 1987 to 2387 magnified from entire strip. Vessel image scale bars, 100  $\mu$ m; ST plot horizontal scale bars, 0.5 s, vertical scale bars, 0.25 mm.

of ST plots. Video coordinates were compiled into a list and then used to mark object locations to visualize the tracking results.

## 2.5 Quantification of Object Speeds

In the absence of eye motion, the speed of an object in a raster scanning system can be explicitly computed using line information from pairs of frames.<sup>20</sup> The correction is based on computing the actual time at which a given line was acquired, as opposed to assuming that the entire frame was acquired at the same time. The true time,  $t$ , can be computed as

$$t = T_f \left( 1 + \frac{l_2 - l_1}{N_l} \right), \quad (3)$$

where  $T_f$  is the time per field,  $l_1$  and  $l_2$  the scan lines of the object center in the first and second frames, and  $N_l$  the number of scan lines per frame. However, this approach [Eq. (3)] must be modified for AOSLO videos since the effect of the raster scan is confounded with eye motion and desinusoidal.

In a nonraster scanning system without eye motion, the speed of an object can be computed by simply computing the slope of the extracted trace from an ST plot. However, in our system, the slope of the trace gives speeds in time units of frames. We present a slope modification procedure to correct speeds, based on computing the acquisition time coordinates in the extracted traces. In order to perform the correction, line numbers on preprocessed videos need to be transformed back to line numbers on the raw videos (not preprocessed). This is important because the correction for intraframe eye motion results in local stretching or compression of pixels, thereby altering line numbers.

The AOSLO uses a fast horizontal scan and a slower vertical scan, from left to right and top to bottom directions, respectively. The main component of the error is due to the slower vertical scan; the error due to the horizontal scan is small and does not need to be corrected (Secs. 2.6 and 3.1). As an example, in the absence of eye motion, a downward moving object will have a larger observed displacement compared to the actual displacement, since the scan is chasing a moving target. More generally, any object that is moving in a nonhorizontal trajectory has a vertical component of velocity that needs to be corrected. If there is eye motion, then the actual displacement is also dependent on the amount that the eye has moved.

Consider coordinates from the extracted traces, given as (frame number,  $s$ ). The acquisition time for each line (in units of partial frames) can be computed as

$$\text{acquisition time} = \text{frame number} + \frac{L}{512}. \quad (4)$$

$L$  is the line number at which the data was taken on the raw video, and can be recovered in the following manner:

1. Recover the line number,  $L_{\text{crop}}$ , of the object on the cropped video by determining the  $y$  coordinate from the inverse transformation  $s \rightarrow (x, y)$ .
2. Correct the line number for cropping by adding back the number of lines at the top of the image that were removed during cropping,  $dL_{\text{crop}}$ , stored in cropping lines (Sec. 2.2.2).

3. Correct for eye motion by applying the inverse transformation,  $S^{-1}$ , from raw to stabilized videos.  $S^{-1}$  was stored during preprocessing in the eye motion trace (Sec. 2.2.1).

Compute  $L = S^{-1}(L_{\text{crop}} + dL_{\text{crop}})$ . The extracted traces are then plotted as (acquisition time,  $s$ ).

For each corrected trace, a linear regression was applied and the slope of the line, with units of pixels/frame, was used to compute the speed of the leukocyte (in units of mm/s) through the selected vessel segment in the following manner:

$$\text{slope}^* \left( \frac{2 * \text{frame rate} * \frac{\text{mm}}{\text{deg}}}{X \text{ scale factor} + Y \text{ scale factor}} \right) = \text{speed in } \frac{\text{mm}}{\text{s}}. \quad (5)$$

Sample conversion parameters for Type 1, 2, and 3 videos are summarized in Table 1. The mm/deg conversion factor on the retina was estimated as previously described.<sup>6,21</sup>

## 2.6 Expected Error due to Raster Scanning (RS)

The raster scan error is significant for AOSLO videos. In this section, we develop a theoretical model to quantify the magnitude of the raster scan error. To understand the nature of the expected raster scan error, consider the case of a vertically oriented vessel with a downward-moving object that starts at the top of the image (Fig. 1). Assuming that there is no eye motion, we derive the expected raster scan error in the vertical and horizontal cases for comparison to actual measured error rates and show that it is significant in the vertical direction, but not the horizontal direction.

We introduce the dimensionless number,

$$RS = \frac{v_o}{v_s}, \quad (6)$$

where  $v_o$  is the speed of the object and  $v_s$  is the speed of the scanning line.

If  $RS > 1$ , then the system is unable to image the object and the error becomes infinite. When  $RS = 0.5$ , the error is exactly 100%. When  $v_o \ll v_s$ ,  $RS \rightarrow 0$ , and the raster scan error is negligible. By convention, leukocyte speeds on the retina are reported in mm/s. For an object speed given in mm/s,  $\bar{v}_o$ , with scan speed,  $v_s$ , given in pixels/frame, with scale factor as defined in Table 1,  $RS$  can be calculated as

$$RS = \frac{\text{scale factor in scan direction}}{\text{frame rate} * \text{mm/deg on the retina}} * \frac{\bar{v}_o}{v_s}. \quad (7)$$

When the object is moving in the same direction as the raster scan, the speed of the object as measured by the system,  $v_e$ , will be overestimated. The percent error due to raster scanning is given by

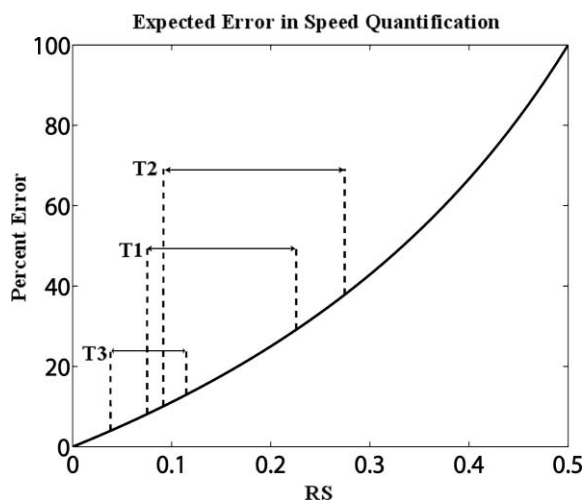
$$\text{percent error} = \frac{v_e - v_o}{v_o}. \quad (8)$$

Assuming the object and scan have speeds of  $v_o$  and  $v_s$ , the relationship between object speed and measured speed is

$$v_e = \frac{v_o}{1 - RS}. \quad (9)$$

The percent error due to the raster scan is

$$\text{percent error} = 100\% * \frac{RS}{1 - RS}. \quad (10)$$



**Fig. 6** Expected raster scan error assuming no eye motion, for a vessel that is oriented in the same direction as the scanner. The range of errors for an object moving between 1 and 3 mm/s is shown for a Type 1 (T1), Type 2 (T2), and Type 3 (T3) AOSLO video. The portion of the plot corresponding to each video type is marked.

For a model object moving downward at 1 mm/s in the vertical direction, the overestimation is 8.1% for a Type 1 video, 10.1% for a Type 2 video, and 4.0% for a Type 3 video, assuming no eye motion. At 3 mm/s, the overestimation is 29.1% for a Type 1 video, 37.8% for a Type 2 video, and 13.0% for a Type 3 video (Fig. 6).

Thus, the raster scan error increases when one or more of the following occur:  $v_s$  decreases,  $\bar{v}_o$  increases, or the field of view decreases (the field of view inversely varies with the scale factor in the scan direction).

A similar analysis can be used to estimate the expected error in the horizontal scan direction. Since the scan speed is defined as the number of rows required to reach the edge of the frame, for the horizontal direction,  $RS_{\text{horizontal}} = RS_{\text{vertical}}/512$ . The percent error in the horizontal direction =  $100\% \cdot \frac{RS_{\text{vertical}}}{512 - RS_{\text{vertical}}}$ . Since  $RS_{\text{vertical}} < 1$  for objects of interest, the percent error in the horizontal direction will always be less than 0.2%. When  $RS_{\text{vertical}} = 0.5$ , the percent error in the horizontal direction drops to 0.098%. Therefore, we do not need to apply the raster scan correction to the horizontal component of calculated speeds.

## 2.7 Virtual AOSLO

The AOSLO is a custom-built, unique instrument with  $\sim 5\times$  the resolution of a commercial scanning laser ophthalmoscope. Typically, ground truth for new systems is generated using manual analysis performed by subject experts. However, due to the low contrast of the moving objects, it is difficult and unreliable to analyze videos by naked eye. Therefore, we used a virtual AOSLO to simulate realistic videos to create a synthetic dataset for use as ground truth in order to validate our methods.

### 2.7.1 Parameters

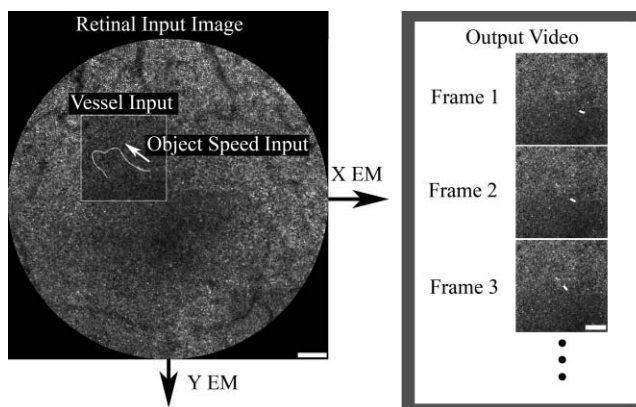
A virtual AOSLO has previously been used to characterize scanning distortions due to raster scanning and eye motion for static

images.<sup>22</sup> We modified this virtual AOSLO to simulate the acquisition of a video in the presence of an object that is moving at the same time as the scanner. For the virtual AOSLO, we selected scanning parameters for a Type 1 video, which is the data set on which the proposed methods were developed. Since we were able to exactly specify the speed and position of the moving object, we considered the simulated videos to be ground truth. Due to the complexity of the AOSLO, the following assumptions were used for the virtual AOSLO:

1. The imaging laser is a perfect, dimensionless dot that samples with true fidelity one pixel of the input image at one time.
2. The retina is rigid and planar across the field of view.
3. Eye motion is strictly translational with no torsional component.
4. Imaging parameters are constant.

The first assumption bypasses sampling and resolution issues introduced by the optics of a human eye. This means that the quality and appearance of the simulated video primarily depends on the input image. Second, for the field of view of the simulation (1.5 deg in each direction), it is reasonable to assume that this region is both rigid and planar. Third, as previously described, the primary components of eye motion are translational.<sup>7</sup> While we have observed torsional motions, they are typically small (unpublished experimental observations). The final assumption is that imaging parameters are constant — in actual practice, due to additional complexities such as calibration and temperature-dependent drift of electronic components, different imaging sessions have minor variations in imaging parameters (these variations are addressed using calibration steps prior to each imaging session).

An overview of the virtual AOSLO is shown in Fig. 7, and a summary of the parameters selected for the simulation is shown in Table 2. The input image was generated using individual frames from overlapping AOSLO videos near the fovea, which



**Fig. 7** Overview of virtual AOSLO used to simulate videos of a moving object in the presence of raster scanning and eye motion. An arbitrary vessel and flow direction was inserted into a retinal input image and an object speed was specified. The rigid input image, vessel, and object were translated according to specified X and Y Eye Motion (EM) inputs. The frames in the output video were generated using the scan parameters and configuration described in Sec. 2.7.1. Scale bars, 100  $\mu\text{m}$ .

**Table 2** Summary of parameters for the virtual AOSLO.

Parameter	Value
Horizontal raster frequency	15.36 kHz
Vertical raster frequency	30 Hz
Video frame size	512×512 pixels
Video acquisition rate	30 fps
Sampling band	Central 80% of forward sweep
Retinal scale factor	0.296 mm/deg
X and Y scale factors	341.33 pixels/deg
Vessel diameter	5 $\mu$ m
Leukocyte length	15 $\mu$ m in the direction of travel
Leukocyte speeds	1.00 to 3.00 mm/s

were scaled to the appropriate size (circle diameter = 3.75 deg). The spatial resolution of the input image was selected to be twice that of the output video and sampled by the virtual AOSLO using nearest neighbor interpolation. Thus, the simulated videos are similar to actual AOSLO videos, which allow us to apply the same correction steps that we would have applied to actual AOSLO videos.

Simulated videos were generated pixel by pixel. The basic steps were to 1. calculate the pixel timing given the raster parameters, 2. convert timings to spatial coordinates on the input image, and 3. sample the input image at the specified spatial coordinates. The time that each pixel is sampled can be directly computed using the raster scan parameters. To convert times to spatial coordinates, two calculations were done. First, the spatial coordinates were calculated assuming a static image. Second, the spatial coordinates were horizontally and vertically translated as specified by the *X* and *Y* components of eye motion, corresponding to the timing at each pixel location. Finally, sampling was performed after insertion of a moving object into the input image. For the moving object, we specified the trajectory and speed and used a high contrast, elongated object oriented along the direction of travel. The time-dependent input image was then sampled at the corresponding spatial coordinate for each pixel to generate a simulated AOSLO video.

### 2.7.2 Experiments

The virtual AOSLO was used to generate a synthetic data set for validation. The synthetic data set consisted of simulated videos with different configurations, varying in object speed, vessel geometry and orientation, eye motion, and noise (Table 3). A single moving object was used for each of the videos.

The goals of these experiments were to 1. verify that the error in measured speed was negligible in the horizontal direction, 2. verify that the theoretical errors in measured speeds were consistent in the vertical direction, and 3. examine the expected effects on calculated speeds due to experimental conditions. H3

**Table 3** Evaluation of speed quantification.

Video code	Vessel orientation	Object speed	Eye motion
H3	Horizontal	3 mm/s	No
V1	Vertical	1 mm/s	No
V3	Vertical	3 mm/s	No
V1_EM	Vertical	1 mm/s	Yes
A2	Arbitrary	2 mm/s	No
A2_EM	Arbitrary	2 mm/s	Yes

was used to quantify the error due to raster scanning in the horizontal case, for an object moving at 3 mm/s – the faster the object moves, the greater the error expected. Most objects traveled at speeds between 1 and 3 mm/s. V1 and V3 were used to quantify the error due to raster scanning alone for an object traveling in the vertical direction, as modeled in Sec. 2.6. For the experimental conditions, the two factors that contribute most to changes in measured speeds were considered: vessel trajectory and eye motion. For the vessel trajectory videos (A2, A2\_EM), we used the vessel centerline extracted from the Type 1 video as the vessel input. For the eye motion, we used the extracted eye motion trace from the first 1.3 s of the Type 1 video. We considered these two factors both separately (V1\_EM, A2), and simultaneously (A2\_EM).

## 3 Results

The proposed methods performed well on both the synthetic data set generated using the virtual AOSLO and on experimental videos acquired on the AOSLO.

### 3.1 Evaluation of Accuracy and Validity Using a Virtual AOSLO

We applied the proposed methods for tracking and speed quantification (Fig. 8). To measure speeds in the simulated videos, we repeated the analysis five times and took the average of computed speeds, in order to reduce the errors due to operator bias and differences in data precision. The data precision varied since there were three times as many data points that could be extracted to measure speeds at 1 mm/s versus 3 mm/s. At 3 mm/s, due to large pixels/frame displacements, the traces on the ST plots were disconnected.

To validate that objects were being correctly tracked, we generated a tracked video and used frame-by-frame examination. As expected, for all videos, the extracted traces corresponded to the moving objects. However, the labeled lines would sometimes lead or lag the moving objects by small amounts. Since the amount of lag/lead was preserved for each moving object, the slope of the traces was accurate. The error was due to the estimation of frame number from the coordinates of the extracted traces, due to the low temporal resolution relative to the speed of the leukocytes. Taking this into consideration, there were no false positives and no false negatives.



**Table 4** Evaluation of speed quantification. Speeds are reported before (no RS) and after (RS) the proposed correction. Actual speeds are the object speeds corresponding to each video, as listed in Table 3.

Video code	H3	V1	V3	V1_EM	A2	A2_EM
No RS [mm/s]	3.0082	1.1428	3.9370	1.1242	1.8956	2.0067
RS [mm/s]	3.0110	1.0552	3.0608	1.0361	1.9766	2.0283
No RS versus RS%	-0.09%	8.31%	28.63%	8.50%	-4.10%	-1.07%
RS versus Actual%	0.37%	5.52%	2.03%	3.61%	-1.17%	1.42%

We compared the corrected speeds to the actual speeds (Table 4). We define the residual error as the percent difference between corrected and actual speeds and found that the residual error was on average 2% for moving objects traveling between 1 and 3 mm/s. The sources of error are most likely due to vessel and trace extraction, which are dependent on user interaction. For experimental data, these sources of error are likely to increase due to 1. lack of prior information about vessel trajectories and 2. variations in trace slopes. For the synthetic data sets, extraction is more accurate due to prior knowledge about the shape of the vessel (since it was specified), and due to the fact that object speeds are uniform (so that there is no variation in trace slopes).

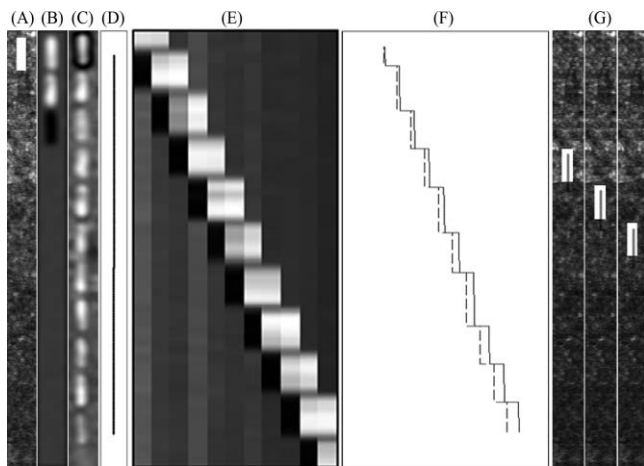
H3 confirms that the error in measured speed is negligible in the horizontal direction, since the calculated error was -0.09%. This is also in agreement with the theoretical model, which specifies an upper bound of 0.2% for the error. Therefore, it is a reasonable assumption to neglect the error due to horizontal scanning.

V1 and V3 confirm the theoretical errors due to the vertical component of raster scanning. We found errors of 8.31 and

28.63%, which are in agreement with the theoretical errors of 8.1 and 29.1%. Therefore, in the absence of eye motion, the computed errors are in agreement with the expected errors.

Eye motion can either increase or decrease the magnitude of the error. If eye motion is random and isotropic, then over time the average speed should not be affected by eye motion. However, if the eye favors motion along a preferred direction, then the computed speed is affected — the computed speed is maximally increased when the object, raster scan, and eye motion are in the same direction (i.e., all vertical and downward). Initially, the vertical component of the eye motion trace input is in the same direction of the scan — as expected, the error for V1\_EM is slightly larger than V1.

In practice, vessels are rarely horizontal or vertical, particularly when considering capillaries. First, the magnitude of the error in calculated speed depends on the trajectory of the vessel at the object location, since only the vertical component of speed is corrected. Therefore, deviations from a vertically-oriented vessel should result in diminishing error magnitudes. Second, the start and end points of the vessel ultimately determine whether speeds are over- or underestimated. The vessel in A2 and A2\_EM has both up- and downward components, but since the endpoint is lower than the starting point, this means that we should expect speeds to be underestimated when comparing A2 to V2. Since the eye motion results in a slight overestimation (comparing V1\_EM to V1 and using the same eye motion input), this explains why the error for A2\_EM is less than the error for A2.

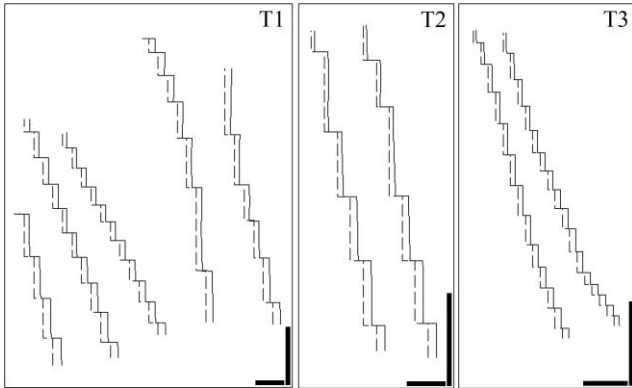


**Fig. 8** Tracking and speed quantification on V1\_EM. (A) First frame in which the object appears from a preprocessed video. (B) Corresponding frame in the multiframe division video. (C) Standard deviation image calculated from 15 frames. (D) Extracted vessel for offline generation of the ST plot. (E) ST plot for the selected vessel. (F) Extracted trace from the ST plot (dotted line on left) and the corrected trace taking into consideration both raster scanning and eye motion (solid line on right). Notice that the speed was overestimated prior to correction, as expected. (G) Three consecutive frames of the video showing tracking results.

### 3.2 Evaluation on Experimental AOSLO Videos

We performed the proposed methods on 40 vessels from ten AOSLO videos; first we report results across all videos and then we show detailed results for one representative vessel for each video Type.

Ten vessels were analyzed from 1 Type 1 video, ten vessels from 3 Type 2 videos, and 20 vessels from six Type 3 videos. The average absolute error in measured speed was 2.59% for the Type 1 video, 3.39% for the Type 2 video, and 2.04% for the Type 3 video, where absolute error was defined as the absolute value of the percent difference between corrected and noncorrected speeds for one trace, and the average absolute error was defined as the average absolute error across all extracted traces for each video Type. For comparison, we estimated the error using the RS parameter as defined in Sec. 2.6, taking  $\bar{v}_o$  to be the average object speed. In the absence of eye motion, for a vertically-oriented vessel, the theoretical error was 12.56% for



**Fig. 9** Effect of raster scan showing extracted traces (dotted lines on left) and raster scan corrected traces (solid lines on right), for Type 1 (T1, frames 800–830), Type 2 (T2, frames 48–60), and Type 3 (T3, frames 2278–2302) videos. The shifts are nonuniform due to the confounding effects of eye motion. Vertical scale bars, 0.1 mm; horizontal scale bars, 0.1 s.

the Type 1 video, 12.84% for the Type 2 video, and 5.27% for the Type 3 video. This suggests that either vessel orientations are horizontally biased or that eye motion is not uniformly distributed across all orientations.

We selected three representative vessels from Type 1, Type 2, and Type 3 videos to further characterize the error in measured speeds. For each vessel, traces were extracted from ST plots and used for tracking on the original videos and speed quantification (Fig. 9). Close examination of Fig. 9 shows that the orientation of the vessel has an effect on the slope modification as long as the effect due to eye motion is small (i.e., one can see whether the slope was over- or underestimated corresponding to a downward- and upward-oriented vessel). We will discuss this effect in more detail considering the actual errors in average speeds that were calculated (Table 5). In experimental videos, there are complexities such as arbitrary vessel shapes and orientations, eye motion, and variations in cell speeds both temporally and spatially. There is also noise due to variations in the intensity of the background photoreceptor tissue, likely due to dynamic scattering changes over time<sup>23</sup> and also coherent artifacts.<sup>24,25</sup> These variations generate noise in the multiframe division videos and affect the appearance of the ST plots. Therefore, the actual error in calculated speeds due to raster scanning and eye motion

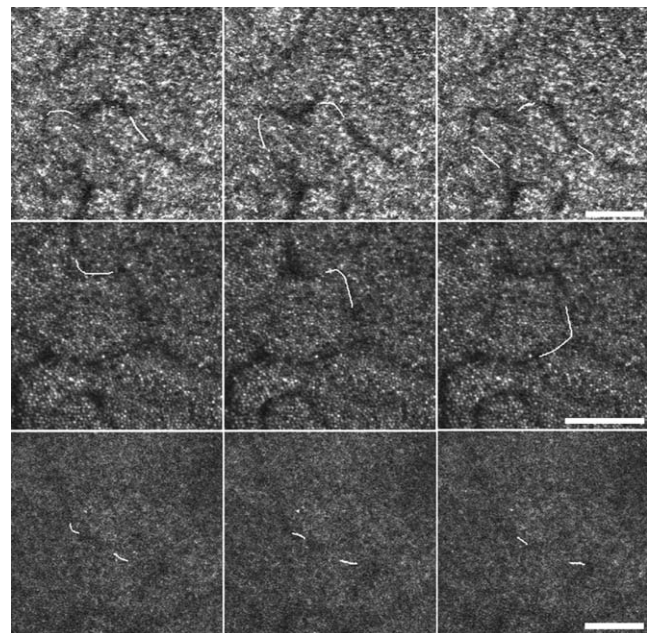
**Table 5** Summary of cell speeds in selected vessel segments with and without the raster scan correction.

Parameter	Type 1	Type 1	Type 2	Type 2	Type 3	Type 3
RS	no	yes	no	yes	no	yes
N	50	50	4	4	12	12
Mean [mm/s]	2.04	1.99	1.89	1.73	2.18	2.12
SD [mm/s]	0.62	0.58	0.30	0.26	0.45	0.43
Min [mm/s]	0.93	0.94	1.62	1.50	1.56	1.53
Max [mm/s]	3.47	3.27	2.16	1.97	2.90	2.80

will be different. We compared our corrected speeds to uncorrected speeds, where uncorrected speeds were simply taken as the slope of the manually extracted trace, which assumes that the entire frame was acquired at the same moment in time.

The actual error in the average speeds is 2.51% for the Type 1 video, 9.25% for the Type 2 video, and 2.83% for the Type 3 video. As previously described, the magnitude and sign of the error is largely determined by the trajectory of the vessel. For all three types, the vessels deviate from a purely vertical vessel, and so the magnitude of the error is diminished compared to the model (Sec. 2.6). In addition, because the end point of each vessel is lower than the start point, we expect an overestimation of speed for all three Types. The Type 2 video has the largest error, as predicted by the theoretical model (Fig. 6). Notice that there is a nonlinear shift that results due to the raster scan correction. Close examination of the Type 1 traces in Fig. 9 show that the slopes at the bottom of individual traces were decreased after application of the raster scan correction, while slopes at the top were increased. This suggests that the entrance side of the path segment underestimated speeds, while the exit side overestimated speeds, corresponding to a net upward and net downward vertical orientations, respectively. As can be seen from Fig. 5, this was exactly the case. As a final comparison, for the Type 1 video, the error was 2.51%, compared to  $-1.07\%$  for the same vessel in A2\_EM. The reason for this is due to a small difference in the starting and ending points of the vessel. Although we used the same trajectory, the end point of the vessel slightly terminates higher than the starting point for A2\_EM, which explains the difference in the sign of the error.

To verify that each extracted trace corresponded to an object on the input video, extracted traces were registered (Sec. 2.4; illustrated in Fig. 10). We individually verified each extracted trace by examining the tracked video frame-by-frame. Overall,



**Fig. 10** Moving objects labeled using white lines in the cropped video using extracted traces from ST plots, for a Type 1 (top), Type 2 (middle), and Type 3 (bottom) video. For visualization purposes, we thickened the line by 2 pixels. Scale bars, 100  $\mu\text{m}$ .

the labeled lines tracked the leukocytes well. There were no false positives; it was not possible to calculate a false negative rate.

#### 4 Discussion

This paper presents a method for quantifying object speeds in AOSLO videos. We demonstrated a multiframe approach for motion contrast enhancement that improves the contrast of moving objects and vessels. Motion contrast enhanced ST plots were used to visualize hemodynamics and individual traces were extracted for analysis. Extracted traces were used to track objects on the input videos and also for speed quantification. Speed quantification was done using a slope-modification technique that corrects for raster scanning in the presence of eye motion. We validated our results using a virtual AOSLO. The combination of selected techniques is significant in terms of putting together a complete system of video and image analysis for noninvasive vascular video imaging.

Our results are similar to other methods. A previously reported method using manual identification and analysis on the same vessel from the same Type 1 AOSLO video used in this paper found a total of 35 objects with a speed of  $1.82 \pm 0.42$  mm/s, without considering the error due to raster scanning or eye motion; our uncorrected speed was  $2.04 \pm 0.62$  mm/s for 50 objects. While the numbers are similar, the discrepancies can be explained with the following considerations: The number of objects identified by the manual method was less than our method, probably due to difficulties in visualizing objects without motion contrast enhancement. It may have been more difficult to visualize objects that were traveling at faster speeds using the manual method. Finally, the vessel trajectory may not have been as accurate in the manual method.

There are a few similar results from different imaging modalities. Using fluorescein-aided scanning laser ophthalmoscopy, blood flow velocity was measured to be  $3.29 \pm 0.45$  mm/s (standard deviation) in the parafoveal capillaries of 21 healthy volunteers.<sup>26</sup> Our measured speeds are similar in magnitude to these results, but we can explain the discrepancies as follows. First, the location and size of the capillaries was different. Second, we measured leukocyte speeds, while they measured whole blood speeds using fluorescein. It is known that leukocytes travel slower through capillaries than erythrocytes,<sup>27</sup> which constitute the majority of blood by volume. Thus, differences in spatial locations, small sample size, and differences in the element of blood that is being measured could account for differences in measured speeds. The blue field entoptic phenomenon is another method to examine capillary flow in the parafoveal region that can be used for estimating blood velocities.<sup>28</sup> The blue field entoptic phenomenon refers to the movement of “flying corpuscles” that can be seen when looking at an illuminated blue background.<sup>29</sup> It is thought that these flying corpuscles are in fact leukocytes. By having observers compare the speeds of these moving objects to those of simulated velocity fields, one can estimate speeds. One study found a speed of  $0.89 \pm 0.2$  mm/s,<sup>28</sup> while another found speeds between approximately 0.5 and 1 mm/s.<sup>30</sup> These speeds are similar in magnitude to those that we obtained, but one needs to be cautious since the blue field technique is subjective in nature.

The methods presented in this paper can be potentially applied to other high-resolution scanning systems with moving

objects. There are many areas for future work, including full automation and application of more advanced detection and tracking methods. There are also important microcirculation studies that can be performed, including development of a family of hemodynamic markers to investigate leukocyte behavior. Such markers could be used to quantify changes in leukocyte behavior for normal and diseased retinas. The human eye allows for a unique opportunity to directly examine the microcirculation, which has been made possible due to improvements in imaging techniques (AOSLO) combined with the image analysis algorithms presented in this paper.

#### 5 Conclusion

Raster scanning and eye motion contribute to significant sources of error when quantifying speed on AOSLO videos. The magnitude of this error depends on the speed of the moving object, configuration of AOSLO imaging parameters, the orientation of the vessel, and the isotropy of the eye motion, but can be as large as 37.8%. Slope modification on ST plots can correct for this error, improving the accuracy of hemodynamics using AOSLO.

#### Acknowledgments

The authors would like to thank Mark Campanelli for developing the initial implementation of the virtual AOSLO, Scott Stevenson for his valuable insights regarding ST plots, and Qiang Yang for his help with video stabilization and desinuosiding. This work was based on research funded in part by the NSF Center for Adaptive Optics AST-9876783 and the NIH Bioengineering Research Partnership EY014375. Johnny Tam is supported in part by a National Defense Science & Engineering Graduate Fellowship (NDSEG), sponsored by the Department of Defense, and in part by a National Science Foundation Graduate Research Fellowship.

#### References

1. J. A. Martin and A. Roorda, “Direct and non-invasive assessment of parafoveal capillary leukocyte velocity,” *Ophthalmology* **12**, 2219–2224 (2005).
2. J. Liang, D. R. Williams, and D. T. Miller, “Supernormal vision and high-resolution retinal imaging through adaptive optics,” *J. Opt. Soc. Am. A* **14**, 2884–2892 (1997).
3. A. Roorda, F. Romero-Borja, W. J. Donnelly III, H. Queener, T. J. Hebert, and M. C. W. Campbell, “Adaptive optics scanning laser ophthalmoscopy,” *Opt. Express* **10**, 405–412 (2002).
4. Y. Zhang, S. Poonja, and A. Roorda, “MEMS-based adaptive optics scanning laser ophthalmoscopy,” *Opt. Lett.* **31**, 1268–1270 (2006).
5. L. C. Sincich, Y. Zhang, P. Tiruveedhula, J. C. Horton, and A. Roorda, “Resolving single cone inputs to visual receptive fields,” *Nat. Neurosci.* **12**, 967–969 (2009).
6. J. Tam, J. A. Martin, and A. Roorda, “Non-invasive visualization and analysis of parafoveal capillaries in humans,” *Invest. Ophthalmol. Visual Sci.* **51**, 1691–1698 (2010).
7. S. Martinez-Conde, S. L. Macknik, and D. H. Hubel, “The role of fixational eye movements in visual perception,” *Nat. Rev. Neurosci.* **5**, 229–240 (2004).
8. S. Stevenson, A. Roorda, and G. Kumar, “Eye Tracking with the Adaptive Optics Scanning Laser Ophthalmoscope,” presented at *Proc. Eye Tracking and Research Applications (ETRA) Meeting* (2010).
9. C. R. Vogel, D. W. Arathorn, A. Roorda, and A. Parker, “Retinal motion estimation in adaptive optics scanning laser ophthalmoscopy,” *Opt. Express* **14**, 487–497 (2006).

10. Q. Yang, D. W. Arathorn, C. R. Vogel, Y. Zhang, P. Tiruveedhula, and A. Roorda, "Retinally stabilized cone-targeted stimulus delivery," *Opt. Express* **15**, 13731–13744 (2007).
11. F. Reinholz, R. A. Ashman, and R. H. Eikelboom, "Simultaneous three wavelength imaging with a scanning laser ophthalmoscope," *Cytometry* **37**, 165–170 (1999).
12. Y. Sato, J. Chen, R. A. Zoroofi, N. Harada, S. Tamura, and T. Shiga, "Automatic Extraction and Measurement of Leukocyte Motion in Microvessels using Spatiotemporal Image Analysis," *IEEE Trans. Biomed. Eng.* **44**, 225–236 (1997).
13. S. A. Japee, C. G. Ellis, and R. N. Pittman, "Flow Visualization tools for Image Analysis of Capillary Network," *Microcirculation* **11**, 39–54 (2004).
14. D. A. Nelson, S. Krupsky, A. Pollack, E. Aloni, M. Belkin, I. Vanzetta, M. Rosner, and A. Grinvald, "Special report: Noninvasive multi-parameter functional optical imaging of the eye," *Ophthalmic Surg Lasers Imaging* **36**, 57–66 (2005).
15. J. Tam and A. Roorda, "Enhanced detection of cell paths in spatiotemporal plots for noninvasive microscopy of the human retina," in *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium* 584–587, Rotterdam, Netherlands (2010).
16. C. G. Ellis, M. L. Ellsworth, R. N. Pittman, and W. L. Burgess, "Application of Image Analysis for the Evaluation of Red Blood Cell Dynamics in Capillaries," *Microvasc. Res.* **44**, 214–225 (1992).
17. P. S. Jensen and M. R. Glucksberg, "Regional variation in capillary hemodynamics in the cat retina," *Invest. Ophthalmol. Visual Sci.* **39**, 407–415 (1998).
18. D. Kleinfeld, P. P. Mitra, F. Helmchen, and W. Denk, "Fluctuations and stimulus-induced changes in blood flow observed in individual capillaries in layers 2 through 4 of rat neocortex," *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15741–15746 (1998).
19. Z. Zhong, B. L. Petrig, X. Qi, and S. A. Burns, "In vivo measurement of erythrocyte velocity and retinal blood flow using adaptive optics scanning laser ophthalmoscopy," *Opt. Express* **16**, 12746–12756 (2008).
20. H. Xu, A. Manivannan, K. A. Goatman, J. Liversidge, P. F. Sharp, J. V. Forrester, and I. J. Crane, "Improved Leukocyte Tracking in Mouse Retinal and Choroidal Circulation," *Exp. Eye Res.* **74**, 403–410 (2002).
21. K. Y. Li, P. Tiruveedhula, and A. Roorda, "Inter-subject variability of foveal cone photoreceptor density in relation to eye length," *Invest. Ophthalmol. Visual Sci.* **51**, 6858–6867 (2010).
22. M. Campanelli, C. Vogel, and A. Roorda, "Dewarping Scanned Retinal Images," unpublished report, Center for Adaptive Optics (CfAO) (2003).
23. A. Pallikaris, D. R. Williams, and H. Hofer, "The reflectance of single cones in the living human eye," *Invest. Ophthalmol. Visual Sci.* **44**, 4580–4592 (2003).
24. R. S. Jonnal, J. Rha, Y. Zhang, B. Cense, W. Gao, and D. T. Miller, "In vivo functional imaging of human cone photoreceptors [Abstract]," *Opt. Express* **15**, 16141–16160 (2007).
25. N. M. Putnam, D. X. Hammer, Y. Zhang, D. Merino, and A. Roorda, "Modeling the foveal cone mosaic imaged with adaptive optics scanning laser ophthalmoscopy," *Opt. Express* **18**, 24902–24916 (2010).
26. O. Arend, S. Wolf, F. Jung, B. Bertram, H. Postgens, H. Toonen, and M. Reim, "Retinal microcirculation in patients with diabetes mellitus: dynamic and morphological analysis of perifoveal capillary network," *Br. J. Ophthalmol.* **75**, 514–518 (1991).
27. J. Ben-nun, "Comparative Flow Velocity of Erythrocytes and Leukocytes in Feline Retinal Capillaries," *Invest. Ophthalmol. Visual Sci.* **37**, 1854–1859 (1996).
28. O. Arend, A. Harris, W. E. Sponsel, A. Remky, M. Reim, and S. Wolf, "Macular capillary particle velocities: a blue field and scanning laser comparison," *Graefe's Arch. Clin. Exp. Ophthalmol.* **233**, 244–249 (1995).
29. M. Loebel and C. E. Riva, "Macular circulation and the flying corpuscles phenomenon," *Ophthalmology* **85**, 911–917 (1978).
30. C. E. Riva and B. Petrig, "Blue field entoptic phenomenon and blood velocity in the retinal capillaries," *J. Opt. Soc. Am.* **70**, 1234–1238 (1980).