Computers in Biology and Medicine 114 (2019) 103448

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Automatic nystagmus detection and quantification in long-term continuous eye-movement data☆

Jacob L. Newman [a,*], John S. Phillips [b], Stephen J. Cox [a], John FitzGerald [b], Andrew Bath [b]

[a] School of Computing Sciences, University of East Anglia, Norwich, NR4 7TJ, United Kingdom
[b] Norfolk & Norwich University Hospitals NHS Foundation Trust, Norwich, NR4 7UY, United Kingdom

ARTICLE INFO

Keywords:
Nystagmus
Dizziness
Vestibular diseases
Event detection
Biomedical signal processing
Electronystagmography
Time series classification

ABSTRACT

Symptoms of dizziness or imbalance are frequently reported by people over 65. Dizziness is usually episodic and can have many causes, making diagnosis problematic. When it is due to inner-ear malfunctions, it is usually accompanied by abnormal eye-movements called nystagmus. The CAVA (Continuous Ambulatory Vestibular Assessment) device has been developed to provide continuous monitoring of eye-movements to gain insight into the physiological parameters present during a dizziness attack. In this paper, we describe novel algorithms for detecting short periods of artificially induced nystagmus from the long-term eye movement data collected by the CAVA device. In a blinded trial involving 17 healthy subjects, each participant induced nystagmus artificially on up to eight occasions by watching a short video on a VR headset. Our algorithms detected these short periods with an accuracy of 98.77%. Additionally, data relating to vestibular induced nystagmus was collected, analysed and then compared to a conventional technique for assessing nystagmus during caloric testing. The results show that a range of nystagmus can be identified and quantified using computational methods applied to long-term eye-movement data captured by the CAVA device.

1. Introduction

Dizziness is a physically and socially debilitating symptom. In England and Wales, dizziness or imbalance are experienced by 30% of the population by the age of 65 years [1]. Dizziness often occurs episodically, reducing the likelihood that an attack with coincide with a visit to a clinician. Furthermore, there are multiple possible causes of dizziness, and they originate from pathologies that affect a number of different organ systems. Patient reporting of dizziness symptoms has shown to be imprecise and unreliable [2]. These aspects of dizziness make its diagnosis challenging, and patients may visit several healthcare professionals, in both primary and secondary care, before receiving a diagnosis [3].

When dizziness is due to a malfunction of the pathway involving the eyes, brain and inner ears, it is usually accompanied by abnormal eye-movements called *nystagmus* [4]. Fig. 1 shows an example of an eye-movement trace corresponding to horizontal right-beating nystagmus. The waveform pictured is commonly referred to as a *sawtooth* wave. This waveform is comprised of a slow component, corresponding to the eyes moving slowly in a leftward direction, and a fast phase in which the eyes move quickly in the opposite direction. The direction of the fast-phase clinically defines the *beat* direction, which can either be left, right, up or down-beating.

Nystagmus can be classified as either pathological or physiological. Pathological nystagmus can arise as a clinical sign of a multitude of peripheral vestibular conditions, such as vestibular neuronitis, Ménière's Disease and Benign Paroxysmal Positional Vertigo (BPPV). The dizzy attacks associated with these conditions are known to vary in duration and severity [5]. Physiological nystagmus can also be induced as a normal response to viewing a full-field, moving, visual stimulus (e.g. as experienced by looking out of the window of a moving train) [6], or by stimulating the balance organs directly through caloric testing [7].

During caloric testing, warm or cold water is used to irrigate the external ear canal and this invokes movement of the fluid within the semicircular canals of the inner ear. This stimulation leads to patterns of nystagmus which can be observed and analysed by an expert. Caloric testing is one of an array of tools used by clinicians to assess the function of an individual's balance organs and to make a diagnosis [8]. The tests described are limited by their inability to provide an assessment of a patient's symptoms in the community during an episode of dizziness.

The CAVA (Continuous Ambulatory Vestibular Assessment) device is part of a new system we have developed to overcome the limitations

☆ This work is funded by the Medical Research Council (MR/P026265/1).
* Corresponding author.
E-mail address: jacob.newman@uea.ac.uk (J.L. Newman).
URL: http://www.cava-project.org (J.L. Newman).

https://doi.org/10.1016/j.compbiomed.2019.103448
Received 17 July 2019; Received in revised form 13 September 2019; Accepted 13 September 2019
Available online 17 September 2019

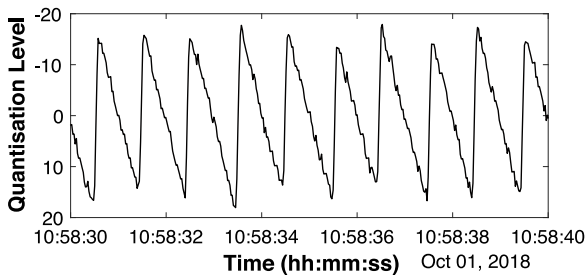

Fig. 1. An example of right-beating nystagmus presented in the horizontal eye-movement channel, as captured by the CAVA device.

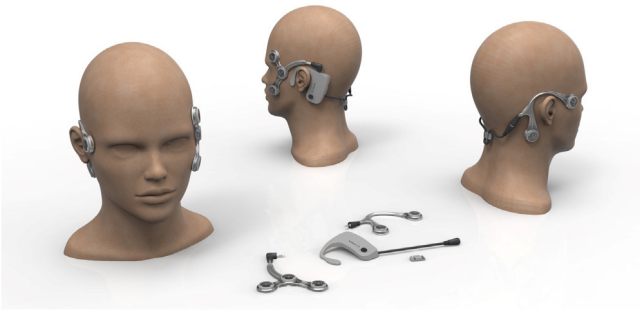

Fig. 2. The CAVA device. The device consists of five electrode pads contained within two, detachable mounts, and an electronic logging unit which sits behind the left ear.

of conventional diagnostic techniques. The device records a patient's eye-movements continuously over 30 days, after which the data can be automatically analysed offline for features of nystagmus, a concept that reflects existing technology for long-term ambulatory cardiac and brain monitoring. A duration of 30 days was considered to be long enough to capture sufficient examples of nystagmus in patients with active dizziness to support a clinician's diagnosis. The rich information provided by the device, which is collected during the user's normal daytime and nighttime activities and has not been obtained before, may allow new insights into the understanding of inner-ear disorders. For example, it may eventually be used to predict episode onset, to identify aetiological factors, to determine prognosis, to define and categorise subtypes of dizziness syndromes, to identify the pathophysiological mechanisms underpinning disease development, and to monitor treatment outcomes.

There is much previous work on the automated detection of nystagmus events from eye movement data [9–13]. Existing techniques usually involve the detection of saccades, by way of information contained in the velocity signal. For example, the fast phase of nystagmus can be identified from peaks in the velocity channel surrounded by velocities of the opposite sign. However, no previous work has been applied to the novel, long-term data provided by the CAVA device, which poses new challenges that previous techniques may not be well-suited to tackle. For example, the large amount of data captured during a 30-day monitoring period imposes limits of the time complexity of the algorithms applied, and the vast variability of normal eye movement data makes conventional techniques susceptible to increased rates of false positive detections. Other areas of research, such as event detection in long-term electrocardiography (ECG) and electroencephalography (EEG) data, have been the focus of extensive research and are perhaps more applicable to this work. Various computational techniques have been applied in these fields, including neural networks [14–17], other machine learning algorithms such as support vector machines and decision trees [18–21] and many time-series classification approaches [22–25].

This article describes in detail the novel computing approaches developed to tackle nystagmus detection in long-term eye movement data, and also new algorithms for detecting and quantifying induced nystagmus captured during a routine clinical test. For this purpose, we have undertaken a clinical investigation involving seventeen healthy volunteers. For a full description of the clinical aspects of this investigation, including a detailed analysis of the reliability, acceptability and safety of the device, please see [26]. The trial was designed to demonstrate the suitability of the device and to evaluate its safety, prior to a larger investigation involving individuals experiencing dizziness. As the healthy volunteers recruited for this study did not suffer from dizziness, nystagmus was artificially induced by way of a visual stimulus, providing data which could be used to determine the specificity and sensitivity of the algorithms developed for detecting nystagmus. This stimulus consisted of a video, viewed on a VR headset, showing a black dot moving repeatedly across a white a screen. An offline, blinded analysis of the data was performed after the trial, the purpose of which was to automatically identify the dates on which nystagmus was induced. Additionally, ten of the participants underwent caloric testing whilst simultaneously wearing the CAVA device and a pair of camera-mounted goggles (Videonystagmography or VNG), which is the current state-of-the-art for assessing nystagmus in clinical settings. Using this *vestibular-induced* nystagmus data, a separate computer analysis was undertaken to compare the two systems.

The remainder of this article is organised as follows: in Section 2, we describe the CAVA device and the important aspects of its functionality and design. Section 3 provides details of the clinical investigation undertaken to evaluate the device and to collect the dataset used in the experiments described here. Section 4 outlines a blinded recognition experiment used to assess the diagnostic accuracy of the system developed, and also the computer analysis applied to the vestibular induced nystagmus data. The algorithms developed for this work are described in Section 5 and the experimental results are presented in Section 6. Section 7 includes a discussion of the results and the article concludes in Section 8.

2. Device design and functionality

The CAVA device (Fig. 2) captures the corneo-retinal potential (CRP) between the back and front of each eye, which is a proxy for eye movement [27]. Horizontal eye-movement is captured by placing sensors at the outer edge of both eyes (lateral canthi), and vertical eye-movement by placing sensors above and below one eye. Hence the device uses five electrodes; two for horizontal movement, two for vertical movement and a fifth to provide a reference voltage. Similar technology has been used routinely in clinical settings for decades, in the form of Electronystagmography (or ENG) [28]. More recently, camera-mounted video goggles (Videonystagmography or VNG) are the preferred approach, as computer advancements have made realtime eye tracking possible. However, CAVA is intended to monitor eye-movements over long periods and this imposes constraints on size, privacy and power consumption; hence ENG was the preferred approach for this work. In addition, CAVA has been designed to be comfortable to wear for extended durations and during normal daily activities; it is small, lightweight and minimally intrusive to the user. Battery life is a minimum of eighteen days and the user is not required to activate or charge the device.

The device contains a three-axis accelerometer, which simultaneously records the accelerative forces experienced by the head. An event marker button is present on the device to allow the wearer to log events of interest (such as experiencing dizziness), for which the device records the date and time of the button press. The device also includes a status LED, which is used to indicate both normal device operation and device malfunction. Data captured by the device is stored on a removable SD card and can be downloaded by way of a USB-B interface.

Neither the SD card nor the USB-B interface are accessible to the subject, minimising the chance of damage to the device or loss of data.

The electronics underlying the CAVA device includes a microcontroller unit (MCU) and an analogue to digital front end (FE). The MCU is a low power, eight-bit chip chosen to offer an integrated USB solution for the device. The MCU also supports a timing function and battery monitoring. The FE used in CAVA is based on an integrated 2-channel device provided by Texas Instruments (ADS1192). The horizontal and vertical eye movement channels are sampled at a frequency of 41.67 Hz each and the three acceleration components at 20 Hz each. These frequencies are a compromise between providing sufficient bandwidth to record the associated movements with low distortion whilst minimising power consumption.

3. Clinical investigation

A clinical investigation was conducted on healthy volunteers to evaluate the CAVA device's suitability for its intended purpose of long-term data capture of eye movement data. The trial produced a large dataset of eye movement data intended to be analysed by computer algorithms to demonstrate the accuracy of the system at detecting 30-s periods of artificially induced nystagmus. The trial also aimed to provide data to establish the reliability, acceptability and safety of the device, the results of which are described fully in [26]. During the trial, seventeen participants wore the CAVA device for 23 h a day, for a period of up to 30 days. A period of 23 h was chosen to give participants an hour in which to shower and to renew the device's electrode mounts. It was necessary to renew the mounts each day to ensure that the electrodes were adequately adhesive and conductive for capturing good quality signals. Participants were not otherwise required to maintain the device, as the battery life was sufficient to last between hospital visits. On eight pre-determined dates, each participant was asked to watch a video on a VR headset. This 30-s video showed a single black dot moving repeatedly across a white screen. The purpose of this video was to visually induce nystagmus. This stimulus differs from conventional optokinetic videos, in which many bars or dots move across the field of view. The instructions for the task were less ambiguous and the participants' nystagmus responses were less varied than might be expected from a conventional stimulus [29]. The number of participants and the number of occasions on which they simulated nystagmus was chosen to demonstrate 98% sensitivity and specificity, with a 95% confidence interval and a 3% margin of error.

Each participant was provided with one of six possible videos. The videos either contained a leftward or a rightward moving dot, and the dot moved at one of three different speeds (0.8 Hz, 1.0 Hz or 1.2 Hz, where 1.0 Hz means that each dot took 1 s to move across the screen). These videos induced nystagmus with a fast phase (beat direction) in the opposite direction of the moving dot, corresponding to the eyes moving to the edge of the screen to catch the emergence of each dot.

During the trial, participants were instructed to press the device's event marker prior to viewing the video and also to keep a written record of the time of viewing. These records allowed the unblinded investigator to validate the data captured after the trial. Validation was achieved by visually inspecting the signals captured for evidence of a nystagmus waveform, which confirmed that participants had successfully viewed the video stimulus. For the first four viewings of the video, participants were instructed to watch whilst sitting or standing still, and for the last four they were told to walk gently on the spot. The purpose of this task was to determine if walking introduced noise that affected the nystagmus signal.

After day 31 of the trial, participants returned to the hospital for their final visit, where they underwent a routine balance assessment procedure, known as a caloric test. Ten participants underwent this procedure whilst simultaneously wearing the CAVA device and a conventional camera-mounted goggle system (VNG). Each participant was first given a warm water irrigation (44 °C) in their right ear, and then in their left ear. In healthy volunteers, such as those in this study, these irrigations are expected to lead to nystagmus beating in the direction of the irrigated ear. Thus, each participant produced right and then left-beating nystagmus. The purpose of this experiment was to allow a retrospective comparison between an established approach to quantifying nystagmus and the results obtained by the CAVA system.

The VNG system was calibrated prior to the start of the irrigations. During the calibration, participants sat approximately 36 inches from a screen on which a dot performed a 30 degree excursion (±15 degrees from the centre). The event marker on the CAVA device was used to record the start time of the calibration, and the event marker was also deployed at the start of each irrigation. Vestibular nystagmus can be consciously suppressed, and therefore, participants undertook a simple word game exercise in order to divert their attention during the test. Once the irrigation was complete, a clinical scientist monitored the participant's eye movements on a computer screen to determine the start and end of the participant's nystagmus response. Finally, the clinical scientist reviewed the VNG recording and, using manual analysis, estimated the beat direction, the maximum slow phase velocity (SPV) and the time it occurred.

4. Recognition tasks

At the end of the clinical trial, two separate computer analysis tasks were undertaken using the data obtained: first, a blinded, nystagmus detection task using the data from each participant's 30-day trial, and second, an unblinded caloric nystagmus analysis task using the data captured during caloric testing. Given 405 separate data files, each containing a day's worth of horizontal eye movement data, the primary recognition task was to identify automatically the 112 files that contained the 30 s of nystagmus waveform embedded in the 24 h of data. In addition, for each file, the beat direction (left or right) and also the speed of the moving dots (0.8 Hz, 1.0 Hz or 1.2 Hz) were to be determined. The data files were re-segmented prior to this experiment, which involved combining multiple files from a single day, joining the first and last half-days together, and adjusting the data to account for daylight savings time. Prior to testing, one investigator was blinded from the true class labels of the data files. This blinding process was achieved by renaming each file to a randomly generated character string. The record of this randomisation was only accessible to the unblinded investigators. As described in Section 3, these tasks were performed as part of a formal clinical investigation of the CAVA medical device.

The secondary task was to separately analyse and compare the caloric nystagmus data to the results obtained by a clinical scientist. Ten participants underwent caloric testing whilst simultaneously wearing VNG goggles and the CAVA device. Each participant was given a hot water irrigation in both ears. Of the twenty irrigations performed, two were undertaken without the VNG goggles, in case the goggles were found to interfere in some way with the CAVA device. Analysis of the data did not reveal any such effects. After each test, the clinical scientist reviewed the VNG recording and manually recorded a number of features of the nystagmus, including the beat direction, the maximum slow phase velocity (SPV), and the start and end time of maximum SPV. The same measurements were then extracted using computational methods applied to data from the CAVA device. The focus of this task was to provide a more detailed analysis of the nystagmus and to compare the CAVA system to a well-established technique used in clinical settings.

5. Nystagmus recognition algorithms

Two separate algorithms were developed for the two tasks described in Section 4. The first algorithm was developed for the blinded nystagmus detection task and comprised modules for testing and training. Feature extraction was applied to the horizontal eye-movement data for both the testing and the training data (Section 5.1). System training

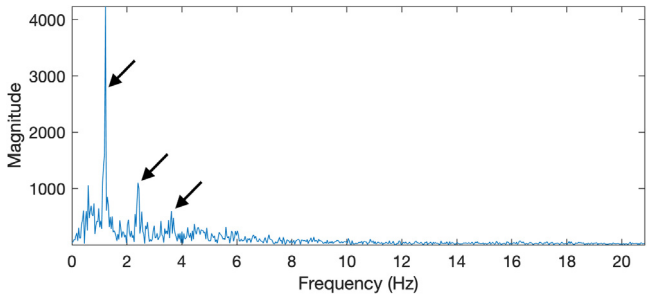

Fig. 3. The frequency spectrum of a typical example of visually-induced nystagmus captured by the CAVA device. The black arrows mark the fundamental frequency of the nystagmus waveform (at 1.2 Hz), an even and an odd harmonic (at 2.4 Hz and 3.6 Hz respectively). Harmonics beyond 3.6 Hz are largely indistinguishable from the background noise.

(Section 5.2) was performed using a small dataset of internally collected data, in which participants viewed several of the video stimuli over a short period of time (usually, less than an hour). Experiments using preliminary data were used to select appropriate model parameters and thresholds. Finally, classification was performed on the randomised test data gathered from the clinical investigation (Section 5.3). These algorithms were developed using a high performance computing cluster to allow many hundreds of experiments to be performed simultaneously. The final algorithm applied to the blinded detection task took approximately 20 s to process a day's worth of data, and approximately 2.25 h to process sequentially all 405 days of data on a 2017 MacBook Pro laptop.

A second algorithm was developed to analyse the caloric data, for detection and quantification of the individual nystagmus beats. This algorithm is described in Section 5.4 and the parameters used were determined empirically, favouring sensitivity over specificity.

5.1. Feature extraction

A Fast Fourier Transform (FFT) was used to transform the horizontal eye-movement time series data into the frequency domain. Analysis of the target nystagmus signals revealed distinctive frequency domain characteristics, owing to its similarity to a sawtooth wave (Fig. 3). Therefore, FFT features were the natural choice for this work. Vertical eye movement data was discarded as this information was not considered necessary for detecting horizontal nystagmus. A sliding window with 80% overlap and a window length of 500 samples was used. As the sampling frequency of the device is 41.67 Hz, these parameters correspond to a window length of 12 s. The window length and overlap lengths used ensure that when a nystagmus event of length 30 s is encountered in the signal, there are sufficient cycles of nystagmus present in several windows to provide a good spectral signature whilst also giving good time resolution. Prior to applying the FFT, the time series signal in each window was zero-meaned to remove any DC drift, and a Hanning window was applied to reduce spectral leakage.

After applying the FFT, each frame was represented by a 250-dimensional vector of positive frequency magnitudes. Coefficients 3 to 84 were retained, corresponding to the frequencies between 0.25 Hz and 7 Hz, spaced at 0.083 Hz intervals. Coefficients corresponding to frequencies above 7 Hz were discarded because the harmonic components caused by nystagmus above this frequency were indistinguishable from noise, whilst coefficients representing frequencies lower than 0.25 Hz represent noise that is related to the DC offset and any remaining signal drift (Fig. 3).

The distribution of the frequency amplitudes in each frame was used as a measure of frame energy. Frames were discarded if less than 50% of the distribution occurred outside of the 0.25 Hz–7 Hz range, and where the sum of the amplitudes in a frame was above or below a threshold value. For testing, discarding meant assigning a non-nystagmus class label to the data, while for training, the data was excluded from the model building process. For all data, each frame was normalised by its maximum value. Finally, the signal velocity was extracted from the time series signal, which was used during testing to determine the beat direction of the nystagmus.

5.2. Training

Three separate classifiers were trained using the features derived from the training data: a Support Vector Machine (SVM), a Linear Discriminant Analysis transformation (LDA) and an Ensemble classifier of boosted trees [30]. The SVM's kernel value was determined automatically by MATLAB, which uses a heuristic subsampling method to find an appropriate value. The use of multiple classifiers allows an ensemble approach to recognition, in which the classifiers vote to determine the final classification. This approach has been shown to often outperform conventional single classifier systems [31]. These classifiers were trained to identify nystagmus, irrespective of direction and speed. As described in 5.3, following this coarse scale identification of putative nystagmus candidates, and subsequent validation of these events, a more detailed classification of beat direction and speed was made, which did not require training.

5.3. Classification

Unseen test data was treated using the same feature extraction process as applied to the training data (Section 5.1), producing several thousand frames of data for each day. Following this, classification consists of three stages: classification, sieving and validation. Firstly, the testing data was classified on a frame-by-frame basis by each of the three classification models described in Section 5.2. All three classifications were combined by way of a majority voting system, whereby the winning class is the one for which two or more classifiers agree. After this first classification stage, each frame was represented by a binary classification, indicating whether that frame contained nystagmus or not.

A sequence of classified frames typically has some frames labelled "nystagmus" and some "non-nystagmus", and so the question is raised of how to decide the class of such a signal segment. We used a *sieve* filter to smooth the output from the first classification stage. For a full description of the operation of a sieve filter, please see [32], but to summarise, the sieve essentially operates here by removing very short durations of negative or positive classifications. After sieving, the continuous runs of positively classified frames that remain represent candidate nystagmus events. The beat direction of each event is classified using the velocity of the time series waveform (Section 5.1). The velocity of the signal provides useful information which can be used to determine the beat-direction of the nystagmus. For example, the slow phase of a normal sawtooth wave contains mostly negative velocities, while for a mirrored sawtooth, the signs are inverted. Each frame was assigned a weighting (between −1 and 1) reflecting the proportion of negative and positive velocities. The average weighting within the segment determined the direction of the nystagmus.

Once each candidate event is assigned a beat direction, it is then subject to a validation procedure (Fig. 4). In this step, Dynamic Programming (DP) [33] was used to compare the corresponding time series signal to a 1 Hz sawtooth waveform of an equal duration. The technique of DP compresses or expands both the test signal and the sawtooth waveform in such a way as to produce the optimum fit between the two signals. The amount of compression or expansion distortion required for the fit is a measure of the "distance" between the two waveforms, i.e. how closely the test signal corresponds to a sawtooth waveform. Either a sawtooth or a mirrored sawtooth waveform were compared, depending on whether the test signal is classified as right or left-beating nystagmus. The DP distances were normalised by the length of each

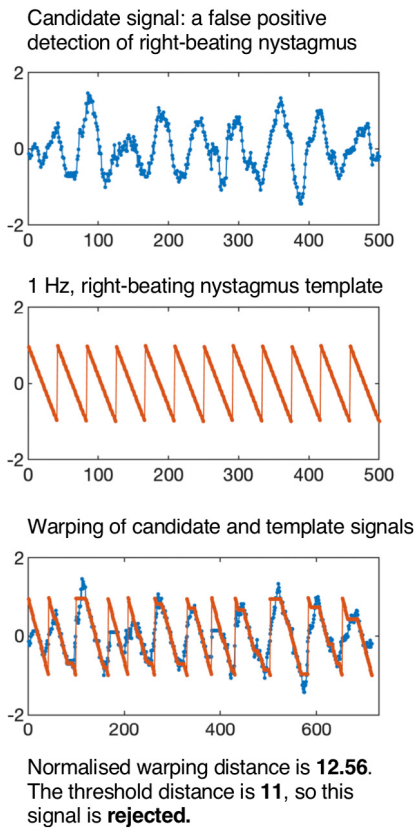

Fig. 4. The process of validating candidate nystagmus signals aims to reduce false positive detections, such as the candidate signal shown here. This signal is evidently periodic, which is likely why the detection has occurred. Dynamic Programming is used to generate a distance between the test signal and a 1 Hz, right-beating nystagmus waveform. This distance is normalised by the length of the test signal and the candidate is accepted if this value is lower than the specified threshold.

signal, and the decision to accept or reject the candidate waveform was made by way of a threshold value. If any positively classified frames remained after these steps, then the entire day's worth of data is classified as a positive detection of nystagmus.

Finally, for the classification of nystagmus frequency, the FFT was calculated for each validated nystagmus event. Using the output from the FFT, the modal frequency bin was identified and the frequency of the nystagmus event was classified as the closest class to that bin. In the case of a tie between bins, the decision was based on the highest neighbouring bin.

5.4. Detection and quantification of nystagmus induced by Caloric testing

The algorithm described in this section was applied to data captured by the CAVA device during a caloric test. It differs from the approach described in Section 5.3, in that it attempts to identify and quantify individual beats of nystagmus. The algorithm exploits the characteristic shape of the nystagmus waveform by finding straight lines with a steep gradient (fast phases) that are followed by shallow gradient lines of the opposite sign (slow phases). Best fit lines are used to determine the straightness of candidate waveforms, and the features of these lines also provide information from which nystagmus can be validated and quantified.

Firstly, the horizontal eye-movement data is pre-processed using a low-pass filter with a 5 Hz cut-off frequency. This filter reduces high frequency noise whilst largely retaining the characteristic shape of the nystagmus waveform. Potential fast phases are identified by using the velocity of the signal to find consecutive velocities of the same sign. The number of samples in each fast phase is iteratively increased and a best fit line is applied in each step. A number of thresholds determine when to stop adding new samples, and following that, whether each candidate is valid (e.g. a minimum and maximum gradient, the R-squared value of the fit, a minimum magnitude etc.).

A fast phase must be followed by a slow phase for it be a valid beat of nystagmus, and thus a best fit line is applied to the 10 samples following each fast phase. As before, the number of samples in the line is iteratively increased and the best-fit line recalculated. Thresholds determine when to this process and if the resulting slow phase is valid.

Following the detection of nystagmus within the data, the characteristics of the nystagmus are quantified. Using the data captured during the VNG calibration, a normalisation factor is calculated manually for each participant. This factor represents the number of quantisation levels per degree of eye movement. This factor is applied to the unfiltered signal, to limit the effects of signal distortion. The beat direction for a given period of nystagmus is determined by majority vote. The SPV of each beat is determined by the slope of the best fit line applied to the slow phase. The maximum SPV is calculated as the maximum value from a 10 s moving average, only considering windows with at least three beats in the same direction. Only beats detected after the start of each irrigation are included in the analysis, and all times are calculated relative to the start of each ear irrigation.

Table 1
Results of the blinded recognition experiment for detecting nystagmus in long-term, horizontal eye movement data. These results relate to the algorithm described in Sections 5.1–5.3, which uses a 3-classifier ensemble.

Task	tp	tn	fp	fn	Sens. (%)	Spec. (%)	F1
Detection	111	289	4	1	99.11	98.63	0.98

*tp = true positive, tn = true negative, fp = false positive, fn = false negative.

Table 2
Results of the blinded recognition experiment for classifying nystagmus direction and beat frequency in long-term, horizontal eye movement data. These results relate to the algorithm described in Sections 5.1–5.3, which uses a 3-classifier ensemble.

Task	Actual	Test **Left**	**Right**		Acc. (%)	Sens. (%)	Spec. (%)	F1
Dir.	**Left**	52	0		99.10	1.00	0.98	0.99
	Right	1	58			0.98	1.00	0.99
		0.8 Hz	**1.0 Hz**	**1.2 Hz**				
Freq.	**0.8 Hz**	35	0	0	98.20	1.00	0.99	0.99
	1.0 Hz	1	35	0		0.97	0.99	0.97
	1.2 Hz	0	1	39		0.98	1.00	0.99

6. Results

Table 1 shows the results of the blinded nystagmus detection task. 400 correct classifications were made by the algorithm, giving a recognition accuracy of 98.80%. Once the blinded investigator had submitted the final analysis and was subsequently unblinded, the false positive and false negative classifications were manually inspected. The single false negative was found to contain a noise impulse, caused by the participant pressing the device's event marker button midway through watching the video. It is therefore likely to have been misclassified due to a corruption of the signal in the frequency domain, causing the event to be discarded on the basis of duration or signal energy. The false positives were found to generally be periodic in nature, and observation of the concurrent accelerometer data showed that they might be due to physical activities, such as running. Visual inspection of the nystagmus induced whilst walking did not show significant evidence of motion artefacts.

For the two-class task of identifying the direction of the moving dots, the system attained an accuracy of 99.10% (110 correct and 1

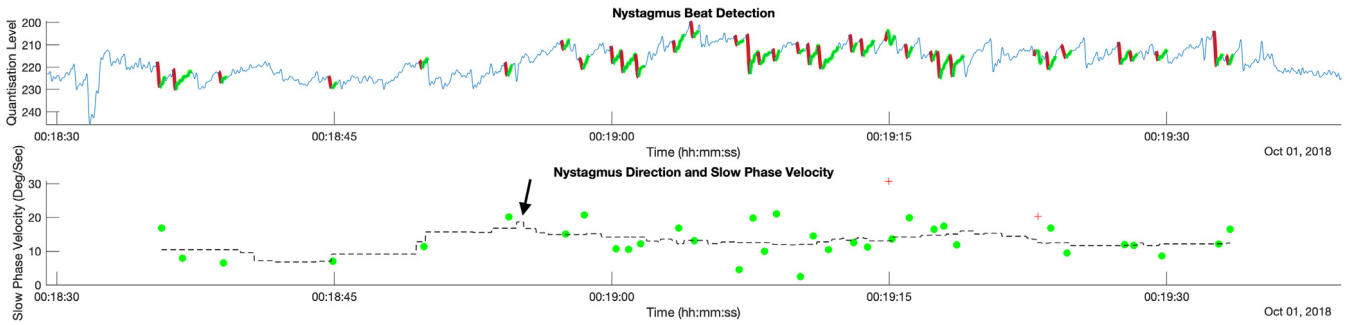

Fig. 5. Plot showing the computer analysis of a caloric irrigation from one participant. The top panel shows the times series waveform, with detected fast and slow phases highlighted (in red and green respectively). The example shown is of left-beating nystagmus, and this is confirmed by the direction of the highlighted fast phases. The bottom panel shows a quantification of the SPV for each of the detected beats (green circles for left-beating, and red crosses for right-beating false positives). A 10 s moving average of the SPV is shown as a black dotted line. The peak of this plot is the maximum SPV (black arrow). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

incorrect classifications). Finally, for the three class task of identifying the frequency of the moving dots, the accuracy was 98.20% (109 correct and 2 incorrect classifications). These data are presented in Table 2.

Fig. 5 provides an example of the computer analysis of the data from one participant, captured during a caloric test. The upper panel in the figure shows examples of fast and slow phases detected by the algorithm, which are highlighted in red and green respectively. The highlighted waveforms are visually representative of left-beating nystagmus, which is consistent with the test performed (a warm water irrigation of the left ear). It is interesting to note that this nystagmus usually spans less than ten quantisation levels, compared to the visually-induced nystagmus in Fig. 2, which spans about thirty levels. This is explained by the fact that the eye movements induced visually during the trial were more extreme than those induced by caloric testing.

A further analysis of this caloric data is presented in the bottom panel of Fig. 5. The green dots in this plot indicate the SPV of the left-beating nystagmus detected, and a small number of red crosses, which are false positive detections of right-beating nystagmus. The black dotted line shows a 10 s moving average of the average SPV, and the maximum SPV on this plot is 18.7 deg/s (black arrow).

Table 3 provides a comparison between the clinical scientist's measurements of the caloric data and the computer analysis. Of the twenty irrigations performed, eighteen were captured simultaneously using the CAVA device and the VNG system. 94% of the irrigations (16 of 17) were correctly detected by the computer analysis (i.e the majority of beats detected were in the expected direction). The two irrigations performed without the VNG system also showed evidence of nystagmus in the expected direction (i.e. left-beating for warm water irrigations of the left ear). For the correctly identified nystagmus events, the maximum slow phase velocity was estimated by a clinical scientist and also by the algorithm described in Section 5.4. Fig. 6 (orange circles) shows the relationship between the predicted velocities and the *ground-truth*, as estimated by the clinical scientist. In the majority of cases (16 of 17), the computer predicted velocities are over-estimated compared to the ground-truth. 29% of the computer estimates were within ±5 degrees of the clinical scientist's estimates. The maximum SPV estimated using the time recorded by the expert (blue crosses) was much closer to their estimate; 65% (12 of 17) of the computer estimates were within ± 5 degrees. An estimate was not generated for one sample as insufficient nystagmus was detected at the specified time. Table 3 shows the difference between the expert and computer predicted times of the maximum SPV. On average, the computer analysis reported the onset of the maximum SPV ± 41.2 s from the clinical scientist's estimate, which broadly explains why their estimates of the maximum SPV were different.

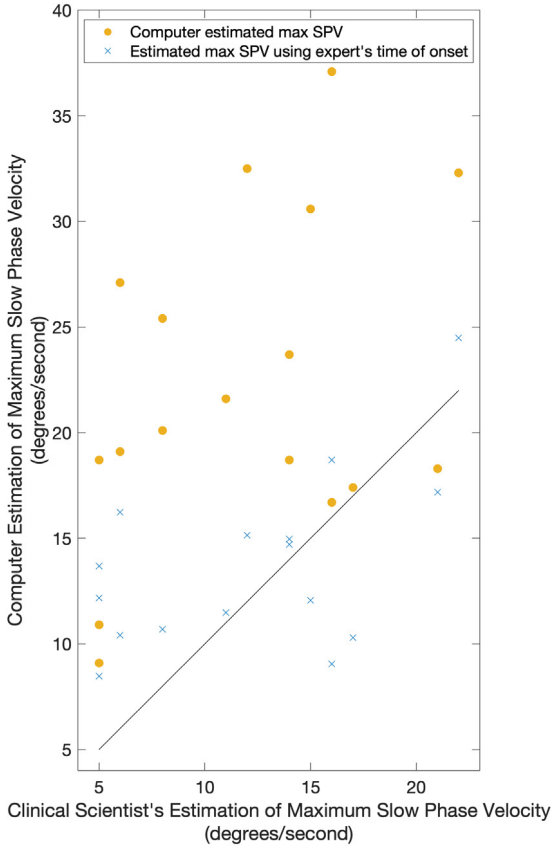

Fig. 6. Scatter plot showing the clinical scientist's estimated values of maximum slow velocities plotted against the values predicted by the computer estimation (orange circles) and estimated at the times provided by the scientist (blue crosses). The black line represents the ideal result, in which the estimates agree with the scientist's predictions.

7. Discussion

The results in Section 6 have demonstrated a very high degree of recognition accuracy for both the tasks of identifying the presence of the nystagmus waveform, and also the speed and direction of the moving dots. The ability to confidently identify short periods of visually-induced nystagmus from within days' worth of normal eye movement data is a promising step towards a complete system for diagnosing dizziness. These promising results also serve to validate the design of the CAVA device itself, which was successfully used to capture

Table 3
Results of a computer analysis of the caloric data compared to the results recorded by a clinical scientist. These results relate to the algorithm described in Section 5.4. The third column shows the difference between the expert's estimate of the time that the maximum SPV occurred, and the time estimated by computer analysis (e.g. a time of −6.54 s means the computer estimated time was earlier than the expert's time).

ID	Ear Irrigation	Max SPV time difference (s)	Nyst. beat direction
1	Right	106.90	Correct
1	Left	−3.86	Correct
2	Right	−27.12	Correct
2	Left	51.58	Correct
3	Right	57.72	Correct
4	Right	176.79	Correct
5	Right	−37.82	Correct
5	Left	−15.84	Correct
6	Right	25.91	Correct
6	Left	40.91	Correct
7	Right	0.66	Correct
7	Left	N/A	Incorrect
8	Right	−6.54	Correct
8	Left	−22.44	Correct
9	Right	54.16	Correct
9	Left	−44.78	Correct
10	Right	−5.87	Correct
10	Left	−21.9	Correct
	Mean Offset:	41.23	**%Acc:** 94.12%
	STD Offset:	43.70	

the trial data. The device was worn by 17 participants for up to 30 days, who in total captured around 9000 h of eye and head movement data.

In patients suffering from dizziness caused by inner ear malfunctions, nystagmus beat direction suggests the affected ear. The results for beat direction classification presented here demonstrate that the CAVA system can provide this valuable information. We have also shown that the beat frequency of periods identified as nystagmus can be recognised with good accuracy. Furthermore, the comparison between our computer analysis of the data captured during caloric testing, and that provided by the clinical scientist, shows that it is possible to estimate detailed characteristics of nystagmus which are of direct clinical importance. Regarding future applications of the data captured by the CAVA system, such as episode onset prediction, it is important to note that CAVA currently operates in a post-processing, rather than a realtime modality. Realtime analysis would increase the power consumption of the CAVA device, thus limiting its long-term data capture capabilities. However, it is conceivable that these limitations will be overcome in time.

Although we were able to detect nystagmus from the caloric data with a high degree of accuracy, the estimates of the maximum SPV and its onset, were less accurate. Table 3 shows that the onset of the maximum SPV reported by the computer analysis was almost always different to the expert's estimate. Prior to calculating the maximum SPV, a clinical scientist will manually validate each detected beat; a step which our automated approach does not replicate. It is therefore possible that the computer estimates are being skewed by incorrect data points or outliers. This explanation is supported by the fact that the estimated maximum SPV was more accurate when using the times recorded by the clinical scientist. Another explanation relates to the variability of the corneo-retinal potential (CRP). The CRP has shown to be influenced by many factors, including light, fatigue and diurnal rhythm [34], and although a calibration step was performed, it is possible that these variations might account for the differences observed. It should be noted that analysing caloric data is a nuanced and subjective task, and our ground-truth is one expert's analysis. Therefore, it is possible that other experts may interpret the data differently, and their opinion may be closer to or further from our computer analysis. The primary purpose of this analysis was to determine whether vestibular induced nystagmus was identifiable from the data captured by the CAVA device. In reality, the CAVA system is intended to identify periods of dizziness rather than to provide a quantitative assessment of inner ear physiology, which are often poorly correlated.

8. Conclusion

The results presented here have demonstrated the suitability of the CAVA device for long-term ambulatory monitoring, and the capability of our algorithms to detect short periods of artificially induced nystagmus buried within days' worth of normal eye movement data. In our future work, we intend to build on this progress by addressing some of the limitations of this study. Firstly, we intend to make the recognition system more robust to short impulses of noise. Secondly, we intend to focus on finding robust methods of normalising the CRP and for validating detected nystagmus, to allow more accurate estimation of the maximum SPV and other characteristics. Thirdly, our next body of work will include a clinical study on patients suffering from *real* dizziness conditions, to test our detection algorithms on nystagmus data from actual patients.

The approach used to detect nystagmus from the trial data exploited the periodic nature of the visually-induced nystagmus produced, which was achieved through the use of frequency domain recognition features. While this approach has shown to be very effective for the detection of nystagmus in our trial data, it is unclear how this technique will perform when applied to nystagmus produced as a sign of real dizziness conditions, such as Ménière's Disease or Benign Paroxysmal Positional Vertigo, especially if such nystagmus is less periodic. Capturing continuous and long-term data related to nystagmus arising from different diseases will also provide insight into the underlying mechanisms driving these diseases which has been impossible hitherto.

The dataset collected during this clinical trial is believed to be the largest dataset of eye movement data ever captured. These data will certainly be useful as we further refine our device and algorithms, but it may also be of interest within the areas of sleep analysis, activity detection, and for other areas of medical research. Additionally, the device could be used to capture new data in these areas, and, as this device is the first of its kind, it is possible that it will have other applications beyond the field of medicine. The long-term aim of this work is to improve patient care by speeding up diagnosis, reducing the number of hospital visits and the number of diagnostic tests performed, which in turn will have cost saving implications. The work presented here represents an incremental step towards this objective.

Declaration of competing interest

None declared.

CRediT authorship contribution statement

Jacob L. Newman: Conceptualisation, methodology, software, data acquisition, analysis and interpretation of data, drafting and final approval of the article. **John S. Phillips:** Conceptualisation, methodology, data acquisition, analysis and interpretation of data, drafting and final approval of the article. **Stephen J. Cox:** Conceptualisation, methodology, software, data acquisition, analysis and interpretation of data, drafting and final approval of the article. **John FitzGerald:** Conceptualisation, methodology, data acquisition, analysis and interpretation of data, drafting and final approval of the article. **Andrew Bath:** Conceptualisation, methodology, drafting and final approval of the article.

Acknowledgements

Thank you to the CAVA project's scientific advisory committee, who provided valuable advice during all stages of the clinical investigation underpinning this work. We would also like to acknowledge Wright Design Limited who designed and produced the CAVA device, and also the UEA's High Performance Computing Cluster for their services. This is work is funded by the Medical Research Council (MRC). The MRC reviewed the study design but were not involved with any other aspects of this work.

References

[1] N. Roydhouse, Vertigo and its treatment, Drugs (ISSN: 1179-1950) 7 (3) (1974) 297–309, http://dx.doi.org/10.2165/00003495-197407030-00004.

[2] D.E. Newman-Toker, L.M. Cannon, M.E. Stofferahn, R.E. Rothman, Y.H. Hsieh, D.S. Zee, Imprecision in patient reports of dizziness symptom quality: a cross-sectional study conducted in an acute care setting., Mayo Clin. Proc. (ISSN: 0025-6196) 82 (11) (2007) 1329–1340, http://dx.doi.org/10.4065/82.11.1329.

[3] H. Benecke, S. Agus, G. Goodall, D. Kuessner, M. Strupp, The burden and impact of vertigo: Findings from the REVERT patient registry, Front. Neurol. (ISSN: 1664-2295) 4 (2013) 136, http://dx.doi.org/10.3389/fneur.2013.00136.

[4] U. Büttner, T.B. Ch. Helmchen, Diagnostic criteria for central versus peripheral positioning nystagmus and vertigo: a review, Acta Otolaryngol. 119 (1) (1999) 1–5, http://dx.doi.org/10.1080/00016489950181855.

[5] D.E. Kentala, D.S.D. Rauch, A practical assessment algorithm for diagnosis of dizziness, Otolaryngol. Head Neck Surg. 128 (1) (2003) 54–59, http://dx.doi.org/10.1067/mhn.2003.47, PMID: 12574760.

[6] B. Cohen, V. Matsuo, T. Raphan, Quantitative analysis of the velocity characteristics of optokinetic nystagmus and optokinetic after-nystagmus, J. Physiol. 270 (2) (1977) 321–344.

[7] N.G. Henriksson, Speed of slow component and duration in caloric nystagmus, Acta Otolaryngol. 46 (sup125) (1956) 3–29, http://dx.doi.org/10.3109/00016485609120817.

[8] S.A. Bhansali, V. Honrubia, Current status of electronystagmography testing, Otolaryngol. Head Neck Surg. 120 (3) (1999) 419–426.

[9] M. Sangi, B. Thompson, J. Turuwhenua, An optokinetic nystagmus detection method for use with Young children., IEEE J. Trans. Eng. Health Med. (ISSN: 2168-2372) 3 (2015) 1600110, http://dx.doi.org/10.1109/JTEHM.2015.2410286.

[10] T. Pander, R. Czabanski, T. Przybyla, D. Pojda-Wilczek, An automatic saccadic eye movement detection in an optokinetic nystagmus signal, Biomed. Tech. (Berl) (ISSN: 1862-278X) 59 (6) (2014) 529–543, http://dx.doi.org/10.1515/bmt-2013-0137.

[11] J. Turuwhenua, T.Y. Yu, Z. Mazharullah, B. Thompson, A method for detecting optokinetic nystagmus based on the optic flow of the limbus, Vis. Res. (ISSN: 0042-6989) 103 (2014) 75–82, http://dx.doi.org/10.1016/j.visres.2014.07.016, http://www.sciencedirect.com/science/article/pii/S0042698914001758.

[12] T. Charoenpong, P. Pattrapisetwong, V. Mahasitthiwat, A new method to detect nystagmus for vertigo diagnosis system by eye movement velocity, in: 2015 14th IAPR International Conference on Machine Vision Applications (MVA), 2015, pp. 174–177, http://dx.doi.org/10.1109/MVA.2015.7153161.

[13] A. Ben Slama, A. Mouelhi, M.A. Cherni, S. Manoubi, C. Mbarek, H. Trabelsi, M. Sayadi, Features extraction for medical characterization of nystagmus, in: 2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), 2016, pp. 292–296, http://dx.doi.org/10.1109/ATSIP.2016.7523094.

[14] U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, H. Adeli, Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals, Comput. Biol. Med. (ISSN: 0010-4825) 100 (2018) 270–278, http://dx.doi.org/10.1016/j.compbiomed.2017.09.017, http://www.sciencedirect.com/science/article/pii/S0010482517303153.

[15] S. Chambon, V. Thorey, P.J. Arnal, E. Mignot, A. Gramfort, A deep learning architecture to detect events in EEG signals during sleep, in: 2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP), 2018, pp. 1–6, http://dx.doi.org/10.1109/MLSP.2018.8517067.

[16] U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, M. Adam, A. Gertych, R.S. Tan, A deep convolutional neural network model to classify heartbeats, Comput. Biol. Med. (ISSN: 0010-4825) 89 (2017) 389–396, http://dx.doi.org/10.1016/j.compbiomed.2017.08.022, http://www.sciencedirect.com/science/article/pii/S0010482517302810.

[17] S. Kiranyaz, T. Ince, M. Gabbouj, Real-time patient-specific ECG classification by 1-d convolutional neural networks, IEEE Trans. Biomed. Eng. (ISSN: 0018-9294) 63 (3) (2016) 664–675, http://dx.doi.org/10.1109/TBME.2015.2468589.

[18] A. Shoeb, J. Guttag, Application of machine learning to epileptic seizure detection, in: Proceedings of the 27th International Conference on International Conference on Machine Learning, in: ICML'10, Omnipress, USA, ISBN: 978-1-60558-907-7, 2010, pp. 975–982, http://dl.acm.org/citation.cfm?id=3104322.3104446.

[19] R.M. Mehmood, H.J. Lee, Emotion classification of EEG brain signal using SVM and KNN, in: 2015 IEEE International Conference on Multimedia Expo Workshops (ICMEW), 2015, pp. 1–5, http://dx.doi.org/10.1109/ICMEW.2015.7169786.

[20] V. Krasteva, I. Jekova, R. Leber, R. Schmid, R. Abächerli, Real-time arrhythmia detection with supplementary ECG quality and pulse wave monitoring for the reduction of false alarms in ICUs, Physiol. Meas. 37 (8) (2016) 1273–1297, http://dx.doi.org/10.1088/0967-3334/37/8/1273.

[21] C. Song, K. Liu, X. Zhang, L. Chen, X. Xian, An obstructive sleep apnea detection approach using a discriminative hidden markov model from ecg signals, IEEE Trans. Biomed. Eng. (ISSN: 0018-9294) 63 (7) (2016) 1532–1542, http://dx.doi.org/10.1109/TBME.2015.2498199.

[22] U. Satija, B. Ramkumar, M.S. Manikandan, Robust cardiac event change detection method for long-term healthcare monitoring applications, Healthcare Technol. Lett. (ISSN: 2053-3713) 3 (2) (2016) 116–123, http://dx.doi.org/10.1049/htl.2015.0062.

[23] D. Ngo, B. Veeravalli, Design of a real-time morphology-based anomaly detection method from ECG streams, in: 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2015, pp. 829–836, http://dx.doi.org/10.1109/BIBM.2015.7359793.

[24] R. Gravina, G. Fortino, Automatic methods for the detection of accelerative cardiac defense response, IEEE Trans. Affect. Comput. (ISSN: 1949-3045) 7 (3) (2016) 286–298, http://dx.doi.org/10.1109/TAFFC.2016.2515094.

[25] A. Garcés Correa, L.L. Orosco, P. Diez, E. Laciar Leber, Adaptive filtering for epileptic event detection in the EEG, J. Med. Biol. Eng. (ISSN: 2199-4757) (2019) http://dx.doi.org/10.1007/s40846-019-00467-w.

[26] J.S. Phillips, J.L. Newman, S.J. Cox, An investigation into the diagnostic accuracy, reliability, acceptability and safety of a novel device for continuous ambulatory vestibular assessment (CAVA), Scientific Reports 9 (1) (2019) 10452, http://dx.doi.org/10.1038/s41598-019-46970-7, https://doi.org/10.1038/s41598-019-46970-7,

[27] O.H. Mowrer, T.C. Ruch, N.E. Miller, The corneo-retinal potential difference as the basis of the galvanometric method of recording eye movements, Am. J. Physiol.-Legacy Cont. 114 (2) (1935) 423–428.

[28] D.L. McCaslin, Electronystagmography/videonystagmography, Plural Publishing, 2012.

[29] E.D. Whalen, Effects of instructions on optokinetic nystagmus (OKN), Ph.D. thesis, 2014.

[30] J. Friedman, T. Hastie, R. Tibshirani, Additive logistic regression: a statistical view of boosting, Ann. Statist. 28 (1998) 2000.

[31] A. Bagnall, J. Lines, A. Bostrom, J. Large, E. Keogh, The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances, Data Min. Knowl. Discov. (ISSN: 1573-756X) 31 (3) (2017) 606–660, http://dx.doi.org/10.1007/s10618-016-0483-9.

[32] A. Bangham, R. Harvey, D. Ling, Paul, R. Aldridge, Morphological scale-space preserving transforms in many dimensions, J. Electron. Imaging 5 (1997) http://dx.doi.org/10.1117/12.243349.

[33] H. Sakoe, S. Chiba, Dynamic programming algorithm optimization for spoken word recognition, IEEE Trans. Acoust. Speech Signal Process. (ISSN: 0096-3518) 26 (1) (1978) 43–49, http://dx.doi.org/10.1109/TASSP.1978.1163055.

[34] L. Proctor, D. Hansen, R. Rentea, Corneoretinal potential variations: significance in electronystagmography., Arch. Otolaryngol. (Chicago, Ill. : 1960) (ISSN: 0003-9977 0003-9977) 106 (5) (1980) 262–265.