Vision Research 103 (2014) 75–82

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

A method for detecting optokinetic nystagmus based on the optic flow of the limbus

Jason Turuwhenua [a,b,*], Tzu-Ying Yu [b], Zan Mazharullah [a], Benjamin Thompson [b]

[a] Auckland Bioengineering Institute, University of Auckland, New Zealand
[b] Department of Optometry and Vision Science, Faculty of Medical and Health Sciences, Auckland, New Zealand

ARTICLE INFO

Article history:
Received 17 June 2013
Received in revised form 11 July 2014
Available online 21 August 2014

Keywords:
Optokinetic nystagmus
Optic flow
Video-oculography
Random-dot-kinetogram

ABSTRACT

Optokinetic nystagmus (OKN) is the sawtooth movement of the eye elicited when an observer views a repeated moving pattern. We present a method for identifying the presence and direction of OKN in recordings of the eye made using a standard off-the-shelf video-camera or webcam. Our approach uses vertical edge detection to determine the limbus/iris boundary, and we estimate the velocity of the edge using Lucas–Kanade optical flow. Heuristic rules are applied to identify saccadic velocity peaks from the resulting velocity signal. The normalized average of the resulting peaks is used to estimate the presence/direction of OKN. Our preliminary testing with six participants observing global motion stimuli with full or partial coherence yields an accuracy of 93% which compares favorably to the performance of an experienced human observer (98% accuracy). Additional tests using high contrast, square-wave gratings show that performance of the technique is consistent at stimulus speeds of 5 and 10 deg/s and that OKN is not reported by the algorithm when participants view stationary stimuli.

1. General Introduction

Optokinetic nystagmus (OKN) is an involuntary eye movement elicited by continuously moving patterns (Büttner & Kremmyda, 2007, chap. Smooth pursuit eye movements & optokinetic nystagmus). The motion consists of a smooth phase (which occurs as the eye tracks a target) followed by a saccade in the opposite direction (resetting event), that allows the eye to refixate on a new feature of the stimulus. As shown by Fig. 1A, the characteristic movement is a sawtooth in the displacement versus time graph.

OKN only occurs when the motion of a continuously moving pattern is detected by the visual system and therefore the presence or absence of OKN can provide an objective measure of visual function (Han et al., 2011). The use of OKN to assess visual function may be particularly useful when testing observers who are unable to provide reliable behavioral or verbal responses such as young children (Yu et al., 2013). Recently, it has been proposed that the visual pathway responsible for motion processing is particularly vulnerable to developmental disorders (Braddick, Atkinson, & Wattam-Bell, 2003; Grinter, Maybery, & Badcock, 2010; Macintyre-Béon et al., 2010). In this context, the ability to accurately assess the presence and direction of OKN in response to a range of dynamic visual stimuli such as global motion (Newsome, Britten, & Movshon, 1989) may provide a useful tool for the assessment of neurological function, particularly in young observers.

Computer based detection of OKN is attractive, particularly for less visible stimuli and for measurements that require a considerable number of trials. Moreover, objective measures can avoid any bias that may exist in a clinical judgement. The standard quantitative methods for detecting OKN are: electro-oculography, which entails the use of electrodes to measure the change in electrical potential as the eye (a strong dipole) rotates (Marg, 1951); corneal reflection techniques, in which near infra-red (IR) light is reflected from the cornea (Young & Sheena, 1975); and video-oculography, in which the eye's position is tracked using digital video (Duchowski, 2007). Video-oculography is attractive because it is a non-invasive approach that can be implemented using low-cost off-the-shelf hardware (Li, Babcock, & Parkhurst, 2006). Commercially available video-oculography devices are typically IR based, and generally require fixed mounting of camera on a chin-rest or the head to maintain a fixed, unobstructed view of the eye. Eye tracking under these conditions is a well characterized problem. Daugman's classic iris recognition work introduced an integro-dif-

* Corresponding author at: Auckland Bioengineering Institute, University of Auckland, New Zealand.
E-mail addresses: j.turuwhenua@auckland.ac.nz (J. Turuwhenua), sandy.yu@auckland.ac.nz (T.-Y. Yu), zmazharullah@gmail.com (Z. Mazharullah), b.thompson@auckland.ac.nz (B. Thompson).

http://dx.doi.org/10.1016/j.visres.2014.07.016
0042-6989/

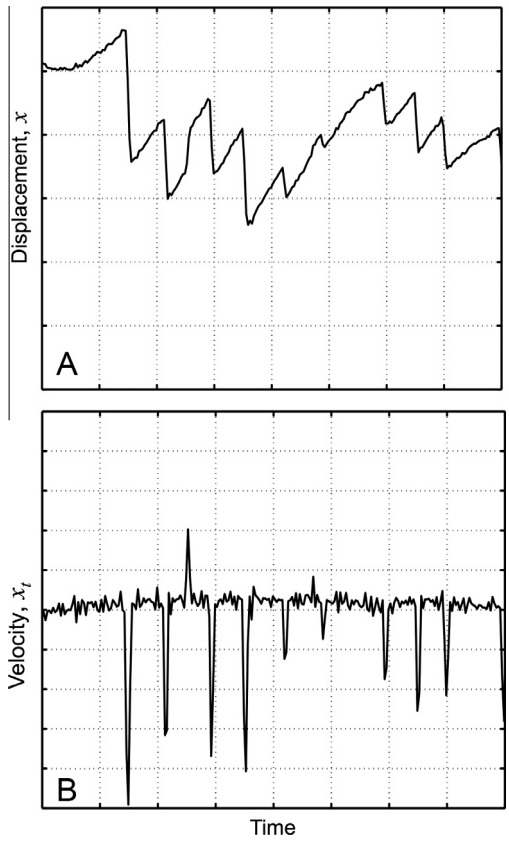

Fig. 1. Signals measured from a commercially available eye tracking system. (A) The sawtooth displacement versus time signal typical of OKN. (B) The velocity versus time signal (calculated here by differentiating displacement), showing the "delta function-like" features we aim to measure and detect. Whilst differentiation of the displacement has been used to illustrate the form of the velocity signal, it is noted that our method estimates velocity *directly* from video footage.

ferential operator, that maximized at boundaries of appropriate shape (*e.g.*, circular regions such as the pupil/limbus, and parabolic regions like eyelids) given the pupil center position (Daugman, 1993). Yuille, Hallinan, and Cohen (1992) introduced deformable eye templates, based on fitting mathematical models to image data (using energy minimization). Chow and Li (1993) also used a deformable template approach, in which parabolas were fitted after the (circular) pupils were detected by Hough transform. More recently, Li, Babcock, and Parkhurst, 2006 described the starburst method. In this approach, rays emanate from an estimate of the pupil center, and searches are performed along these paths to detect intensity edges, namely the pupil edge or the limbus. These edges serve as additional seeds for further additional edge searches. The random sample consensus (RANSAC) algorithm (Fischler & Bolles, 1981) robustly rejects unlikely candidate edge points found by this approach, by determining an ellipse for the pupil/limbus that fits a majority of feature points within a given threshold error.

In young subjects, head-mounting of IR sources and/or cameras is not well tolerated (Franchak, Kretch, Soska, & Adolph, 2011), the use of chin rests is not practical and the pre-measurement calibration routines that are required by many eye tracking systems require a level of cooperation that is unlikely to be achieved. This often precludes the use of standard eye tracking equipment with young children and indicates the need for remotely positioned camera monitoring of the eye, as well as simpler calibration if possible. It is, furthermore, natural to ask whether specialized equipment is required to provide the eye movement data required for the objective detection and quantification of OKN, or whether low cost off the shelf consumer video is sufficient.

A key issue that is faced when using low-cost video equipment for recording OKN is illustrated in Fig. 2B which shows a representative RGB video frame of an eye. The contrast between the pupil and iris is poor: severely reducing the accuracy of any pupil tracking algorithm, particularly so for darker irises. Furthermore, the image contains reflections and eyelashes that will further confound efforts to determine a precise position-based displacement signal for the eye. On the other hand, the image shows a readily identifiable limbal edge, which represents a viable feature for tracking. Limbal detection and tracking methods using parametric modeling (Iskander, 2006) and light integration over the entire eye have been described (Abe, Ohi, & Ohyama, 2007). Furthermore, methods are available that may account for the presence of confounding features such as reflections and eyelashes (Kong & Zhang, 2001a; Kong & Zhang, 2001; He, Tan, Sun, & Qiu, 2008) within such images. Our aim was to build on this previous work to develop a reliable and robust method for the detection of OKN within RGB video footage of the eye that could provide a basis for the objective assessment of OKN in both clinical and research environments.

Our novel approach attempts to directly estimate the horizontal *velocity* of the limbal edge $\overline{V}_x(t)$ over a video sequence using a method based on Lucas–Kanade optical flow (Barron, Fleet, & Beauchemin, 1994). It is noted that whilst Lucas–Kanade *feature tracking* has been used to track head movement (Morris, Blenkhorn, & Zaidi, 2002), and the torsion of the iris (Lee, Choi, & Park, 2007), we are not aware of optical flow being used for the analysis of OKN. This optic flow based approach represents a significant departure from model based/feature tracking methods which critically rely on precise localization of specific features from one frame to the next in order to determine displacement. Here we present the technique itself and behavioral results from adult observers obtained using both a low cost off the shelf consumer video camera and web-cam.

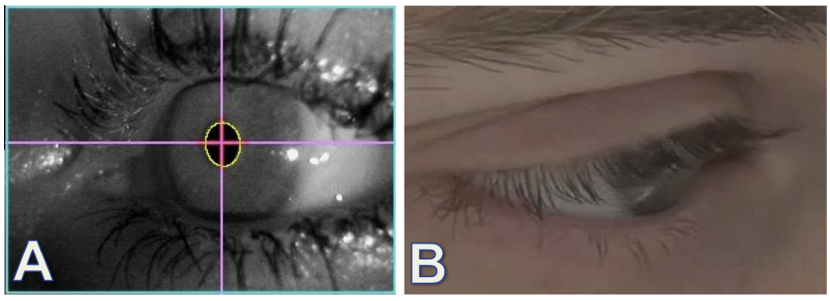

Fig. 2. (A) Image from a standard IR eye tracking system and (B) Video footage of the eye, typical of our experiments.The contrast between the pupil/iris edge is low. Reflections and contamination from eyelashes are also present. At the same time, we observe that the limbus often presents a good contrast edge.

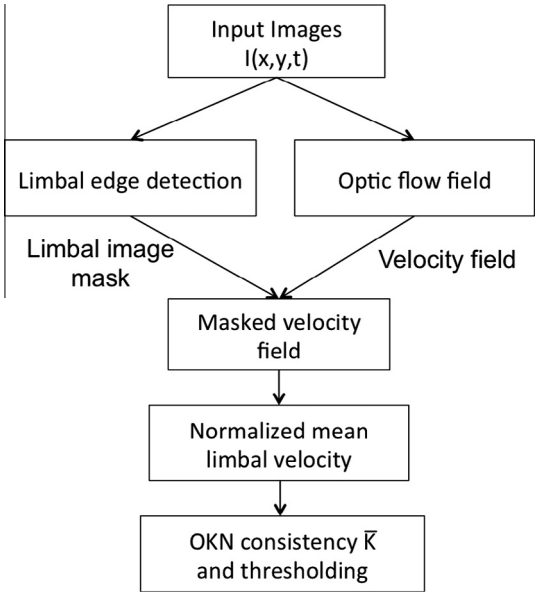

Fig. 3. The major elements of the algorithm we have implemented for determining the presence of OKN from the limbal velocity signal. The strategy is to determine optic flow over an image containing the eye, which we further mask with a binary estimate of the limbal edge. By doing this, we aim to select only those velocity vectors corresponding to the limbal edge (the masked velocity field). The presence of OKN is determining by calculating and thresholding the consistency measure $\overline{K}$, the normalized average of saccadic peaks in the resulting signal.

2. General methods

2.1. Algorithm overview

The process we have adopted is summarized by Fig. 3. The first step is to convert RGB video footage into a stream of reduced/gray-scale images $I(x,y,t)$ (here (x,y) refers to pixel locations and t is the frame number). In this work, we have adopted the use of an image pyramid approach to perform the data reduction transformation (Camus & Wildes, 2002; Hsu, Abdel-Mottaleb, & Jain, 2002). The subsequent method is as follows: (1) we use the Lucas–Kanade flow to determine a velocity field $V(x,y,t)$ over an *entire* image (for each given time t). We then (2) select $V(x,y,t)$ in regions of limbus edge (found using standard edge detection algorithms) in order to improve the signal-to-noise ratio. The next step is to (3) take the vector average of the masked velocity field, and use the horizontal component to yield the limbal signal $\overline{V}_x(t)$. Finally, given a suitable interval of $\overline{V}_x(t)$ containing OKN we (4) apply a heuristic analysis to the signal to detect the "delta function-like" peaks in the velocity signal associated with the OKN reset event as illustrated by Fig. 1B. By determining the normalized average of these peaks (which we refer to as $\overline{K}$) and thresholding the absolute value $|\overline{K}|$, we obtain a measure that indicates the presence of consistent saccadic movements and the direction of these saccades. Lower absolute $\overline{K}$ values ($|\overline{K}| < 1$) indicate saccades in both rightward and leftward directions and/or small saccadic eye movements. Higher absolute $|\overline{K}|$ values ($|\overline{K}| < 1$), on the other hand, indicate pronounced saccadic eye movements occurring in consistent direction as would be expected from a brisk OKN response.

2.1.1. Limbal edge detection

Edge detection is used to determine a binary edge map $q(x,y,t)$ identifying likely limbus pixels within each image of the eye $I(x,y,t)$. By using this edge map to select out the relevant regions of the optic flow field, we aim to increase the specificity of our velocity determination. In this work we identify the limbus as a falling vertical edge (*i.e.*, going from sclera to pupil) within the image $I(x,y,t)$ by applying the Prewitt operator (Prewitt, 1970). The resulting edge map is binarized using hysteresis thresholding (Condurache & Aach, 2005). Connected regions below an empirically determined weight, as well as regions connected to the image border are removed.

2.1.2. Optic flow estimation

Our method departs from standard eye tracking methods by our use of the optic flow over the region of the eye to identify motion due to saccadic eye movements. The optical flow equation is (Horn & Schunck, 1981; Barron, Fleet, & Beauchemin, 1994),

$$I_t = -\nabla I(x,y,t)\cdot V \tag{1}$$

where $I_t = \partial I/\partial t(x,y,t)$ and $\nabla I = (\partial I/\partial x, \partial I/\partial y, \partial I/\partial z)$ are the time and spatial gradients of the image respectively, and $V(x,y,t)$ is the estimated (vector) velocity field. The assumption of Lucas–Kanade optical flow is that V is constant over a small region around the pixel (x,y), which yields a linear system that can be solved for $V(x,y,t)$ using a standard least squares approaches (Björck, 1996).

Our modification is to estimate a limbal velocity $\overline{V}(t)$ from the optic flow derived velocity $V(x,y,t)$ by weighting this velocity with the masking function $q(x,y,t)$. The masking function is designed to select only velocities in the region of the limbus. In particular, we set it to $q(x,y,t) = 1$ when a pixel is determined to belong to the limbus, and $q(x,y,t) = 0$ otherwise. In practice $q(x,y,t)$ is the binary edge map introduced in Section 2.1.1. The equation for $\overline{V}(t)$ is therefore:

$$\overline{V}(t) = \begin{cases} \frac{\sum_I q(x,y,t)V(x,y,t)}{\sum_I q(x,y,t)} & \text{if } \sum_I q(x,y,t) > 0 \\ \text{undefined} & \text{if } \sum_I q(x,y,t) = 0 \end{cases} \tag{2}$$

This equation is defined when limbal pixels have been detected in a frame, and in this case $\overline{V}(t)$ is simply the average velocity of pixels classified as belonging to the limbus. Otherwise, $\overline{V}(t)$ is undefined. We also note that the signal described here is a vector quantity, but in further analysis we use only the horizontal component $\overline{V}_x(t)$.

2.1.3. Analysis of the limbal velocity signal

The next step of our analysis is to decide whether OKN is present within the recorded section of data $\overline{V}_x(t)$, the horizontal component of $\overline{V}(t)$, and if so, the direction of the saccadic eye movements. Because the raw velocity signal $\overline{V}_x(t)$ contains velocity peaks due not only to OKN, but spurious eye movements, blinks, and signal noise that need to be removed we have taken a number of steps to remove unwanted information. These are described in detail below, but in summary, velocity peaks are excluded from analysis: (1) if peaks attain a maximum velocity below an empirically derived threshold value σ_V pixels/frame, (2) if the generated peak did not cross the zero velocity axis, and (3) if subsequent detected peaks are too close temporally (with threshold of σ_t frames) to be due to OKN. If there are surviving peaks, we average their maximum values (multiplied by the normalization value $1/\sigma_{\overline{V}}$ an empirically derived velocity threshold). If peaks are weak (*i.e.*, near threshold), or equally distributed between positive and negative then a low averaged value will result. Conversely, if peaks are consistently positive or negative, the measure will exceed threshold, and the sign will correlate with the direction of the saccadic eye movements occurring due to OKN. Finally, the entire velocity record is discarded if there are not enough remaining peaks to indicate consistent saccadic movements indicative of OKN. The detailed algorithm is given below.

1. Detect all velocity maxima/minima within the given interval of $\overline{V}_x(t)$ signal and label them $Q(j)$, where $j = 1 \ldots M$. A maxima/minima is defined as a point with neighboring points less than/greater than the test point by an absolute threshold of σ_{peak} pixels/frame.
2. Threshold these peaks, *i.e.*, eliminate all peaks $|Q(j)| < \sigma_{\overline{V}}$ where $\sigma_{\overline{V}}$ is an empirically determined (horizontal) velocity threshold. Reset M to the number of remaining peaks.
3. Reject minima/maxima if they do not indicate a reversal in the direction of the eye. A valid velocity peak should in principle, cross the axis defined by zero velocity: the slow-phase and quick-phase occur in opposite directions (and hence are of opposite sign). This rule also assists in identifying "jagged edges" or noise that may occur on the face of a valid velocity peak. Again, reset M to the number of remaining peaks.
4. Reject peaks that are less than a given number of frames apart (σ_t). Again, reset M to the number of remaining peaks. In this work, we used a threshold of 4 frames corresponding to 133–160 ms (for camera frame-rates of 30 Hz and 25 Hz). We note that the expected slow phase interval for a stimulus moving at 10 deg/s was estimated to be greater than 250 ms by Waddington and Harris (2012).
5. Reject a solitary maxima/minima (*i.e.*, with no other maxima of the same sign in the trial). We assume that an isolated peak, is not enough evidence to indicate the presence of OKN. In this case we do not continue to the estimation of normalized average peak velocity and the algorithm stops.
6. Determine the normalized average peak velocity $\overline{K}$. If there are remaining peaks then take $\overline{K}$ as the mean of the velocities after normalization by the velocity threshold $\sigma_{\overline{V}}$ (which was set to 0.5 pixels/frame in this work). If there are no remaining peaks, set the value of $\overline{K}$ to zero.

$$\bar{K} = \begin{cases} 1/(M\sigma_{\overline{V}})\sum_{j=1}^{M} Q(j) & M > 0 \\ 0 & M = 0 \end{cases} \quad (3)$$

In this work we take an absolute value of $\overline{K}$ of less than 1 to indicate the absence of consistent quick phase movements indicative of OKN, whilst a value of $\bar{K} \geqslant 1$ indicates the presence of consistent quick phase movements, and hence OKN. The sign of $\overline{K}$ indicates the direction.

3. Experimental validation

All participants gave full written informed consent, all study protocols were approved by the University of Auckland Ethics Committee and complied with the declaration of Helsinki.

3.1. Experiment 1

3.1.1. Overview

The aims of experiment 1 were; (1) to assess whether the OKN detection algorithm could correctly classify OKN eye movements as being leftwards or rightwards, and (2) to compare the performance of the OKN detector to that of an experienced human observer (TY). For this experiment we used a random dot kinematogram (RDK) with variable motion coherence to elicit OKN. We chose this stimulus as the perception of global motion within such stimuli is thought to rely on dorsal areas of the extrastriate visual cortex that may be particularly vulnerable to abnormal development (Braddick, Atkinson, & Wattam-Bell, 2003; Grinter, Maybery, & Badcock, 2010; Macintyre-Béon et al., 2010). Furthermore, we have recently found that RDKs can elicit reliable slow and fast-phase optokinetic eye movements that can be used to measure motion coherence thresholds in young children who cannot yet provide behavioral responses to psychophysical tasks (Yu et al., 2013).

3.1.2. Stimulus and recording equipment

Following Yu et al. (2013), the RDK consisted of 250 moving white dots (0.5 deg diameter, 8 deg/s speed), presented at 100% contrast on a 16″ cathode ray tube (CRT) display (see the inset of Fig. 4). The stimuli were presented using software developed in MATLAB (Mathworks, Natick, MA) using the Psychophysics Toolbox extensions (Brainard, 1997). The dots were presented in a 8.3° circular window for 8 s and had a limited lifetime whereby each dot had a 25% chance of dying on any given frame and being replotted in a new, random location within the stimulus. Dots that reached the edge of the stimulus aperture were wrapped around. The *coherence* level of the stimulus, *i.e.*, the proportion of dots moving in the same direction versus the total population of dots could be adjusted to vary the strength of coherent motion present in the stimulus (Newsome, Britten, & Movshon, 1989). We tested two coherence levels: (1) full coherence (*i.e.*, 100% of the dots moved in the same direction) and (2) low coherence (between 12% and 15% of all dots moved in concert, whilst the rest moved randomly). The direction of coherent motion was randomized across trials. The low coherence level was chosen to be suprathreshold for our observers (*i.e.*,the global motion direction was still clearly visible) while still allowing us to test the algorithm on eye movements elicited by degraded motion signals.

The 16″ CRT display was viewed from 50 cm. Video footage was collected using a SONY digital high definition camera (HDR-CX7EK, Sony Corporation, Tokyo, Japan), that delivered video consisting of RGB images (1920 × 1080 pixels) at 25 frames per second. The camera was placed to the side of the CRT, and centered on the closest eye. The experimental setup is shown in Fig. 4.

3.1.3. Procedure

Six adults (mean age = 25 years) with normal vision viewed the RDK stimuli binocularly with their head restrained by a chin rest. A fixation point was presented in the center of the CRT screen before and after each RDK presentation trial. Participants were instructed to fixate on the point in between trials and stare at the center of the screen during trials. The fixation point was not presented during a trial.

Footage obtained from the trials was prepared for further processing by manually cropping the video around each eye within every trial. The video for each trial was cropped temporally to include only frames recorded when the participant was viewing the motion stimulus. A total of 115 trials (of 8 s length) were obtained from our participants (73 at full coherence and 42 at

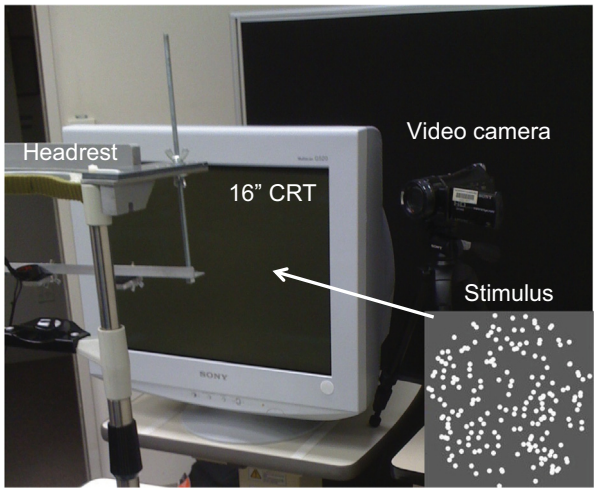

Fig. 4. The experimental setup. The inset shows an example of the RDK stimulus pattern presented to observers.

low coherence). The participants were able to perceive the direction of the signal dots in each trial as indicated by behavioral responses provided during testing.

The velocity estimator and OKN detector were applied offline to each resulting video record. The detector estimated $\overline{K}$. The absolute value of $\overline{K}$ was thresholded to indicate the presence ($|\overline{K}| \geqslant 1$) or absence ($|\overline{K}| < 1$) of OKN. The sign of the measure indicated the direction. The first 17 trials were recorded using the 100% coherence stimuli and were used to calibrate the parameters for the method. Our software was written using MATLAB, and we used an implementation of the Lucas–Kanade optical flow solver written by Dollár (2012). The video footage was also viewed by an experienced human observer (TY) who made a 2-alternative forced choice decision as to whether the OKN response was consistent with a leftwards or rightwards moving RDK. The observer was not aware of the results generated by the detection algorithm when viewing the videos.

3.1.4. Results

We achieved 100% accuracy for our calibration set of 17 trials (9 correct left detections, 8 correct right detections), after appropriate adjustment of all algorithm parameters. Fig. 5 presents a sequence of video frames obtained from the calibration trials that illustrate important aspects of the detection process. The figure shows the estimated limbal region overlayed on grayscale images of the eye. The color of the limbal region correlates with limbal velocity, where cool colors (blue) are low velocities and high velocities are hot (red). Frame 10 shows a resting eye, whilst Frames 54–56 show the occurrence of the OKN reset event (a saccadic eye movement). Frame 26 demonstrates the loss of the limbal edge during an eye blink. These observations are supported by inspection of the velocity traces (*i.e.*, horizontal and vertical) for this sequence presented in Fig. 6A and B. The frame numbers that appear in the bottom left corner of the stills in Fig. 5, correspond directly to the horizontal axis of Fig. 6A and B. The peaks accepted as "OKN-like" are shown boxed. The empirically derived velocity threshold of 0.5 pixels/frame is also shown as a dotted line.

The performance of the detector over all 115 trials was 93%, (54 correct left detections, 53 correct right detections, 8 errors) compared to 98% (53 correct left detections, 60 correct right detections, 2 errors) for the human observer. This difference approached, but did not reach, statistical significance (Chi-Square = 3.76, $p = 0.052$). The performance of the detector for the full coherence trials was 96% (38 correct left detections, 32 correct right detections, 3 errors) compared to 100% (38 correct left detections, 35 correct right detections) for the human observer, a non-significant difference (Chi-Square = 3.06, $p = 0.08$) For the low coherence trials the detector performed at 88% correct (16 correct left detections, 21 correct right detections, 5 errors) compared to 95% correct for the human observer (15 correct left detections, 25 correct right detections, 2 errors), also a non-significant difference (Chi-Square = 1.40, $p = 0.24$). The poorer performance for the low coherence trials is likely due to the weaker global motion signal inducing less reliable OKN responses (Yu et al., 2013).

The distributions of absolute OKN consistency values $|\overline{K}|$ for incorrectly and correctly classified trials are summarized by Fig. 7A and B respectively. For the full coherence trials 3/3 (100%) incorrectly classified trials had K values less than 1 compared to 18/70 (26%) for the correctly classified trials (Chi-Square = 7.7, $p < 0.005$). Similarly for the low coherence trials 5/5 (100%) incorrectly classified trials had K values less than 1 compared to 17/37 (46%) for the correctly identified trials (Chi-Square = 5.0, $p = 0.02$). An inspection of the velocity traces for the incorrectly classified trials (*e.g.*, Fig. 6 panels C–E) indicated that the low $|\overline{K}|$ values were due to velocity peaks being relatively equally distributed around zero. These bi-directional peaks appeared to cancel each other out resulting in a low OKN consistency measure.

To confirm this observation, the video footage of the 8 trials incorrectly classified by the detector were reviewed for a second time by our experienced human observer (TY). Firstly, it was confirmed that bi-directional velocity spikes resulted from actual eye movements visible within the video footage, and were not caused by spurious errors of the method. Secondly, it was found that the recordings indeed contained sequences of OKN in which direction changed.

3.1.5. Discussion

The results of experiment 1 were encouraging as they indicated that judgments of OKN direction generated by the detector were not reliably different from those made by an experienced human observer. Furthermore, classification errors made by the detector were due to irregular or weak OKN responses. For the low coherence trials, the irregular OKN responses (*i.e.*,the occurrence of OKN in two different directions within one trial) was likely due to the presence of noise dots within the stimulus. The noise dots may have happened to generate a motion signal that was sufficiently strong to elicit OKN in a direction different to the signal direction for a short period during stimulus presentation.

3.2. Experiment 2

3.2.1. Overview

Experiment 2 had three aims: (1) to assess whether the OKN consistency measure $\overline{K}$ was affected by stimulus velocity, (2) to assess whether the OKN detector would correctly reject trials during which participants passively viewed a stationary stimulus, and (3) to assess whether the detection technique could be applied to footage obtained using a standard webcam. A high contrast square-wave grating stimulus was used for experiment 2 as this type of stimulus is used routinely to induce OKN in clinical studies.

3.2.2. Stimulus and recording equipment

An IBM P275 cathode ray screen (20″ viewable area, a resolution of 1600 × 1200 and 75 Hz refresh rate) was viewed from 90 cm on which were presented a 100% contrast square wave grating with a fundamental spatial frequency of 0.833 cycles/deg as used by Hyon et al. (2010). Three presentation velocities (0 deg/s, 5 deg/s and 10 deg/s) were used, and the grating always moved right to left. Video footage of the OKN eye responses was obtained using an LogiTech HD Pro Webcam C920 (recording facilitated using the Logitech Webcam Software Version 2.51). The stimuli were coded and presented using the MATLAB Psychophysics toolbox.

Fig. 5. Video frames corresponding to the data shown in Fig. 6A and B. Frame 10: The eye in a resting phase, as indicated by the cool color. Frame 26: The limbus estimate vanishes during an eye blink. Frames 54–56: The occurence of OKN, where red indicates a large velocity between frames. A rapid movement to the right occurs over these frames. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

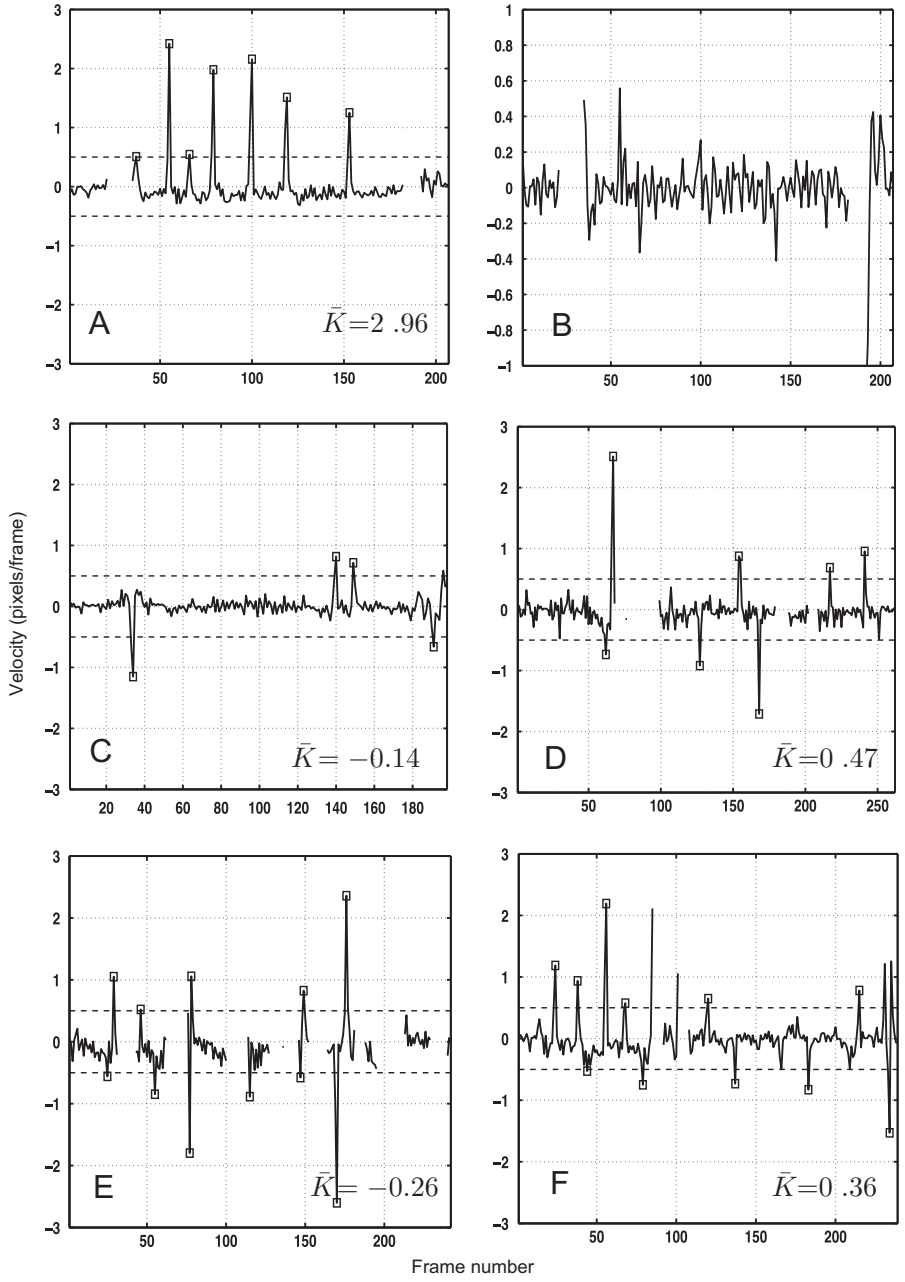

Fig. 6. A series of example traces. (A) The horizontal component of the limbal velocity for a successfully detected trace. Notice the large positive value for $\bar{K}$ (the signed value of the OKN consistency measure). (B) The corresponding vertical trace does not contain OKN reset information, in this case no $\bar{K}$ value is shown. (C–E) incorrectly detected examples, where peaks are found in both directions, leading to small absolute OKN consistency values ($|\bar{K}| < 0.5$). Empty data is due to eye blinks, but blur also reduces edge detection efficacy (see in particular (E)). (F) A correctly determined trace, but with only a marginal value for $\bar{K}$. The dotted horizontal lines below and above the zero velocity axis (panels (A) and (C–F)) show the velocity threshold $\sigma_{\overline{v}} = 0.5$ pixels/frame.

3.2.3. Procedure

A group of five adult observers (mean age = 23) with normal vision who did not participate in experiment 1 viewed the stimuli patterns binocularly, and were instructed to stare at the center of the drifting stimulus pattern during the stimulus presentation period. Observers heads were restrained using a chin rest. Each observer completed one trial of each stimulus velocity (0, 5 and 10 deg/s). Eye movements were recorded for a 20 s period and the recordings were processed using the optic flow methods described above.

The limbus detection parameters were modified to account for the larger image of the eye that resulted from positioning the camera closer to the eye. In experiment 1, a single pyramidal reduction was used, and groupings of pixels with a weight below 20 were discarded. In this experiment, an additional pyramidal reduction was performed, and groupings of pixels less than 50 pixels were discarded. It is also noted that two additional heuristic rules were introduced. Firstly, blinks were detected by thresholding the vertical velocity signal of the eye ($\overline{V}_y(t) > 1$ pixels/frame) as determined by masking the vertical component of the optical flow with a horizontal edge detector (again using the Prewitt operator). By doing so, we aimed to detect frames containing the rapid downward sweep of the (horizontal) lid edge. The regions of the $\overline{V}_x(t)$ signal corresponding to the detected eye blinks were ignored from further processing (in this work, a two frame interval on either side of a detected blink was also ignored). Secondly, a strategy for removing widely spaced peaks was introduced. If two subsequent

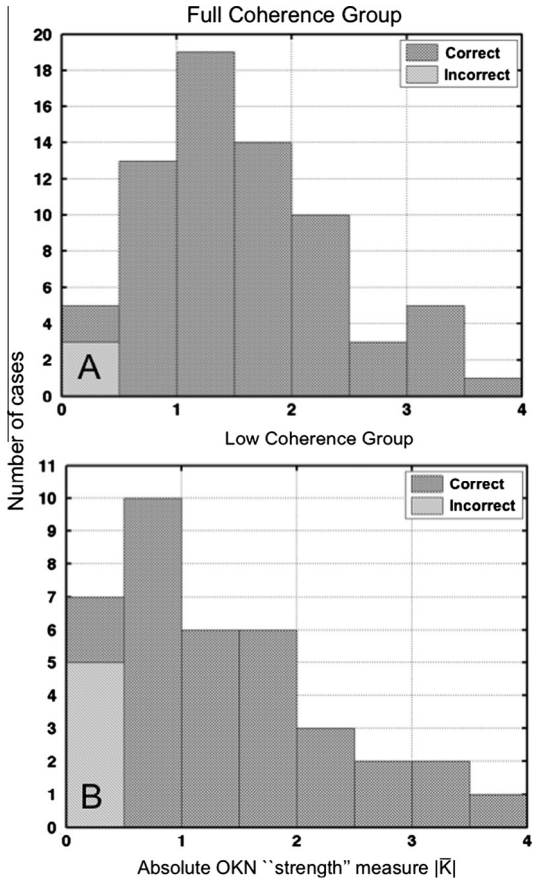

Fig. 7. The distribution of absolute OKN consistencies $|\overline{K}|$ for: (A) The full coherence group. (B) The low coherence group. Shown within each histogram are the correct detections, as well as the incorrect detections. The incorrect detections appear at low values of OKN consistency, particularly so for the full coherence group.

peaks were detected (in the same direction), but the time interval between subsequent saccadic peaks was greater than a 4 s threshold (approximately the time taken to travel across the full screen at the 5 deg/s velocity), then it was assumed that the interval between the peaks was not consistent with being due to the slow phase of OKN. In the absence of additional neighboring peaks (of the same sign) these two peaks would be discarded.

3.2.4. Results

For the stationary stimuli, velocity traces either did not pass through the algorithm to allow estimation of $\overline{K}$ (3/5 traces) or $\overline{K}$ did not exceed threshold (2/5 traces, $|\overline{K}| \leqslant 0.6$ pixels/frame). The direction of all moving stimuli was classified correctly by the detector and the $|\overline{K}|$ values for the 5 deg/s trials did not differ significantly from those for the 10 deg/s trials (5 deg/s mean $|\overline{K}| = 2.38$, SD = 0.15; 10 deg/s mean $|\overline{K}| = 2.85$, SD = 0.62; $t = 1.45$, $p = 0.22$).

3.2.5. Discussion

The OKN detector was sensitive to the presence or absence of OKN when provided with web-cam footage of participants viewing square wave grating stimuli. In the case where valid OKN was detected, the $|\overline{K}|$ values produced by the OKN detector did not change when the stimulus velocity was doubled suggesting that this measure for detecting OKN is relatively robust to stimulus velocity.

4. General discussion

We report a technique for detecting the presence and direction of OKN that can be implemented using off-the shelf video equipment (*i.e.*, camcorders or web-cams). Our initial testing indicates that the technique can detect OKN elicited by RDKs and square wave gratings and has an accuracy that is comparable to that of an experienced human observer.

The technique may be useful for the assessment of visual perception in observers who are not able to complete clinical or psychophysical tasks that require subjective behavioral responses. Previous studies have used the presence or absence of OKN in response to specific psychophysical stimuli to address a range of issues such as visual acuity in patients with low vision (Wester, Rizzo, Balkwill, & Wall, 2007) and global motion perception in infants (Banton & Bertenthal, 1996; Mason, Braddick, & Wattam-Bell, 2003; Yu et al., 2013). However, the use of OKN to measure visual function can be challenging. Well calibrated eye tracking equipment can be used to precisely measure the motion of the eye and provide displacement over time information that can subsequently be used to derive the presence or absence of OKN. However the required equipment is expensive and often requires levels of compliance from the observer that can be difficult to achieve. Alternatively human observers can judge the presence of OKN, either in real time or on the basis of video footage of the eyes. While reliable (see experiment 1), this approach can be laborious, particularly if a considerable number of trials are required to estimate a psychophysical threshold. Our technique provides a third alternative whereby the presence or absence of OKN can be determined objectively on the basis of eye movement recordings made using inexpensive equipment that does not require detailed calibration. We are currently working to integrate our OKN detection approach with a technique that we have developed which stabilizes the position of the head within video footage of participants with unrestrained heads (Sangi, Thompson, Vaghefi, & Turuwhenua, 2014). This will allow for OKN detection without the need for head restrains such as chin rests which are poorly tolerated by young children.

We found that the simple edge map approach we used worked effectively. Fortuitously, in experiment 1, we found that if an eye blink occurred (see for example, Fig. 6A) the limbal edge vanished, thereby leaving a null response in the velocity record $\overline{V}_x(t)$. Therefore the absence of the limbal edge served as a simple blink detection mechanism. In experiment 2, an additional simple blink detection mechanism was introduced to compensate for the close proximity of the camera to the eye.

On the other hand, we found that non-limbal features such as eye-lashes and general reflections could, in some instances, distort the limbal edge mask. Furthermore, these spurious distortions could appear and disappear temporally in a highly unpredictable manner. Even so, we found that our estimates for $\overline{V}_x(t)$ remained reliable in these instances: the shape of the edge map *did not* need to be particularly precise (spatially or temporally) in order to yield valid results. It is likely that the process of averaging over the edge map made the velocity estimates robust to instantaneous changes in the size of the mask and the presence of stationary non-limbal edges such as eyelashes. Nevertheless, it is possible that our method could be improved by the addition of suitable eyelash (Kong & Zhang, 2001), and corneal reflection (Li, Winfield, & Parkhurst, 2005) detectors.

The main disadvantage of our technique is that it acts as an OKN detector and does not provide estimates of OKN gain, a measure that has been widely used to assess the integrity of the visual pathways involved in retinal image stabilization (Rivaud, Müri, Gaymard, Vermersch, & Pierrot-Deseilligny, 1994; Valmaggia

et al., 2004). Accurate measurements of OKN gain require a temporal resolution greater than that provided by most off the shelf video equipment and therefore require specialized eye tracking equipment. Rather than measuring OKN gain, our technique is designed to indicate the presence of OKN and its direction in cases where eye tracking equipment is not available or cannot be used due to issues with observer compliance. Work is underway to understand whether the methods developed here could be used with higher frame rate equipment.

5. Conclusion

We have developed a promising approach for determining the presence and direction of OKN in video sequences recorded using low cost off-the-shelf equipment. The method uses optic flow combined with a limbal masking strategy to directly estimate limbal velocity. Simple heuristics are used to determine the presence and direction of OKN within a trial. The approach aims to avoid pitfalls of methods based on model/feature tracking, in-particular, the confounding effects of eyelashes on feature detection. The results are encouraging, and suggest that more work is warranted. This method could be used as the basis of a system for the low cost assessment of OKN.

Acknowledgments

This work was supported by a grant from the University of Auckland Faculty Development Research Fund to BT and JT. The authors would like to thank Dr. Misha Vorobyev for helpful comments, and Mr. Nouzar Irani and Mr. Girish Ramlugun for their assistance with data collection.

References

Abe, K., Ohi, S., & Ohyama, M. (2007). An eye-gaze input system using information on eye movement history. In *Proceedings of the 4th international conference on Universal access in human-computer interaction: Ambient interaction. UAHCI'07* (pp. 721–729). Berlin, Heidelberg: Springer-Verlag. <http://dl.acm.org/citation.cfm?id=1763296.1763379>.

Banton, T., & Bertenthal, B. I. (1996). Infants' sensitivity to uniform motion. *Vision Research, 36*(11), 1633–1640.

Barron, J. L., Fleet, D. J., & Beauchemin, S. S. (1994). Performance of optical flow techniques. *International Journal of Computer Vision, 12*(1), 43–77.

Björck, Å. (1996). Numerical methods for least squares problems. No. 51. Society for Industrial Mathematics.

Braddick, O., Atkinson, J., & Wattam-Bell, J. (2003). Normal and anomalous development of visual motion processing: Motion coherence and 'dorsal-stream vulnerability'. *Neuropsychologia, 41*(13), 1769–1784. braddick, Oliver Atkinson, Janette Wattam-Bell, John Review England.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Büttner, U., & Kremmyda, O., (2007). Neuro-Ophthalmology Neuronal Control of Eye Movements (Vol. 40, pp. 76–89). Karger.

Camus, T., & Wildes, R. (2002). Reliable and fast eye finding in close-up images. *16th international conference on pattern recognition, 2002. Proceedings* (Vol. 1, pp. 389–394). IEEE.

Chow, G., & Li, X. (1993). Towards a system for automatic facial feature detection. *Pattern Recognition, 26*(12), 1739–1755. <http://www.sciencedirect.com/science/article/pii/003132039390173T>.

Condurache, A., & Aach, T., (2005). Vessel segmentation in angiograms using hysteresis thresholding. In MVA (pp. 269–272).

Daugman, J. (1993). High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 15*(11), 1148–1161.

Dollár, P., (2012). Piotr's Image and Video Matlab Toolbox (PMT). <http://vision.ucsd.edu/pdollar/toolbox/doc/index.html>. <http://vision.ucsd.edu/pdollar/toolbox/doc/index.html>.

Duchowski, A. (2007). *Eye tracking methodology: Theory and practice*. Springer-Verlag New York Inc.

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM, 24*(6), 381–395.

Franchak, J. M., Kretch, K. S., Soska, K. C., & Adolph, K. E. (2011). Head-mounted eye tracking: a new method to describe infant looking. *Child Development, 82*(6), 1738–1750.

Grinter, E. J., Maybery, M. T., & Badcock, D. R. (2010). Vision in developmental disorders: Is there a dorsal stream deficit? *Brain Research Bulletin, 82*(3-4), 147–160. http://dx.doi.org/10.1016/j.brainresbull.2010.02.016. 0361-9230.

Han, S. B., Han, E. R., Hyon, J. Y., Seo, J.-M., Lee, J. H., & Hwang, J.-M. (2011). Measurement of distance objective visual acuity with the computerized optokinetic nystagmus test in patients with ocular diseases. *Graefe's Archive for Clinical and Experimental Ophthalmology, 249*(9), 1379–1385.

He, Z., Tan, T., Sun, Z., & Qiu, X. (2008). Robust eyelid, eyelash and shadow localization for iris recognition. In *15th IEEE international conference on image processing, 2008. ICIP 2008* (pp. 265–268). IEEE.

Horn, B., & Schunck, B. (1981). Determining optical flow. *Artificial Intelligence, 17*(1-3), 185–203.

Hsu, R., Abdel-Mottaleb, M., & Jain, A. (2002). Face detection in color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24*(5), 696–706.

Hyon, J. Y., Yeo, H. E., Seo, J.-M., Lee, I. B., Lee, J. H., & Hwang, J.-M. (2010). Objective measurement of distance visual acuity determined by computerized optokinetic nystagmus test. *Investigative Ophthalmology & Visual Science, 51*(2), 752–757.

Iskander, D. R. (2006). A parametric approach to measuring limbus corneae from digital images. *IEEE Transactions on Biomedical Engineering, 53*(6), 1134–1140.

Kong, W., & Zhang, D. (2001a). Accurate iris segmentation based on novel reflection and eyelash detection model. In *Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing* (pp. 263–266). 2001: IEEE.

Kong, W., & Zhang, D., (2001b). Eyelash detection model for accurate iris segmentation. In *Proceeding of ISCA 16th international conference on computers and their applications* (pp. 204–207). Citeseer.

Lee, I., Choi, B., & Park, K. (2007). Robust measurement of ocular torsion using iterative Lucas–Kanade. *Computer Methods and Programs in Biomedicine, 85*(3), 238–246.

Li, D., Babcock, J., & Parkhurst, D. J. (2006). openEyes: a low-cost head-mounted eye-tracking solution. In *Proceedings of the 2006 symposium on Eye tracking research & applications. ETRA '06* (pp. 95–100). New York, NY, USA: ACM. <http://doi.acm.org/10.1145/1117309.1117350>.

Li, D., Winfield, D., & Parkhurst, D. J. (2005). Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches. *Computer Vision and Pattern Recognition Workshop, 0*, 79.

Macintyre-Béon, C., Ibrahim, H., Hay, I., Cockburn, D., Calvert, J., Dutton, G. N., et al. (2010). Dorsal stream dysfunction in children. A review and an approach to diagnosis and management. *Current Pediatric Reviews, 6*(3), 166–182.

Marg, E. (1951). Development of electro-oculography. *Archives of Ophthalmology, 45*, 169–185 (no issue number).

Mason, A. J. S., Braddick, O. J., & Wattam-Bell, J. (2003). Motion coherence thresholds in infants–different tasks identify at least two distinct motion systems. *Vision Research, 43*(10), 1149–1157.

Morris, T., Blenkhorn, P., & Zaidi, F. (2002). Blink detection for real-time eye tracking. *Journal of Network and Computer Applications, 25*(2), 129–143. <http://www.sciencedirect.com/science/article/pii/S108480450290130X>.

Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature, 341*(6237), 52–54.

Prewitt, J. (1970). *Object enhancement and extraction* (Vol. 75). New York: Academic Press.

Rivaud, S., Müri, R., Gaymard, B., Vermersch, A., & Pierrot-Deseilligny, C. (1994). Eye movement disorders after frontal eye field lesions in humans. *Experimental Brain Research, 102*(1), 110–120. <http://dx.doi.org/10.1007/BF00232443>.

Sangi, M., Thompson, B., Vaghefi, E., & Turuwhenua, J. (2014). A head tracking method for improved eye movement detection in children. In J. Goh (Ed.), *The 15th international conference on biomedical engineering. IFMBE proceedings* (Vol. 43, pp. 508–511). Springer International Publishing.

Valmaggia, C., Rütsche, A., Baumann, A., Pieh, C., Shavit, Y. B., Proudlock, F., et al. (2004). Age related change of optokinetic nystagmus in healthy subjects: A study from infancy to senescence. *British Journal of Ophthalmology, 88*(12), 1577–1581.

Waddington, J., & Harris, C. M. (2012). Human optokinetic nystagmus: A stochastic analysis. *Journal of Vision, 12*(12).

Wester, S. T., Rizzo, J. F., 3rd, Balkwill, M. D., & Wall, C. 3rd, (2007). Optokinetic nystagmus as a measure of visual function in severely visually impaired patients. *Investigative Ophthalmology & Visual Science, 48*(10), 4542–4548.

Young, L. R., & Sheena, D. (1975). Survey of eye movement recording methods. *Behavior Research Methods, 7*(5), 397–429.

Yu, T.-Y., Jacobs, R. J., Anstice, N. S., Paudel, N., Harding, J. E., & Thompson, B.CHYLD Study Team. (2013). Global motion perception in 2-year-old children: a method for psychophysical assessment and relationships with clinical measures of visual function. *Investigative Ophthalmology & Visual Science, 54*(13), 8408–8419.

Yuille, A., Hallinan, P., & Cohen, D. (1992). Feature extraction from faces using deformable templates. *International Journal of Computer Vision, 8*(2), 99–111.