Tracking retinal motion with a scanning laser ophthalmoscope

Zhiheng Xu, MS;[1] Ronald Schuchard, PhD;[2] David Ross, MS;[2] Paul Benkeser, PhD[1]*

[1]*Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA; [2]Rehabilitation Research and Development Center, Atlanta Department of Veterans Affairs Medical Center, Atlanta, GA*

Abstract—The vast majority of people with low vision retain some functional vision to perform everyday tasks. To study the effectiveness and efficiency of the visual tasks performed by people with low vision, knowing the movement patterns of their preferred retinal locus (PRL) used for fixation, saccade, and pursuit is critical. The scanning laser ophthalmoscope (SLO) has been used to acquire retinal images while a subject is performing a visual tracking exercise. SLO data has traditionally been analyzed with the use of manual techniques that are both time-consuming and prone to errors due to operator fatigue. To improve the speed and accuracy of the analysis of retinal motion from SLO image sequences, we developed an automated image processing technique and tested it using MATLAB[TM] (The MathWorks, Natick, MA) software. The new software technique was experimentally tested on both normal- and low-vision subjects and compared with the results obtained using manual techniques. The findings indicate that the new technique works very well for most subjects, fairing poorly only in subjects where the quality of the SLO images was substandard.

Key words: block matching, distortion, eye tracking, low vision, manual registration, PRL, rehabilitation, retinal motion, scotoma, SLO.

INTRODUCTION

Vision is a complex sense, encompassing the ability to perceive detail (acuity), color, and contrast and to distinguish objects [1]. These capacities can diminish naturally with age. More than three million Americans suffer with low vision. Among all the diseases related to low vision, age-related macular degeneration (AMD) is the most common, primarily affecting people over the age of 60. With macular degeneration, a scotoma, an isolated retinal area of diminished vision, may appear in the central visual field. The visual system of a person with a central scotoma from a sensory deficit is thought to choose a preferred eccentric retinal area to perform the visual tasks that the nonfunctioning fovea used to perform [2]. The notation of preferred retinal locus (PRL) has been developed for the measurement of the location and extent of these eccentric retinal areas.

The movement of PRLs has been the subject of a number of investigations in low-vision research [3–4]. Previous studies of eccentric PRLs due to a central scotoma have documented that the PRLs for fixation are significantly larger than foveal PRLs for fixation [5]. In the measurement of the movement of PRLs, the scanning laser ophthalmoscope (SLO) is the only instrument that one

can use to determine the retinal location of images of visual targets [2,6]. The SLO employed for this study, manufactured by Rodenstock Instruments (Munich, Germany), has a maximum resolution of about 2 min of arc (1 min of arc = 1/60 of a degree) for measurement of retinal dimensions and the positioning of targets [2]. However, the major limitation of using the SLO for eye-movement tracking has been the need for researchers to use rigid manual registration to analyze the motion from SLO image sequences. Frame-by-frame manual registration has been used for determination of the retinal movement every 1/30 of a second [5]. Because of the tedious and time-consuming nature of this approach, this registration was typically only applied to a small fraction of the total number of frames. Accurately tracking the eye motion over long intervals of time with manual techniques is simply impractical. Several research groups have investigated the use of passive digital image-registration techniques to measure motion between image frames [7–8]. However, these techniques have not proven robust in evaluating long periods of retinal motion. Hammer et al. reported very good results in tracking retinal motion by employing an active, high-speed, hardware-based tracker integrated with an SLO [9]. For many applications, whether the improvement in accuracy made possible by such hardware-based trackers is worth their high cost is still unclear. Thus, for the improvement of the efficiency and accuracy of the analysis of long periods of retinal motion, an automatic passive system for tracking motion from SLO images is needed.

METHODS

Image Acquisition

The SLO with graphics capabilities obtains 640- × 480-pixel retinal images at a rate of 30 frames/s with an invisible infrared laser (780 nm) and scans graphical stimuli images (e.g., a fixation target) onto the retina with a modulated visible red-light laser (633 nm) [2,6–7]. We used two types of visual stimuli (5 × 3 and 5 × 5 array of white crosses on a black background) to obtain the retinal motion images. Subjects were given this instruction: "Fixate on each successive fixation target across a row and then go to the first target on the next row, similar to reading each word on a line of a page. I will tell you when you should go to the next fixation target. While fixating the target, keep your eye as still as possible on the target." The fixation pattern is illustrated in **Figure 1**. **Figure 2** represents a typical frame from an SLO retinal image.

Image Processing

SLO images typically exhibit some noise and distortion that are a function of the age and physical condition of the eye. The most important factor is the mild opacities (sometimes called mild cataracts) that happen with age or other physical conditions. These opacities cause the laser light to scatter, introducing noise into the retinal image. We had to minimize the effects of this noise and distortion before proceeding with the motion tracking. All image processing and tracking techniques were implemented with the use of MATLAB™ (The Math-Works, Natick, MA).

One of the common characteristics of SLO retinal images is low contrast. To compensate for this low contrast,
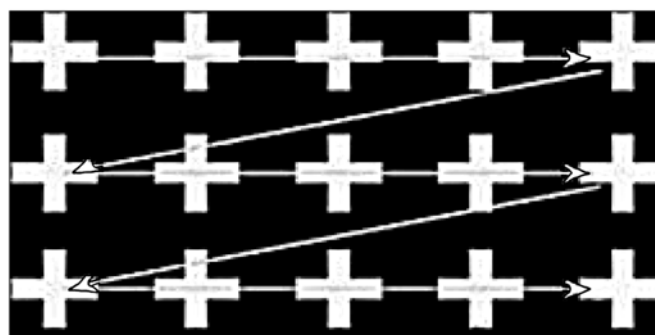

Figure 1.
5 × 3 stimuli used in scanning laser ophthalmoscope experiments. Arrows represent progression of stimuli subject followed.
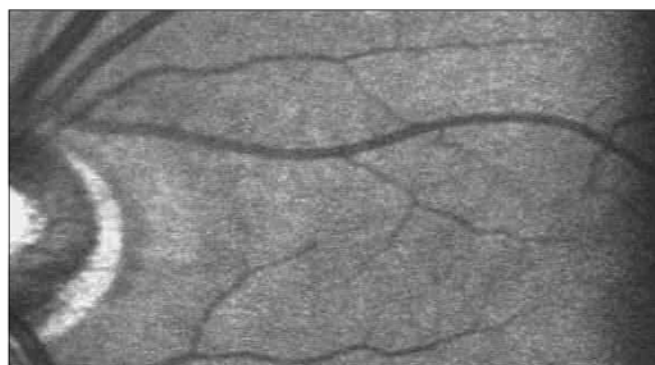

Figure 2.
Scanning laser ophthalmoscope retinal image.

we applied the following contrast stretch algorithm to every frame:

$$I_0(I) = \frac{I - I_{min}}{I_{max} - I_{min}} \ , \qquad (1)$$

where I is the original image intensity, I_{min} and I_{max} are the minimum and maximum intensity values in the original image, and $I_0(I)$ is the new image intensity [10]. In consideration of the large number of frames to be processed and the speckle/shot noise properties [11] in the SLO image sequences, median filtering was applied to every SLO image frame.

In typical SLO image sequences, the visualization of the retinal vessel structures in a subset of the frames cannot be improved by contrast stretching and median filtering. To provide more clearly visible structures for use in motion tracking, we used segmentation to extract the desired vessel structures from the images. We obtained the segmentation masks (M_1, M_2, and M_3) by rotating two-dimensional (2-D) Gaussian filters by 0°, 60°, and 120°. Our purpose in applying segmentation masks to these images was to enhance the image quality for motion analysis. We obtained the segmented images by convolving these masks with the original image. I_1, I_2, I_3 are the segmented images we obtained by convolving the Gaussian filter at 0°, 60°, and 120°, respectively, with the original image:

$$I_1(x,y) = I_0(x,y)**M_1(x,y) = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} I_0(x-k_1, y-k_2)M_1(k_1,k_2)$$

$$I_2(x,y) = I_0(x,y)**M_2(x,y) = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} I_0(x-k_1, y-k_2)M_2(k_1,k_2)$$

$$I_3(x,y) = I_0(x,y)**M_3(x,y) = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} I_0(x-k_1, y-k_2)M_3(k_1,k_2) \ ,$$

$$(2)$$

where the "**" symbol represents a convolution operator; (x,y) represents the pixel position; $I_0(x,y)$ is the original image value at pixel (x,y); and $I_1(x,y)$ is the segmented image value at pixel (x,y) after convolving the original image with the segmentation mask with Gaussian filter at 0°, and $I_2(x,y)$ with the segmentation mask with Gaussian filter at 60°, and $I_3(x,y)$ with the Gaussian filter at 120°.

The filtered image was calculated with

$$I(x,y) = \max(I_1(x,y), I_2(x,y), I_3(x,y)) \ , \quad (3)$$

where $I(x,y)$ is the new filtered image value at pixel (x,y).

Optimal thresholds that minimized the total squared error between the filtered and original images were then applied. Typical results are shown in **Figure 3**.

Motion Tracking

Block matching, one of the most standard motion-tracking methods, was employed in this study. In a generic block-matching algorithm, an $M \times N$ block of pixels is defined in frame k. This block is used to determine the movement between frame k and frame $k + 1$ by finding the best match for this block from a search area that surrounds this block in frame $k + 1$.

The correlation coefficient between blocks, r, was used to find this best match:

$$r = \frac{\sum_i \sum_j (X(i,j) - \bar{X}) \cdot (Y(i,j) - \bar{Y})}{\sqrt{\sum_i \sum_i (X(i,j) - \bar{X})^2 \sum_i \sum_i (Y(i,j) - \bar{Y})^2}} \ , \qquad (4)$$
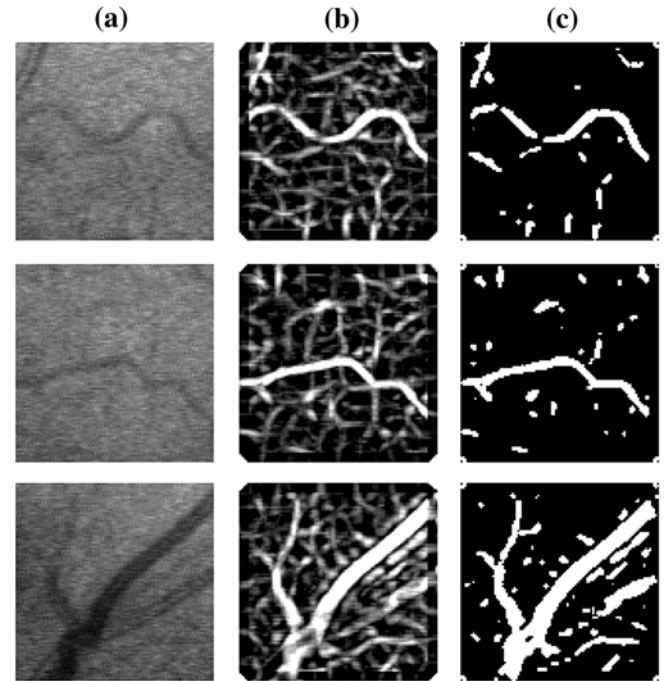

Figure 3.
Enhanced effect of Gaussian filters on retinal images: **(a)** original images, **(b)** Gaussian-filtered images, and **(c)** thresholded images.

where $X(i,j)$ and $Y(i,j)$ are the intensity values at the coordinate (i,j) of two blocks, and

$$\bar{X} = \frac{\sum_i \sum_j X(i,j)}{M \cdot N} \ , \qquad (5)$$

$$\bar{Y} = \frac{\sum_i \sum_j Y(i,j)}{M \cdot N} \ . \qquad (6)$$

The new position of the block $< \overset{\prime}{i}, \overset{\prime}{j} >$ at frame $k+1$ was calculated with

$$< \widehat{i}, \widehat{j} > = \max_{i,j} \{ r[X(i,j),Y(i,j)] \ \}. \qquad (7)$$

Two phases were used in this motion-tracking study. In the first phase, we performed a low-temporal resolution tracking to help select appropriate image features to track. In the second phase, with the aid of the feature points selected in the first phase, we tracked the motion between every frame. This two-phase process helped to shorten the overall time required to process the data.

In the first phase, the operator was asked to select two "feature" points at frame 50. These points were selected at distinct image features that were deemed easiest to track (e.g., vessel bifurcations). Tracking a single point allows only vertical and horizontal movement to be measured. Tracking two points permits torsional movement and image magnification to be measured as well. These additional measurements help identify errors introduced into the tracking measurements. The sources of these errors include head movement, and two errors that are associated with the unique characteristics of SLO imaging: (1) aliasing and (2) changes in image magnification. For example, torsional movements of the eye normally do not exceed 10°. Thus torsional movements in excess of 10° may indicate head movement, and the suspect frames can be flagged for manual inspection. One can detect changes in image magnification by monitoring the distance between the two tracked points in each frame. This requires considering the trapezoidal distortion of the SLO laser beam raster. This distortion, as much as 10 percent, affects only the horizontal dimension and is present because an off-axis mirror in the SLO

projects the raster onto the final concave mirror before the beam exits the device [12]. Thus frame-to-frame changes in the horizontal distance between the two tracked points in excess of 10 percent could indicate either a change in image magnification or an aliasing artifact. Again, the suspect frames can be flagged for manual inspection.

We used two 40×40 blocks, with each feature point at their centers, to measure the horizontal, vertical, and torsional movement between frame k and frame $k + 50$, using the block-matching algorithm with, initially, a 40×40 search region. The frames that were tracked (i.e., every 50th frame) are called *key frames*. Tracking at rates greater than every 50th frame increased the processing time with a diminishing yield on improved tracking accuracy. The tracking of key frames progressed through the video sequence automatically, if the maximum correlation coefficient from block matching was greater than the operator-defined threshold (typically 0.8). If the maximum correlation coefficient fell below the threshold, a larger 80×80 search region was automatically employed. If the coefficient was still less than the threshold, the operator was required to select two new points at frame $k + 50$ to create two new blocks for block matching of subsequent key frames. In this case the operator had to manually measure the movement between frame k and frame $k + 50$.

We used the motion data and point positions for key frames obtained in the first phase in the second phase to track the motion between every frame. In this second phase, the motion was tracked for the 50 frames around each key frame. If the correlation coefficient fell below the threshold in this second phase, the motion in that frame was flagged for later verification and possible manual registration.

RESULTS

The experimental testing of this technique used two groups of subjects, as detailed in **Table 1**. The first group consisted of 25 normal subjects with functional fovea, and the second group consisted of 3 relatively elderly low-vision subjects with nonfunctional fovea.

Tracking Results

Figure 4 shows a typical retinal movement of a Group 1 (i.e., normal) subject for the 5×3 stimuli test. In

Table 1.
Description of test subjects.

Group	Age	Gender	Vision	Number of Subjects
1	20–85	Male/Female	Normal	25
2	70–85	Male	Low	3

a 5 × 3 stimuli test, the ideal retinal movement should have a stair-step appearance, five steps in the x- (i.e., horizontal) direction and three steps in the y- (i.e., vertical) direction. The translational movements in **Figure 4** indicate that the fovea of this subject moves in accordance with the spatial location of the visual stimuli.

Figure 5 represents the visual motion patterns of Group 2 (i.e., low-vision) subjects. This subject has AMD, which produced a central scotoma in his fovea area. The PRL is used to perform the visual tasks that the now nonfunctioning fovea used to perform. Although a stair-step appearance does exist in **Figure 5**, distinguishing the five steps in the horizontal direction and three steps in the vertical direction is difficult. Subjects with functional fovea usually are able to keep the fixation target within a retinal area 2° or smaller [3]. However, subjects with low vision due to a central scotoma have eccentric PRLs within a retinal area as large as 9° [5,13–14]. Results such as those in **Figure 5** will be the subject of future research to analyze the PRL motion patterns of people with low vision.

Assessment of Accuracy

To evaluate the accuracy of this technique, we conducted two case studies to compare the results obtained from manual registration methods with the automated technique.

Case Study 1: Young Subject with Normal Vision

To make the comparison under the most challenging conditions, we selected a sequence of frames associated with a time period during which movement took place from one visual stimulus to another. For this case, the first stepwise shift in horizontal direction appeared at frame 330, so the 20 frames around this frame were selected. Three volunteers were asked to use a manual registration technique (i.e., use a computer's mouse to position a cursor over the target point—vessel bifurcation) to track the movement occurring within these 20 frames. These manual segmentation results, together with those from the automated technique, are shown in **Figure 6**. The range
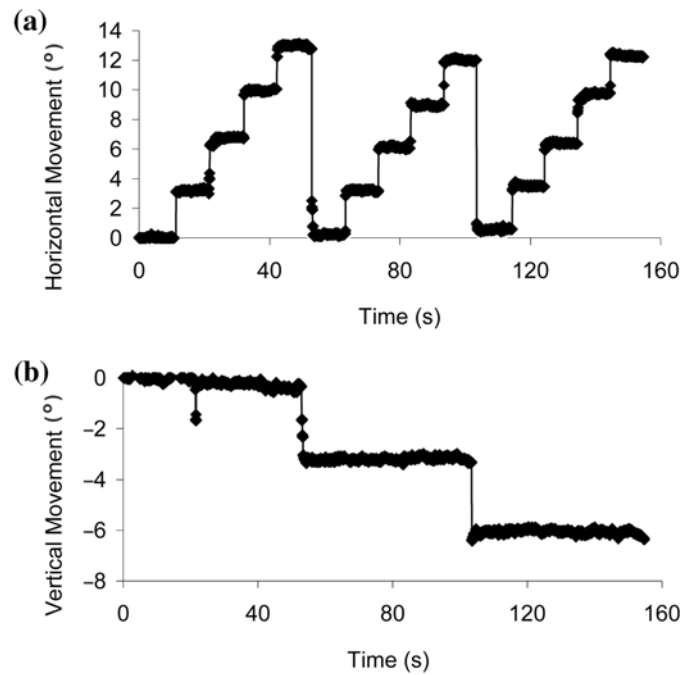
Figure 4.
Retinal movement tracking results in young subject group with normal vision for 5 × 3 test for **(a)** horizontal and **(b)** vertical movement.
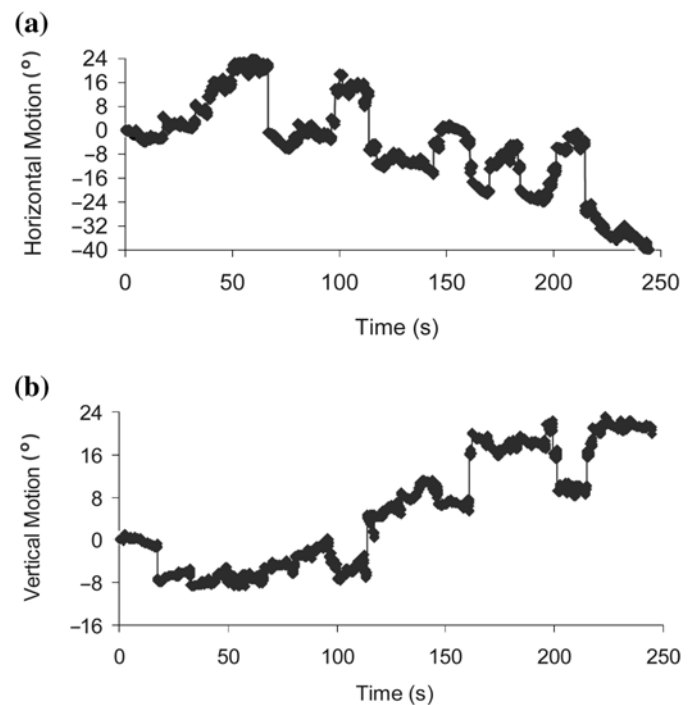
Figure 5.
Retinal movement tracking results in elderly subject group with normal vision for 5 × 3 test for **(a)** horizontal and **(b)** vertical movement.

4 pixels in every frame except frame 330, where the first stepwise shift occurred.

The data reveal almost perfect agreement between the two techniques, with the largest difference occurring at frame 329. A large horizontal movement occurred during the acquisition of frame 329. The video sampling rate of the SLO (30 frames/s) was not fast enough to capture the velocity of this eye movement (several hundred degrees per second). This resulted in an aliasing effect in frame 329 that made it impossible for either the manual or the automatic tracking techniques to accurately determine the relative motion. The results for vertical eye movement are shown in **Figure 7**. This figure shows that in the majority of the frames, negligible difference existed between the measurements made by the two techniques, with the maximum difference in any frame being fewer than 3 pixels.
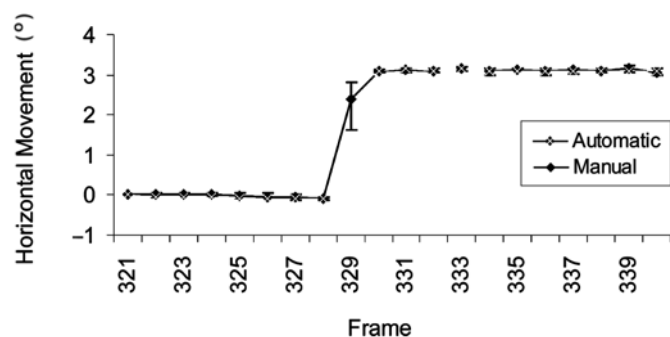

Figure 6.
Comparison of manual registration with automatic tracking for horizontal movement in Case 1 (subject with normal vision).
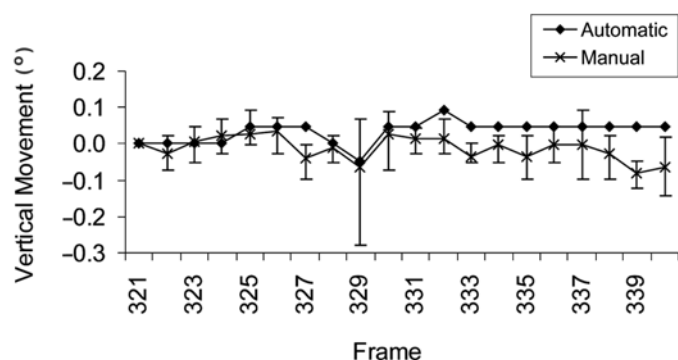

Figure 7.
Comparison of manual registration with automatic tracking for vertical movement in Case 1.

Case Study 2: Elderly Subject with Low Vision

For this subject, we selected 51 frames (frames 900 to 950) for analysis using the same selection criteria as for Case 1. **Figure 8** shows the comparison results for the horizontal, vertical, and torsional movements. This figure illustrates the high level of agreement that was achieved between the manual and automatic results. The range in the manual registration measurements was fewer than 11
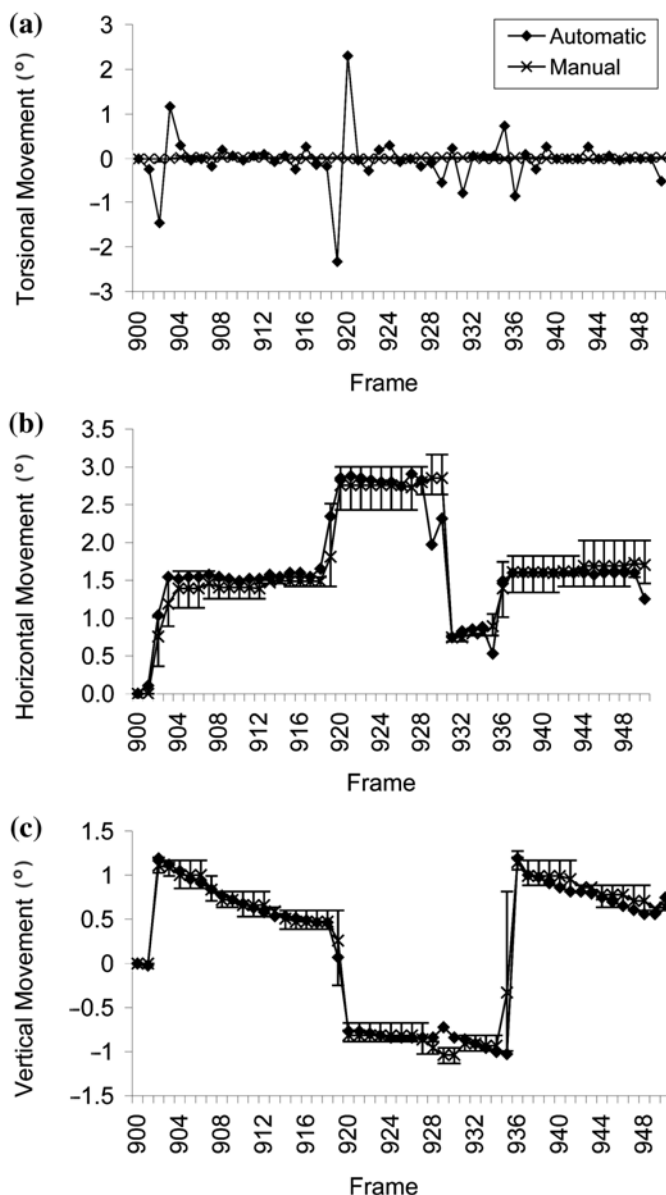

Figure 8.
Comparison of manual registration with automatic tracking for (**a**) torsional movement, (**b**) horizontal movement, and (**c**) vertical movement in Case 2 (subject with low vision).

pixels. Once again, the frames where the greatest discrepancies occurred were those that were captured during fast eye movements where the SLO's frame rate was not sufficient to avoid aliasing artifacts. Such an artifact is clearly seen in frame 935 (**Figure 9**). These artifacts can be detected by the automatic technique through measurements of the distance between the two points that are tracked in each frame, as is illustrated in **Figure 10** for Case 2. Without artifacts present, the distance between these points should remain relatively constant.

DISCUSSION

The technique developed provides a tool for researchers of low vision to analyze the entire record of motion of the retina while subjects perform visual tasks. The results of comparisons between this automated technique and the manual technique on both normal and low-vision subjects indicate that the automated technique can accurately and efficiently tracking eye motion. The time required for the automated analysis ranged from approximately 2 s/frame for normal subjects to 3 s/frame or low-vision subjects.

With the accurate motion data from this technique, researchers of low vision can begin to investigate the motion patterns of PRL in people with low vision. For example, studies could be performed to determine whether PRLs are able to make efficient and effective eye movements while reading, tracking objects, or locating
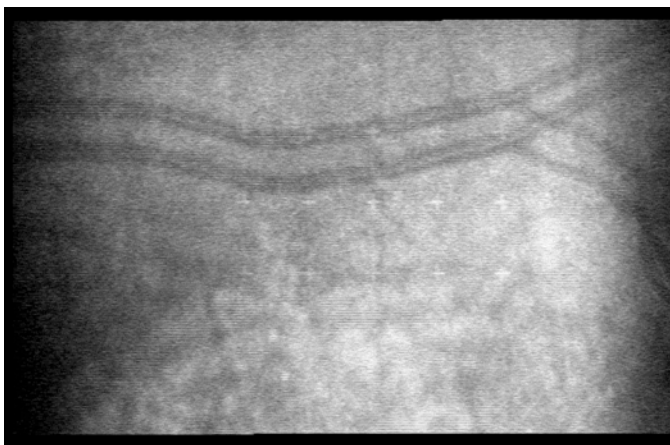
Figure 10.
Distance between two tracked points in given frame for Case 2.

objects in the person's visual field. Only a comparison between subjects with developing PRLs will indicate which one or more of these possible mechanisms guide the characteristics and abilities of PRL for subjects with a central scotoma [2,13–15].

CONCLUSION

This technique, which enables a rapid and reliable analysis of SLO image sequences in low-vision research, will help researchers investigate the visual patterns of people with low vision. The results may lead to the development of rehabilitation therapies.

REFERENCES

1. Guyton AC, Hall JE. Human physiology and mechanisms of disease. 6th ed. Philadelphia (PA): W.B. Saunders Company; 1997. p. 400–15.

2. Schuchard RA, Fletcher DC. Preferred retinal locus: A review with applications in low vision rehabilitation. Ophthalmol Clin North Am. 1994;7(2):243–56.

3. Fletcher DC, Schuchard RA. Preferred retinal loci relationship to macular scotomas in a low vision population. Ophthalmology. 1997;104:632–38.

4. Fletcher DC, Schuchard RA, Watson G. Relative locations of macular scotomas near the PRL: Effect on low vision reading. J Rehabil Res Dev. 1999;36:356–64.

5. Schuchard RA, Cooper S, Lakshminarayanan V. Time series analysis of PRL movement during fixation. OSA Tech Digest Series. 1996;1:20–23.

6. Webb RH, Hughes GW, Delori FC. Confocal scanning laser ophthalmoscope. Appl Opt. 1987;26:1492–99.

7. Peli E, Augliere RA, Timberlake GL. Feature-based registration of retinal images. IEEE Trans Med Im. 1987;6:272–78.

Figure 9.
Frame 935 of scanning laser ophthalmoscope from Case 2, illustrating aliasing artifacts associated with frame rates too low to capture fast eye movements.
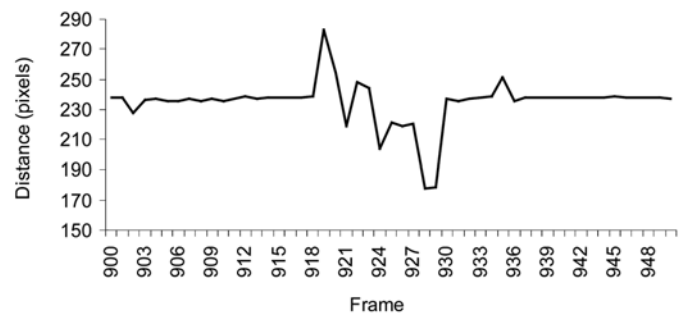
8. Wornson DP, Hughes GW, Webb RH. Fundus tracking with the scanning laser ophthalmoscope. Appl Opt. 1987; 26(8):1500–504.

9. Hammer DX, Ferguson RD, Magill JC, White MA. Image stabilization for scanning laser ophthalmoscopy. Opt Expr. 2002;10(26):1542–49.

10. Smith MJC, Docef A. A study guide for digital image processing. 2nd ed. Riverdale (GA): Scientific Publisher, Inc.; 1999. p. 333–34.

11. Bueno JM, Campbell MC. Confocal scanning laser ophthalmoscopy improvement by use of Mueller-matrix polarimetry. Opt Lett. 2002;27(10):830–32.

12. Timberlake GT, Sharma MK, Gobert DV, Maino JH. Distortion and size calibration of the scanning laser ophthalmoscope (SLO) laser-beam raster. Optom Vis Sci. 2003; 80(11):772–77.

13. Vasudevan R, Rhatak AV, Smith JD. A stochastic model for eye movements during fixation on a stationary target. Kybernetick. 1972;11:24–31.

14. Schuchard RA, Lim JM. Exploring the characteristics of secret eye movements during fixation: A new approach of chaotic time series. In: Lakshminarayana V, editor. Basic and clinical applications of vision science: The Professor Jay M. Enoch Festschrift volume. Amsterdam: Kluwer Publishing Group; 1997. p. 177–80.

15. Robinson DA, Gordon JL, Gordon SE. A model of the smooth pursuit eye movement system. Biolog Cybernet. 1986;55:43–57.