NEAR-REAL-TIME SACCADE DETECTION
WITH SCANNING LASER OPHTHALMOSCOPES

V. C. Aitken
Systems and Computer Engineering,
Carleton University, Ottawa, ON,
vaitken @ sce.carleton.ca

A.B. Markov
Amtech Aeronautical Limited
Medicine Hat, Alberta
alex.markov @ amtech-group.com

Abstract

This paper describes a method to analyze raster scans from a scanning laser ophthalmoscope (SLO) to derive a near-real-time shutter control signal to block a treatment laser within 5ms of the onset of a saccadic shift. The proposed method exploits the scanning nature of the SLO and assumes direct access to each scan line as it is received. A small number of scan lines are grouped to form horizontal bands of the scanned area. Several figures-of-merit are introduced to generate robust decisions on saccade occurrence within each band. Results are shown for both simulated and real SLO image sequences. Fast robust saccade detection has further applications in high accuracy retinal motion estimation and studies of temporal properties of saccadic eye motion in humans

Keywords: Image motion; ophthalmology; saccadic eye motion; real-time; scanning laser ophthalmoscope.

1. INTRODUCTION

Involuntary movements of the eye, called *nystagmus*, include the *fixation reflex* that occurs when an object comes into the peripheral view. Nystagmus also includes three dominant components when the subject attempts to maintain steady fixation on a stationary distant point: 1) high frequency (30-70Hz) movements of small amplitude (~20 sec of arc); 2) saccadic motion (~ 1-2 min of arc) at high angular rates and at fairly regular intervals (~1Hz); and 3) slow regular drift. Saccadic motion is known to be extremely rapid and precise. Several studies have evaluated analytical models of the fixation reflex [1-5] based on 30Hz imagery from fundus cameras or high rate eye motion measurements from external sensors. The involuntary saccadic reflex (Item 2 above), rather than the fixation reflex in response to external stimuli, is the focus of this paper.

Scanning Laser Ophthalmoscopes (SLOs) [6] form images by scanning the fundus with a laser beam and typically provide standard video outputs (30Hz) that can be captured by computer systems. Photocoagulation surgery is used to correct retinal tears or to reduce vision loss due to macular degeneration or diabetic retinopathy through delivery of treatment laser energy to precise points on the retina using bursts of duration from 100-300 ms. Without anaesthesia the saccadic eye reflex can result in involuntary, very fast, and significant eye motion during surgery. Related literature [7-10] recommends that irradiation of incorrect retinal positions should be limited to 5% of the total treatment irradiation time to avoid excessive damage to healthy cells. For a nominal 200ms firing time, an incorrectly directed treatment laser should be blocked in less than 10ms, 3 times faster than the minimum available through analysis of 30Hz SLO imagery.

This paper describes the saccade detection algorithm originally designed during the CSA/STEAR Automated Laser Eye Treatment (ALET) Project [11,12]. The objective of the ALET Project was to develop and test algorithms to estimate eye motion from SLO imagery and provide feedback control for a treatment laser so that photo-coagulation surgery could be conducted safely without the need for anaesthesia to immobilize the eye.

Our work differs from previous methods [4,5,7-10,13-15] due to the fast temporal response needed to satisfy shutter timing requirements. In order to achieve shutter response times on the order of 5ms or less, the scanning nature of the SLO device is exploited. This approach assumes that an interface can be provided to the internal A/D conversion of laser data before any significant delays or buffering needed to produce the video signal. The ability to formulate these decisions in a robust algorithm and shutter the laser within 5ms of the onset of a saccade is to unique to the ALET work.

2. ALGORITHM REQUIREMENTS

SLO systems, such as those produced by LDT[1] or Rodenstock[2] have internal optical arrangements that generate video output signals in significantly different ways. Figure 1 shows simple but representative image formation geometry for a typical SLO system. The imaging laser scans a roughly rectangular region of the retina. The entrance aperture at the pupil, called the *beam scan waist*, is the origin of the raster coordinate system. The beam is focused onto a roughly 10μm spot (15 degree field of view) at the retina which lies about 24mm beyond the beam scan waist. Video signals produced by SLOs also differ in terms of timing, pixel resolution, number of lines per image, scanning pattern (left-to-right only with horizontal flyback, or alternating left-to-right followed by right-to-left) and interlaced vs progressive scan [11,12,15]. The time required

---

[1] Laser Diagnostic Technologies Inc., San Diago CA.
[2] Rodenstock USA Inc., Danbury CT.

to produce each scan line varies between $63.5\,\mu$s for the Rodenstock to $125.6\,\mu$s for the LDT AngioScan. These temporal constraints dictate that the algorithm must be simple and fast, but not at the expense of robustness.
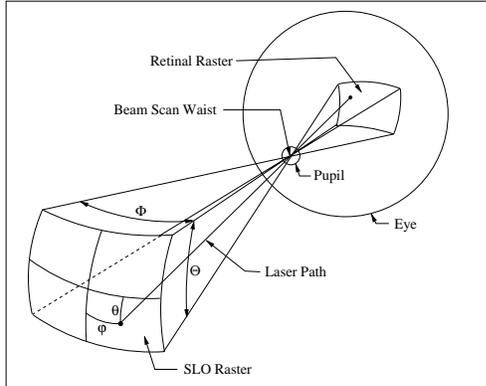
Figure 1: Representative SLO image formation.

The proposed method is based on analysis of both intensity and edge data in each laser scan line (row) in order to estimate motion. Within each SLO raster, groups of N_r consecutive rows form horizontal bands of the scanned area. A band size of 64 Rodenstock lines or 32 LDT AngioScan lines will give band data sets at just over 4ms intervals, thereby leaving 1ms for shutter actuation response time to achieve the 5ms maximum latency. Several figures-of-merit (FOMs) are derived to assess the analysis of scan data and generate robust decisions on the occurrence of a saccade within each horizontal band.

The algorithm compares edge information within each scan line to that from the corresponding line in the previous SLO raster in order to estimate motion. A transition from no motion in the previous band to significant motion in the current band indicates that a saccadic shift started within the current band and this event triggers the shutter close signal.

3. ALGORITHM DESCRIPTION

The saccade detection algorithm consists of three processes: one process is driven by the SLO pixel clock and operates as each row of the image is received; another process is executed at the end of each row; and a final process operates after each band has been received to make the shutter decisions.

As each pixel is received, a small median filter is first applied to remove shot noise followed by an edge mask convolution from left to right along the row. The edge mask [7-9], $M_e = [-1,...,-1,0,1,...1]$, has an odd length, N_M, with a zero centre element called the *zero position* of the mask. Dark to bright transitions in the image produce a positive response from the mask (called a *Left Edge*), while bright to dark transitions produce a negative response (called a *Right Edge*). Figure 2 shows a raw SLO image and the corresponding *edge map* obtained using the edge mask. Any

non-horizontal edge in the image will produce at least a small response in the edge map.
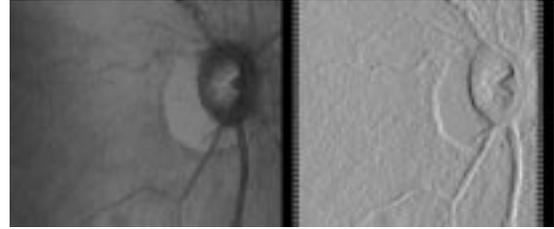
Figure 2: Raw SLO image and edge map.
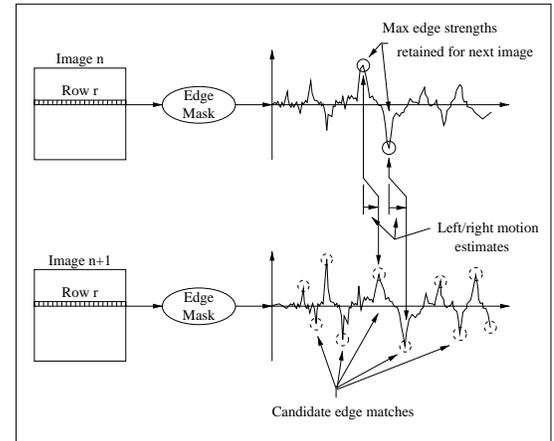
Figure 3: Basic algorithm.

The algorithm is based on processing information associated with the maximum left edge strength and maximum right edge strength within each row of each SLO raster as illustrated in Figure 3. In what follows, subscript L refers to left edges, subscript R refers to right edges, and subscript L/R is used when a process refers to either right or left edges. Superscripts refer to the SLO raster number.

Two information sets are retained for each row, r, from raster $n-1$ until the corresponding row of the next raster, n, has been completely processed. One set corresponds to the maximum left edge strength within the row, and the other corresponds to the maximum right edge strength within the row. These sets are defined as

$$I_{L/R}^{n-1}(r) = \left\{ E_{L/R}^{n-1}(r), P_{L/R}^{n-1}(r), Q_{L/R}^{n-1}(r) \right\},$$

where $E_{L/R}^{n-1}(r)$ represent the extreme values of the edge map within the row, $P_{L/R}^{n-1}(r)$ are the horizontal positions at which the zero position of the mask was applied to obtain the extreme edge values, and $Q_{L/R}^{n-1}(r)$ are the vectors of N_M intensity pixel values that contributed to these extreme edges.

Once row r of image n has been received and convolved with the edge mask, $2N_c$ candidate information sets are computed for the strongest left and right edge *candidates*, c, of the form

$$I_{L/R}^n(r,c) = \left\{ E_{L/R}^n(r,c), P_{L/R}^n(r,c), Q_{L/R}^n(r,c) \right\}.$$

The best left and right candidate numbers, $\hat{c}_{L/R}$, are those values of $c \in \{1,2,...,N_c\}$ that minimize the *candidate pixel intensity metric* for left and right edges defined as

$$\delta_{p,L/R}(c) = \left\| Q_{L/R}^n(r,c) - Q_{L/R}^{n-1} \right\|.$$

Motion estimates, $m_{L/R}^n(r)$, for row r in image n, relative to row r in image $n-1$, are computed from both the best left and right edge candidates as

$$m_{L/R}^n(r) = P_{L/R}^n(r,\hat{c}_{L/R}) - P_{L/R}^{n-1}(r).$$

These estimates correspond to *apparent* horizontal motion for that row. During motion of the retina, any non-horizontal edge that, locally, does not lie strictly along the direction of motion will result in an apparent horizontal shift.

The next step in the algorithm is to derive intermediate assessments of algorithm performance and image integrity. Failure in image integrity can occur, for example, when a patient blinks or when the image becomes out of focus. In these cases the values produced by the edge mask tend to zero indicating a loss of edge information in the row.

Each FOM test is defined as a triplet $\{\delta, \lambda, \omega\}$ where δ represents the FOM metric, while parameters λ and ω are, respectively, the FOM threshold, and the FOM error weight that, upon test failure, is added to the *row error vote*, e_r. Row error votes are accumulated to form a *band error vote*, e_b. Motion estimates retained from rows within a band are stored in the *band motion set*, M_b and are combined to derive a final band motion estimate. The four FOM tests are described as follows.

Pixel Intensity Test – The left/right metrics are $\delta_{p,L/R} = \delta_{p,L/R}(\hat{c}_{L/R})$. This test quantifies how well underlying intensity values for the top candidate edges match the corresponding values from this row in the previous SLO raster. If either metric exceeds the threshold then the corresponding L/R row motion estimate is discarded and the L/R weight is added to the row error vote, otherwise the row motion estimate(s) is/are entered into the band motion set.

Edge Distance Test – The single distance metric is

$$\delta_d = \left\| \left| P_R^n(r,\hat{c}_R) - P_L^n(r,\hat{c}_L) \right| - \left| P_R^{n-1}(r) - P_L^{n-1}(r) \right| \right\|.$$

This test compares the distance between the left and right edges extracted from SLO raster $n-1$ to the distance between those for the top candidate edge positions in raster n. If the metric exceeds the threshold, then the weight is added to the row error vote.

Edge Order Test – This test examines the relative horizontal positions (order) of the edges within a row. The order "metric", δ_o, in this case is binary. If the edge order does not match, then the weight is added to the row error vote.

Row Information Test –The row information metric is defined as

$$\delta_I(r) = \max\left\{ \left| E_L^n(r,\hat{c}_L) \right|, \left| E_R^n(r,\hat{c}_R) \right| \right\}.$$

This metric tends to zero when the image degrades or when a blink occurs. The weight is added to the row error vote whenever the metric is **less** than the threshold.

Following the application of all FOM tests for a row, the accumulated row error vote, e_r, is compared to a threshold Λ_r and if $e_r > \Lambda_r$ then the band error vote, e_b, is incremented by one. Following processing of each row within a band, the band error vote is available with a maximum value of N_r, and up to $2N_r$ motion estimates are available in the band motion set M_b. The band error vote, e_b, is compared to a threshold, Λ_b, and if $e_b > \Lambda_b$ then the decision is made to shutter the treatment laser. The band motion estimate is taken as the median of the band motion set M_b. If the estimated band motion is greater than a predefined threshold, then the decision is made to shutter the treatment laser.

4. OVERVIEW OF RESULTS

Algorithms developed during the ALET project were tested on simulated SLO data and on 30Hz recorded imagery from an LDT AngioScan. The SLO/eye simulator was implemented in Matlab™ with simulated time resolution of ¼ of the SLO pixel clock. A true SLO image is used as a *base retinal map* for the simulation. Using the geometry illustrated in Figure 1, the saccade motion models of [1], restoring slow drift motion, representative scanning dynamics for the SLO, and weighted nearest neighbour averaging for interpolation of pixel intensities of the base map, the eye motion and laser scanning imaging processes are combined to generate a temporally accurate image sequence. Randomness is added to saccade amplitude, and saccadic motion has both horizontal and vertical components, but we force the saccade direction to alternate to ensure that the simulated laser remains within the boundaries of the base retinal map. An example eye motion profile is shown in Figure 4.

Band motion estimates using the proposed algorithm are compared to the true inter-frame motion for this simulated sequence in Figure 5. Not only are the temporal requirements satisfied, but the band motion estimates agree very well with the actual motion. Figure 6 shows band motion estimates and the shutter flag obtained from applying the algorithm to a real LDT AngioScan image sequence for a healthy eye and a band motion threshold of 3 pixels. The shutter flag at about 4.4 seconds is due to a blink and results from the failure of the row information test.
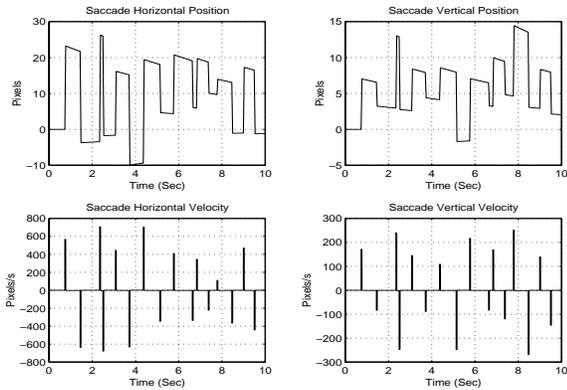
Figure 4: True eye motion for simulation.
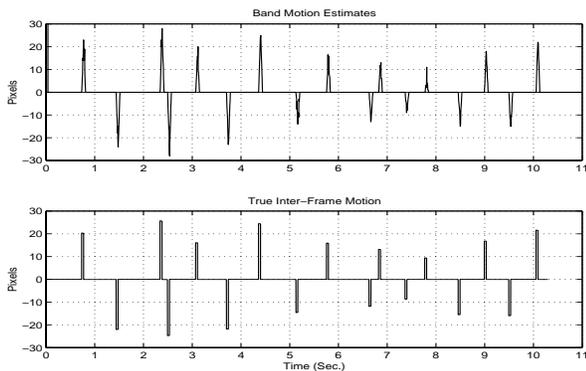
Figure 5: Results for simulated image sequence.
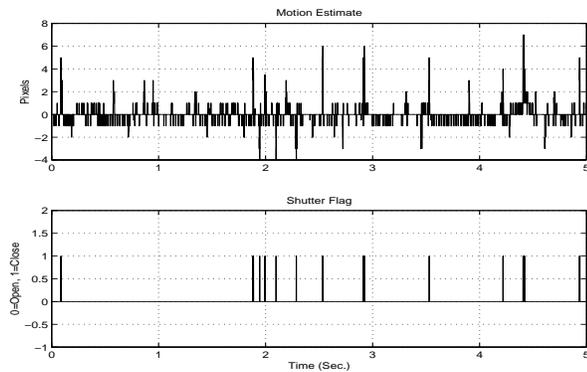
Figure 6: Results for real SLO data.

5. CONCLUSIONS

SLO image sequences can provide rich temporal information of retinal motion processes. The algorithm presented here was designed for near-real-time detection of the onset of saccadic motion in order to shutter a treatment laser with maximum latency of 5ms. Results shown for simulated and real SLO raster sequences demonstrate the effectiveness of the proposed approach. The saccade detection technique also has applications in further studies of the temporal properties of saccadic eye motion in humans and possible correlation of the saccade processes with emotional and/or medical conditions of the patient.

REFERENCES

[1] P.F. Gangemi, *et al.*, "Comparison of two nonlinear models for fitting saccadic eye movement data," *Computer Methods and Programs in* Biomedicine, Vol. 34, 1991, pp 291-297.

[2] V.V. Molebny, "Laser tracking in the refractive eye microsurgery," *Proc. SPIE: Industrial Applications of Laser Radar,* Vol. 2271, Sept 1994, pp 176-183.

[3] B.L. Philipp, C.S. Hsu, "Discrete system dominance index for model reduction," *Applied Mathematical Modelling,* Vol. 20, No. 6, June 1996, pp 429-439.

[4] D. Sauter, *et al.,* "Analysis of eye tracking movements using innovations generated by a Kalman Filter," *Medical and Biological Engineering and Computing,* Vol. 29, No. 1, Jan 1991, pp 63-69.

[5] K. Arai, E.L. Keller, J.A. Edelman, "Spatio-temporal neural network model of saccade generation," *Proc. 1993 IEEE Int. Conf. Neural Networks,* San Francisco CA, 1993, pp 70-74.

[6] D. Ott, *et al.,* "Scanning laser ophthalmoscopy (SLO) – A new tool in vision and oculomotor research," *Proc. SPIE MedTech '89: Medical Imaging,* Vol. 1357, 1989, pp218-227.

[7] *M*.D. Markow, H.G. Rylander, A.J. Welch, "Real-time algorithm for retinal tracking," *IEEE Trans. Biomedical* Engineering, Vol. 4, No. 12, Dec 1993, pp 1269-1281.

[8] S.F. Barret, *et al*., "Computer-aided retinal photo-coagulation system," *J. Biom. Optics*, Vol. 1, No. 1, Jan. 1996, pp. 83-91.

[9] C.H. Wright, *et al*., "Hybrid eye tracking for computer-aided retinal surgery," *Proc. 33rd International ISA Biomedical Sciences Instrumentation Symposium,* Colerado Springs, CO USA, Vol. 32, 1996, pp225-236.

[10] R.D. Ferguson, *et al.*, Hybrid tracking and control systems for computer-aided retinal surgery, *Proc. SPIE: Ophthalmic Technologies VI*, Vol. 2673, June 1996, pp32-41.

[11] V.C. Aitken, "Eye motion detection and estimation from scanning laser ophthalmoscope (SLO) imagery," Amtech Aeronautical Ltd. Rep. No. TR976004, CSA/STEAR contract 9F006-5-0545/009/ST, 1997.

[12] V.C. Aitken, K.L. Russell, C. Milner, "ALET:Motion detection and image tracking algorithm detailed design," Amtech Aeronautical Ltd. Rep. No. TR990705, CSA/STEAR contract 9F006-5-0545/009/ST, 1999.

[13] D.Ott, M. Lades, "Measurement of eye rotations in three dimensions and the retinal stimulus projection using scanning laser ophthalmoscopy," *American Journal Ophthal. Physiol. Optics,* Vol. 10, 1990, pp 67-71.

[14] X. Xie, R. Sudhakar, H. Zhuang, "Real-time eye feature tracking from a video image sequence using a Kalman filter," *IEEE Transactions on Systems Man and Cybernetics,* Vol. 25, No. 12, Dec. 1995, pp 1568-1577.

[15] M. Stetter, et al., "SLO saccadic profile measurements and the effects of retinal raster size and distortion," in *Proc., SPIE Lasers in Ophthalmology III,* Vol. 2632, Eds. R. Birngruber, A.F. Fercher, Jan. 1996, pp 98-109.