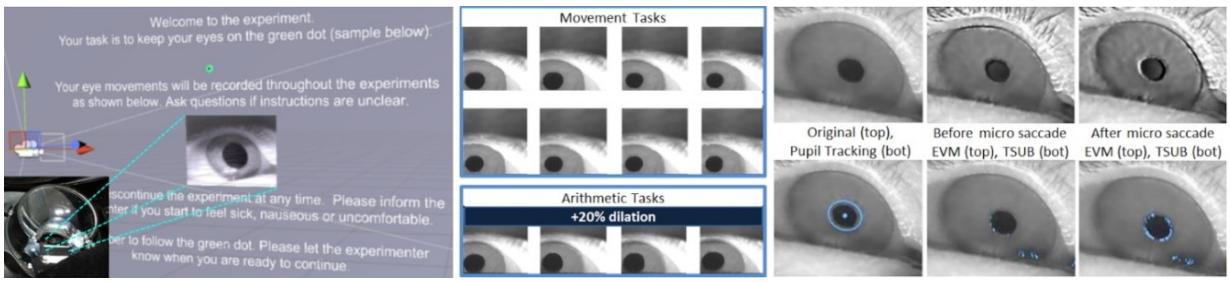
Emulation of Physician Tasks in Eye-tracked Virtual Reality for Remote Diagnosis of Neurodegenerative Disease

Jason Orlosky, Yuta Itoh, Maud Ranchet, Kiyoshi Kiyokawa, John Morgan, and Hannes Devos



A) B) C)

Fig 1. A) The introduction screen of our VR remote diagnosis interface, which emulates tasks used by physicians to diagnose patients with neurodegenerative disease. A custom IR eye camera is integrated into the lens of an Oculus Rift, as shown in the lower left corner. B) Frames taken from the eye camera showing pupillary response to virtual movement tasks (top) and arithmetic tasks (bottom). Evoked pupil response is visible in the bottom row, and is one of many state-of-the-art methods used to study neurodegeneration [18]. C) Frames taken from videos of a patient with Parkinson's disease when completing a VR fixation task. The left column shows the original (top) and pupil-tracked overlay (bottom). These frames are then enhanced or amplified to reveal subtle motion abnormalities to physicians. The middle and right columns show pre- and post-microsaccadic movement enhanced with our custom thresholding and adaptive background subtraction enhancement (TSUB, bottom), and Eulerian Video Magnification (EVM, top) [38]. *Note: Patients' iris contours are masked to protect privacy.*

Abstract—For neurodegenerative conditions like Parkinson's disease, early and accurate diagnosis is still a difficult task. Evaluations can be time consuming, patients must often travel to metropolitan areas or different cities to see experts, and misdiagnosis can result in improper treatment. To date, only a handful of assistive or remote methods exist to help physicians evaluate patients with suspected neurological disease in a convenient and consistent way. In this paper, we present a low-cost VR interface designed to support evaluation and diagnosis of neurodegenerative disease and test its use in a clinical setting. Using a commercially available VR display with an infrared camera integrated into the lens, we have constructed a 3D virtual environment designed to emulate common tasks used to evaluate patients, such as fixating on a point, conducting smooth pursuit of an object, or executing saccades. These virtual tasks are designed to elicit eye movements commonly associated with neurodegenerative disease, such as abnormal saccades, square wave jerks, and ocular tremor. Next, we conducted experiments with 9 patients with a diagnosis of Parkinson's disease and 7 healthy controls to test the system's potential to emulate tasks for clinical diagnosis. We then applied eye tracking algorithms and image enhancement to the eye recordings taken during the experiment and conducted a short follow-up study with two physicians for evaluation. Results showed that our VR interface was able to elicit five common types of movements usable for evaluation, physicians were able to confirm three out of four abnormalities, and visualizations were rated as potentially useful for diagnosis.

Index Terms—Virtual reality, eye tracking, diagnosis, visualization

1 Introduction

In neurology, accurate diagnosis of neurodegenerative conditions is still a difficult problem. The clinical picture of many parkinsonian disorders, such as Parkinson's disease (PD), Progressive Supranuclear Palsy, Multiple System Atrophy, and Cortico-basal Degeneration, is often quite similar. These disorders are therefore classified as parkinsonism. Although neuroimaging techniques may help differentiate between these neurodegenerative conditions, the diagnosis is usually made in the clinic based on visible clinical signs and symptoms. However, patients often have to wait for years for a correct diagnosis. Improving the speed, consistency, and portability of methods for diagnosis is critical for delivering the right treatment and medication at the right time.

An intrinsic part of the skill set of neurologists is the identification of eye movement abnormalities [37], which are common in Parkinsonism and can sometimes be elicited by physical clinical tests. An example of the most basic of these is to have a patient fixate on the physician's finger and to watch the patient's eye movements for abnormalities during the task. Other tools such as eye tracking interfaces are available to give the physician a better view of the patient's eye movements [30]. However, even with current eye tracking interfaces, physicians are still prone to error. Misdiagnosis is also still common, especially at initial stages of certain diseases,

- *Jason Orlosky is with Osaka University. E-mail: orlosky[at]lab.ime.cmc.osaka-u.ac.jp.*
- *Yuta Itoh is with Keio University. E-mail: itoh[at]imlab.ics.keio.ac.jp.*
- *Maud Ranchet is from The French Institute of Science and Technology for Transport, Development, and Networks. E-mail: maud.ranchet[at]ifsttar.fr.*
- *Kiyoshi Kiyokawa is from Osaka University. E-mail: kiyo[at]ime.cmc.osaka-u.ac.jp.*
- *John Morgan is from Augusta University. E-mail: jmorgan[at]augusta.edu.*
- *Hannes Devos is from University of Kansas Medical Center. E-mail: hdevos[at]kumc.edu.*

Manuscript received 19 Sept. 2016; accepted 10 Jan. 2017.
Date of publication 26 Jan. 2017; date of current version 18 Mar. 2017.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TVCG.2017.2657018

including PD. A physician's finger movements may not be adequate in evoking a response and minute abnormalities are not always visible due to limited resolution or perception. Both physician and patient must also be physically present in the same area. To address these problems, this paper presents a portable virtual reality (VR) interface for assisting with remote clinical assessment. The interface includes three core components, including a 3D environment designed to provide consistent emulation of physician tasks, an integrated near-eye eye tracking camera (and custom eye tracking software) to allow for automated and post-task analysis of eye movements, and algorithms and visualizations to assist with accurate detection and discovery of abnormalities. Whereas some other systems for evaluation utilize game performance to determine certain neurodegenerative characteristics like memory loss [4][27], we rely on the fact that the eye movements of patients with parkinsonism have consistently abnormal responses [37], which can even be used to aid in the differential diagnosis of these movement disorders [1].

After consulting with specialists in neurology and neuroscience, we created and tested a set of 3D tasks designed explicitly to evoke the same responses (abnormal saccades, square wave jerks, ocular tremor, etc.) evaluated by clinicians in state-of-the-art methodology. These are simulated with a controlled VR environment, and do not require a physician to be present. Secondly, our eye camera is software focusable and integrated into the Oculus Rift's lens system to obtain a direct, high resolution view of the pupil, as shown in the images in Figures 1 and 2. For evaluating oculomotor metrics, this custom camera mount meets or exceeds the resolution and clarity of many expensive commercial half-mirror implementations [19], which usually just provide an upscaled view of the eye camera output.

Our system also employs several different computer vision techniques to help physicians visualize these changes, such as motion magnification (EVM) [38], a method we call thresholded background subtraction (TSUB), and virtual overlay of the pupil contour.

Lastly, we conduct both a clinical experiment with patients and a follow-up physician evaluation to measure the practicality and usability of the system as a whole. As a baseline, the clinical experiment with patients is designed to test the use of the VR interface with individuals who have a diagnosis of Parkinson's disease (PD), as well as a control group with no diagnosis of PD. The experiment included a number of different tasks designed to evoke the same responses as typical physician tasks, and videos of the patients' eyes were recorded for post-processing. The follow-up experiment with physicians was designed to confirm detection of abnormalities found in the primary experiment, evaluate the different visualizations for their usefulness in detecting abnormal movements, and to compare algorithmic and human recognition of abnormalities. In short, our primary contributions include:

- a framework to combine eye tracking, VR, and visualization to evoke eye movements, reveal abnormalities, and assist with clinical diagnosis of various neurological disorders,
- a series of experiments testing the interface with a group of patients with a diagnosis of Parkinson's disease vs controls,
- and a pilot test with physicians to verify detection results and evaluate visualizations.

2 Prior Work

A majority of prior work falls into two categories: 1) the use of AR and VR for evaluation and rehabilitation of patients and 2) the study of eye movements related to neurodegenerative diseases. Since eye tracking has become widely available, a number of studies have appeared using the eyes as a method for discovering new biomarkers for certain diseases, especially those affecting motor control.

2.1 Eye Tracking and Virtual Interfaces in Clinical Evaluation

For example, Granholm et al. studied the characteristics of the pupil's reflex to light in PD patients using an IR tracker [10]. A similar study in 2009 by Fotiou et al. used eye tracking to measure pupil metrics, discovering that constriction acceleration was the best predictor in classifying a subject as normal [6]. Such eye tracking studies have also motivated several studies in VR that use game performance or other metrics to evaluate patients [29][31], or even to modify existing memory [21].

For example, in 2000 Gupta et al. developed an eye tracked VR system called the VREye that determined a user's gaze patterns for image search tasks [9]. Patients with memory loss or aphasia would view the images differently than controls, for example by only processing one side of the image during search. Very similarly, Cushman et al. used game performance to determine cognitive deficits in the elderly and individuals with Alzheimer's disease [4]. In 2012, Plancher et al. used VR to characterize memory profiles by immersing patients in different virtual environments and measuring recall and recognition for a variety of tasks [25]. More recently, similar interfaces have been developed for studying eye trajectory and gait [34], retinal screening [13], and measuring the eye's refractive error [21].

Gitchel et al. have developed a PC based eye tracking interface to help evaluate patients using a head mounted eye tracker in combination with a monitor [8]. Their eye tracker has very high temporal resolution, but the eye-cameras are not near-eye, limiting resolution. This is one of the techniques we seek to replicate in VR. Moreover, existing interfaces are limited in portability, and a specialist must be present to calibrate the device and administer tests. The system developed by Marx et al. is likely the closest to what we are proposing [19]. They used a half-mirror based head mounted eye tracker that was used to evaluate patients with both a fixed protocol and "in the wild," meaning they used data taken from everyday activities for analysis, though this takes time. These data were evaluated to help differentiate between Parkinson's disease, Progressive Supranuclear Palsy, and healthy controls.

In contrast with these studies, our proposed VR system can give physicians much more control (such as vergence adjustment) over the types of eye gaze tasks used for evaluation. The tasks can also be administered remotely and consistently between patients, and take less than 20 minutes to complete. Moreover, we develop an in-house eye tracking method to compensate for poor lighting or patients that failed to wear the device properly, which can provide better results for automated detection in the future. Lastly, in addition to the virtual environment and analysis of the eye signal itself, we provide augmented visual feedback for physicians so that they can make more informed decisions about diagnoses.

2.2 Neurodegeneration and Oculomotor Measures

Parkinson's disease (PD) is the second most prevalent neurodegenerative condition [34]. Hallmark characteristics of the disease are shaking (tremor), slowness (bradykinesia), muscle stiffness (rigidity), postural instability, and gait disorders. Many other neurodegenerative conditions present with a similar clinical picture, but are not considered PD. These are classified as Parkinson plus disorders and also fall under the umbrella term parkinsonism. Accurate diagnosis of these different forms of parkinsonism is critical to deliver the best medical treatment. Early attempts to measure motor skills via computers date back to 1986 [11], and different techniques have been developed over the years to assist with differential diagnosis of these conditions, including the detection of eye movement abnormalities.

In parkinsonism, several brain areas from the brain stem to higher cortical areas that are crucial for saccade generation and smooth pursuit may be disturbed [15]. In the differential diagnosis of

parkinsonism and their related eye movement disorders, clinicians currently focus on the execution of saccadic movements and smooth pursuit. The role of saccades is to direct the fovea of the eye to an object of interest, and smooth pursuit eye movements are responsible for keeping the fovea on the object. Smooth pursuits are defined as conjugate eye movements that track a moving object in order to maintain fixation [24]. There are different forms of saccadic movements, and visually guided prosaccades and antisaccades are of particular interest in the differential diagnosis of parkinsonian disorders. Prosaccades are reflexive saccades towards the object of interest. Antisaccades are saccades to the opposite side of the object of interest. The initiation of these saccades may be delayed (increased saccadic latency), the saccades may be slower (bradykinesia) than anticipated, or may not shift exactly on the target (hypometria) in patients with parkinsonism.

Unique findings as well as subtle differences in oculomotor abnormalities within apparently similar conditions may be elicited if the patient is examined carefully [5]. For example, patients with progressive supranuclear palsy, a type of parkinsonism, present with vertical gaze palsy, which is less likely to be present in other types of parkinsonism. Some of these eye movement abnormalities are more subtle and hard to detect with the naked eye. Recently, Gitchel found miniscule ocular tremor (oscillatory behavior) in PD patients while fixating on an object [7]. This ocular tremor, which appears only to be present in PD, can currently only be detected with infrared eye trackers and Fourier analysis due to the tremor's small amplitude (~0.27 degrees). Square wave jerks are also often present, which are defined as initial saccades that move the fovea away from the object and are followed by a second saccade in the opposite direction that realigns the fixation position with the fovea. Simply put, they are involuntary, horizontal, saccadic intrusions that interrupt fixation, and are a frequent feature in parkinsonism.

One of our goals is to create virtual tasks that evoke many of these oculomotor abnormalities, which can provide a mechanism with which physicians can remotely detect these movements and ultimately make a better diagnosis.

3 System Framework

Here we describe the system as a whole, including the hardware for VR and eye tracking, software environment used to emulate physician tasks, eye tracking process and algorithms, and visualizations for motion enhancement.

As our VR device, we used the Oculus Rift DK2. Several companies such as Sensomotoric Instruments (SMI) actually offer eye-tracking attachments for the VR displays, but these modifications are largely for determining gaze points in the 3D virtual world. Devices like SMI's eye tracker also tend to use a half-mirror setup, which increases the distance between the camera and eye. In contrast, we needed a less expensive, more flexible camera that could record closer, higher resolution videos of the eye for analysis and allowed us to control focal distance via software on a per-user basis.

Thus, we used the Pupil Labs' "Pupil" development kit eye tracker. This device, which uses the Microsoft Lifecam 6000 as its base camera, has software adjustable exposure and focal depth, which were necessary for properly adjusting the camera to the participant's eyes. In order to get the highest resolution image possible, we first carved out the lower portion of the Rift's eye-cup as shown in the lower left of Figure 1 and in Figure 2, and ran this at 640x480 pixel resolution. This allowed us to position the Pupil tracker closer to the wearer's eye than would normally allow. Though the camera occludes a small portion of the lower field of view of the right eye, participants were still easily able to conduct tasks. Most ocular abnormalities affect both eyes, so we only used a monocular camera, which also allowed users to maintain persistence of vision for objects in the lower field of view.

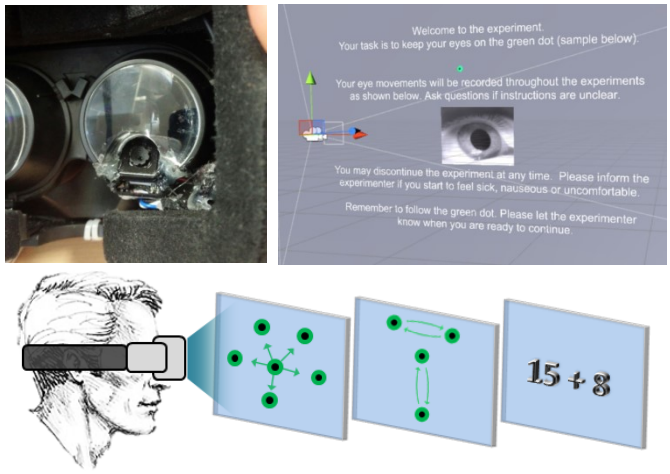


Fig 2. Top row: view from the lens side of the Rift showing the integrated eye tracker (left), and a side perspective view of the introduction screen in the 3D environment (right). Bottom row from left to right: Diagram of a user wearing the display, along with several sample positions of the icons for fixation, pro-saccade (changes in location), smooth pursuit (vertical/horizontal movement), and arithmetic tasks.

3.1 Environment and 3D Task Emulation

To replicate the same tasks used by doctors to detect oculomotor abnormalities, we had to emulate a number of different movements and activities. As described previously, three very revealing abnormalities include increased square wave jerks, ocular tremor during fixation tasks, and "catch-up" saccades during smooth pursuits. To elicit these abnormal movements, doctors primarily use a finger, pen, or object with a dot or point on which to focus, and move that point in different ways. These movements include tasks like holding the object in place and asking the patient to keep his or her eyes focused (a fixation) or moving the object back and forth or up and down (smooth pursuit task). In a majority of these tasks, the physician will simply observe the patient's eyes while conducting the task and watch for abnormal movements. Unfortunately, subtle abnormal movements such as ocular tremor can go unnoticed by the naked eye, despite the skill of the physician.

We then set out to emulate these tasks in VR. Our system improves on current methods in two primary ways: 1) the virtual tasks are automated and can be used remotely so that the physician does not have to be physically present or concentrate on moving an object of concentration and 2) the eye videos are recorded and visible from a near-eye camera position. To accomplish this, we built a series of tasks using Unity. The general layout of the VR environment is shown in Figure 2. We then built a C++ application using OpenCV and interfaced it directly with Unity to handle the synchronization, recording, and storage of eye-videos and head movement data to file. Integration of stereoscopic eye trackers into commodity VR displays will soon allow us to take videos and conduct analysis of both eyes.

The VR interface first included an instruction screen (Figure 2, top right) used to relay initial instructions to the participant and to give them a view of their own eye so they could adjust the display properly. During the instruction step in live experimentation, we also used the camera's software adjustable lens control to ensure that the image was in focus. Participants followed the small green dot visible in Figures 1 and 2, which was used to emulate the point-object that doctors use. To replicate fixation, pursuit, and other tasks, this dot was moved programmatically to emulate the way a physician would move his or her finger. A neutral, gray backdrop with two horizontal and two vertical pillars was used at an approximate distance of 1 Unity unit in front of the participant.

Moreover, this allows us to emulate tasks that are physically impossible. For example, we can instantly move the dot to a new location to elicit a saccade, whereas a physician has to use two fingers and ask a participant to switch between the two instead of instantly moving his or her finger from one location to another. For the experiment, we used tasks that would elicit fixations, smooth pursuits (vertical and horizontal), saccades, and evoked pupillary response. The exact nature and parameters of these tasks are described in detail in the Experiments section.

3.2 Custom Eye Tracking

Initially, we attempted using commercial eye tracking that makes use of the RANSAC algorithm [12][33]. Unfortunately, this algorithm is not robust to significant changes in lighting as shown in Figure 3, and failed to build a correct eye model for many of our videos. The short video duration (10-20s) may have also been a reason this algorithm failed to properly identify the pupil. We then decided to develop our own in-house algorithm to handle low contrast and defocused images on a frame by frame basis as described below.

To do this, our first and most important step was to detect a portion of the pupil region and set it as the region of interest (ROI) for further processing. Our pupil detector is essentially a low contrast area-detection algorithm that finds dark regions with minimal texture in dynamic scenes. The pupil region in most eye tracking applications is almost always the darkest, most uniform region in the image, so this algorithm works very well for finding the pupil when parameterized.

In short, the algorithm computes a darkness factor for a square region representing the minimum pupil size possible in any video. Each nth pixel in both x and y directions in the image is sampled and added to a total darkness factor for that particular sampling region. In other words, every nth pixel is sampled in both x and y directions when searching and also when computing darkness. The benefit of this strategy is that it can deal more effectively with shadows and eye lashes since the space between lashes often contains bright pixels.

The next step is to extract the shape of the pupil region before fitting an ellipse. This is typically handled by applying a threshold or line recognition algorithm to extract the borders of the pupil. However, in low contrast images, a single thresholding operation often fails to extract the full pupil region. Fitting an ellipse directly to such a contour would not only be unstable, but also inaccurate. If the number of points is too small, algorithms like RANSAC will fail to compute an appropriate ellipse.

To solve this problem, we have developed a dynamic thresholding algorithm that can more effectively extract correct points on the pupil ellipse despite low contrast.

3.2.1 Thresholding and Line Detection

The first step in this process is to use a multi-level thresholding and line detection process. In high contrast regions of an image where the pupil line is clear, it is easy to extract the pupil, however, sometimes part of the pupil is low contrast and part is high contrast, in which case a multi-step approach is necessary. To address these cases, we use two auto-thresholded binary images as shown in the bottom row of Figure 4. These thresholds are dynamically set every frame based on the darkest pixel area region mentioned above. This provides a number of very good pixels in low contrast regions, and a larger number of bad and good pixels in high and low contrast regions. Contours for the largest region are then added to a candidate point list. To maximize the number of correct pixels, candidate points from the low threshold image are then logical ANDed with an image processed with the Canny algorithm with very low parameters (i.e. more lines), as shown in the top row of Figure 4. Rather than a point by point ANDing process, any points within several pixels of the candidate point list that are also present in the Canny image are also included in a final list. Conversely, the thresholded image with both good and bad points is logical ANDed with an image processed with more restricted Canny parameters. The resulting points are then passed to a refinement process.

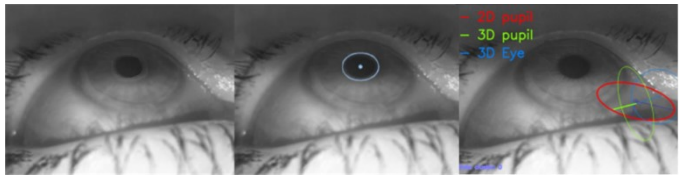


Fig 3. Image of a more challenging video frame (left), our pupil tracking and fitting method (middle), and failure of the RANSAC pupil tracking algorithm (right) [12][33]. *Note: Iris masked for privacy.*

3.2.2 Refinement

Once the lists of candidate points have been merged, they are passed through a refinement process, which primarily consists of two parts: 1) a line refinement, and 2) an ellipse refinement.

The line refinement process uses a similar to many line fitting algorithms by calculating a weighted sum of the differences in pixel intensities. This process is also parameterized with the width and length of the search window and the spacing (sparse sampling) between weighted points. Once this refinement process has been executed for all candidate points, each point goes through a secondary process using ellipse refinement. First, an ellipse is fit to all candidates using OpenCV's fitEllipse function [3]. While this function is good for fitting an ellipse to points that are guaranteed to be on the ellipse, it can fail easily when several points are outliers.

However, one notable characteristic of this method is that if a significant number of correct points (inliers) lie on the true ellipse, some number of inliers will lie on the fitted ellipse that represent a better fit, while a number of outliers will not lie on the fited ellipse at all. By taking advantage of this tendency, we can actually use comparison with the fitted ellipse to exclude many of the outliers present while still including a large number of inliers. An algorithm is then run that checks every candidate point against the fitted ellipse. If a candidate point is within several pixels (3 in our implementation) of any point on the fitted ellipse line, it is kept as an inlier and passed to a final line refinement stage. If the point is not near the fitted ellipse, it is excluded as an outlier and removed from any further refinement. In effect, this process continually excludes outliers, refines points on the pupil border, and comes closer to a perfect ellipse fit, which is high enough to detect < 0.1 degree eye movements. The smallest movements we are trying to study are

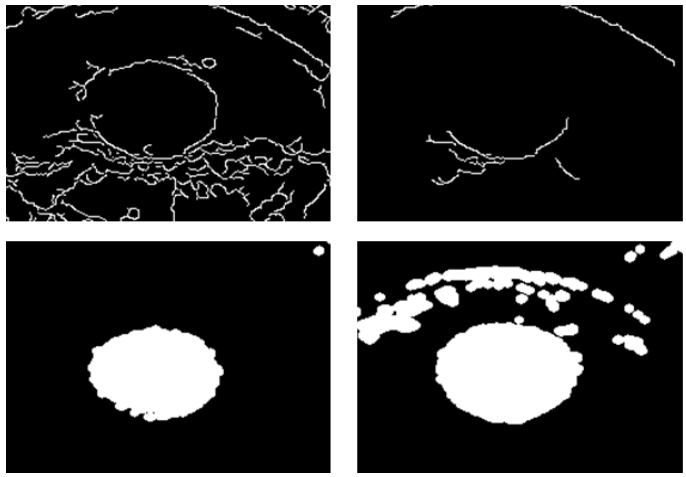


Fig 4. Images of Canny and Threshold [3] algorithms for the eye in Figure 3, which are ANDed (top-to-bottom) and then ORed (left-to-right) for a better representation of the pupil contour.

vertical oscillations (tremor), which have been shown to be approximately 0.27 degrees in amplitude by Gitchel et al. [7]. Our overall accuracy comes out to approximately one pixel, or ~.084 degrees. Other than one control participant and several videos for a PD participant where the pupils were almost entirely covered in shadow, we were able to properly extract the pupil region.

4 Visualizations

Though high resolution near-eye video recordings are a good step towards improving detection of abnormalities, doctors can often still miss small movements that may be representative of neurological disease. For example, the 0.27 degree amplitude of ocular tremor is nearly impossible to detect with the naked eye, and still difficult even with near-eye videos.

To assist with the detection and discovery of these types of abnormalities, we decided to implement and test a number of video enhancements that could emphasize or help visualize small movements. In addition to graphical representations and plots, video enhancements can potentially help doctors make better judgements and more consistent diagnoses. In addition to the original near-eye videos, we came up with three visualizations to reveal movement. These include a virtual overlay of the pupil center and contour ellipse, a custom background subtraction and filtering method (TSUB), and application of Eulerian Video Magnification (EVM) [38]. The TSUB and EVM implementations are described next in detail.

4.1 Thresholding and Background Subtraction

Though a number of motion magnification techniques already exist, we wanted to develop a method that was specifically designed for revealing oculomotor abnormalities. We started with a simple implementation of background subtraction, thinking that frame to frame changes would make for a good visualization. However, due to the various movements of the eye, including normal instability, basic background subtraction resulted in a very noisy result, where it was difficult to distinguish minute abnormalities from other movements.

To address this, we next incorporated a portion of the eye tracking algorithm into this background subtraction in hopes that it would amplify changes to the extent that detection was possible. In essence, the background subtraction algorithm is instead run on the thresholding algorithm shown in Figure 4, which results in the images in Figure 5. Though we originally intended just to show these images as-is, we decided that this would be too far from the original video to allow for a good comparison. During initial testing, even placing these images next to the original would cause observers to have to look back and forth. Instead, the white overlay is turned blue and then re-applied to the original video, as can be seen in the left column of Figure 6, so that physicians have a layered image of both the original eye as well as the highlighted differences in motion.

4.2 Eulerian Video Magnification

As an objective comparison, we also applied Eulerian Video Magnification, the implementation of which was written by Pan [23] as shown in the right-hand column of Figure 6. Motion amplification has been used to exacerbate various types of movements as early as

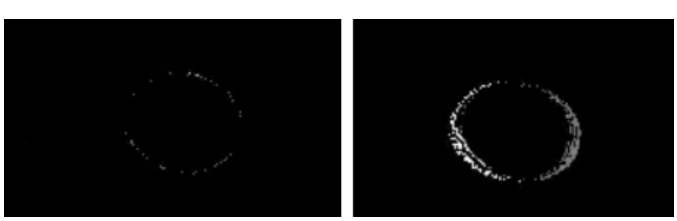


Fig 5. Intermediate stage of the TSUB algorithm, which shows movement differences of a thresholded image over time. The left image shows a static pupil and the right shows a microsaccade.

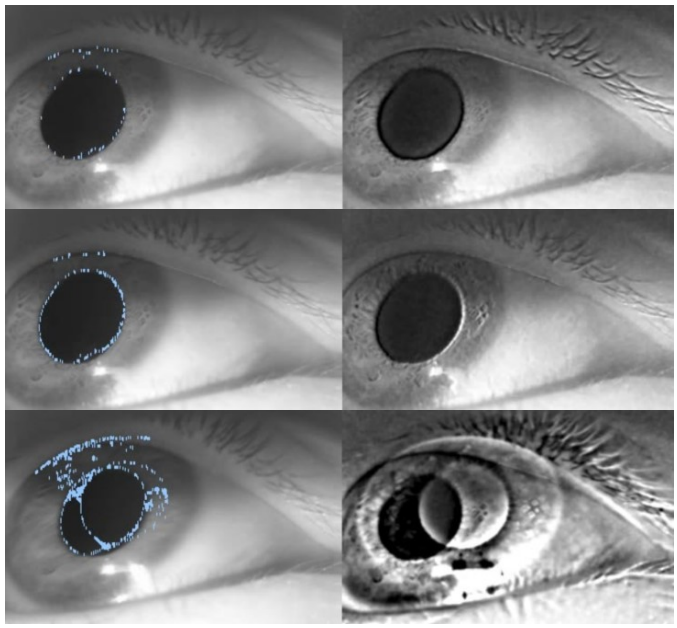


Fig 6. Images comparing TSUB (left column) and EVM (right column) visualizations for a static pupil (top row), microsaccade (middle row), and regular saccade (bottom row).

2005 [17]. More recently, it has been used for specific medical applications, such as visualizing pulse [38]. It has actually been applied to eye movements, but these movements have never been compared to other amplification techniques as a type of visualization. Moreover, we wanted to test to see how well EVM could actually assist physicians in a real world scenario.

Our next step was to design experiments that would first test the system in a movement disorders clinic, and to follow up with physicians afterward to confirm the usefulness of our core data and gather feedback on the visualizations.

5 Experiments

In addition to an overall test of the system, our experiments were designed to answer several research questions:

1) Can VR be used to emulate the same tasks doctors use in practice, and is a VR interface well received amongst patients in clinical practice?

2) Did the system we designed consistently evoke and detect responses representative of neurological disease?

3) Can physicians recognize and consistently evaluate these ocular abnormalities in a remote setting, and would visualizations potentially be useful in clinical practice?

To answer these questions, we set up two experiments. The first experiment tested the VR interface with a group of 10 patients with a prior diagnosis of PD as well as a group of 7 healthy controls. Participants conducted a number of virtual tasks, and their eye movements were recorded during the process. Note that due to excessive mascara and poor eye positioning, data from two patients were excluded from analysis, for a total of 9 patients and 6 controls.

The second experiment was a follow-up to have physicians confirm visual abnormalities and to see if video enhancements would be useful in clinical practice. We gave two physicians at different levels of seniority a detailed web-based questionnaire, which included eye-movement training, explanation of enhancements, and a series of evaluations designed to test detection success and consistency. These two experiments are described in detail below.

5.1 Primary Experiment: Testing the VR interface in a Clinical Setting

We first needed to test whether a VR interface would be usable in a clinical setting. The main purposes of this experiment was to determine whether the interface could successfully emulate the same tasks neurologists typically perform, evaluate whether patients were willing to use the device, and determine whether the interface successfully evoked eye abnormalities. Next, we describe the experiment setup as well as the tasks we designed to help elicit oculomotor responses.

5.1.1 Setup

As described previously, our system uses the eye-tracked Oculus Rift we designed in combination with a custom built 3D environment. This Rift was attached to an Alienware PC (intel core i7-5820k, 6-Core @ 3.8GHz, GeForce GTX 960) that was transported to an unused room in a neurology clinic that sees around 20 - 30 patients per day. Patients with suspected or diagnosed Parkinson's disease, sometimes with their spouses or friends, first had a consultation with a resident and/or primary neurologist as part of their scheduled appointment. Their clinical history, current medication levels, recent progress, and other data were assessed by the neurologist, and the patient (and healthy friend/spouse if present) was then asked if he or she would be willing to participate in Virtual Reality research for the purposes of studying eye movements.

A total of 17 individuals were willing to participate, and they all read and signed an informed consent prior to the start of the experiment. This experiment and informed consent were previously approved by an ethics approval committee. Once the informed consent process was complete, participants were then seated in a chair in a room with the experimenter, who explained the details of the tasks as described below.

Participants were asked to complete four different types of virtual tasks, each designed to elicit or study a different oculomotor movement. For the first three tasks, participants were simply instructed to follow the green icons shown in the bottom row of Figure 2, regardless of location, as described next.

5.1.2 Fixation

The first task was a basic fixation on a stationary point. As mentioned previously, this type of task can be used to elicit square wave jerks and ocular tremor. The green icon first appeared in one of the five locations, each spaced across the display's virtual field of view to ensure we had at least one trackable eye position despite different head shapes or eye positions. All of these were located at a depth of 12 in Unity coordinates, measured from the user's field of view. The Unity physics system treats each Unity unit as one meter in the real world, however it is likely that the actual perceived distance is different for each user due to interpupillary distance and the fact that the Oculus places the image for each eye at a wider disparity than if viewing the real world. Each set of icons, corresponding to those in Figure 2, had the following (x, y) positions: (-1, 3), (1, 3), (3, 0), (0, 3), (0, -2). In other studies, head movement (particularly oscillatory behavior) has been shown to cause the eyes to oscillate [7], making it more difficult to differentiate between vestibular reflex and ocular response characteristic of a disease.

To compensate for this, we disabled head tracking for the icons so that they were displayed in head-relative space, much like a 2D user interface element. This was done specifically to reduce the influence of vestibular reflex on eye movement.

Locations were randomized to alleviate ordering effects. The green icon appeared for a total of ten seconds (one trial), and ten seconds of video were recorded during the trial. At ten seconds, the icon disappeared, recording stopped, and a five second period followed where nothing was displayed except the gray backdrop. The total duration of this task for all ten trials was less than five minutes, and participants were allowed a short break before continuing to the next task.

5.1.3 Smooth Pursuit

The second task was a smooth pursuit of the same green icon. The icon would appear, oscillate for ten seconds from -15 degrees to +15 degrees relative to head-center, and disappear for five seconds before displaying the next icon. We included two different velocities, 10 and 40 degrees per second, and horizontal and vertical oscillation for a total of four trials in order to have a variety of different smooth pursuit tasks. Both velocity and direction of oscillation were randomized to alleviate ordering effects. The total duration of this task was under two minutes.

5.1.4 Saccade

The third task was a series of elicited (pro-)saccades. Saccadic latency and initial direction are both known to vary in different neurodegenerative diseases, and can potentially be used to evaluate progression of degeneration. Much like the fixation task, the icon would appear and stay at screen center, and five seconds later it would switch to a different location (the same locations as described in the fixation tasks). Five seconds later, the icon would disappear, followed by five more seconds of displaying nothing except the background. This task was repeated twice for five different locations, and the total duration was under five minutes. The first icon always initially appeared at screen center, and subsequent locations of the icon were randomized.

5.1.5 Arithmetic

The fourth task was a set of arithmetic problems increasing in difficulty. This was done to record the pupil as a measure for studying cognitive load. Instead of a green icon, random numbers, as shown on the right of Fig. 2, would appear in the center of the backdrop. Participants were instructed to add the numbers and say the sum out loud (continuing to add new numbers to the previous sum). For example, if the sequence was 1, 5, 9, and 2, the participant would say 1, 6, 15, and 17. The speed at which numbers appeared and the range of numbers to add increased over the course of four trials, with the first trial showing four numbers between 1 and 5 at 5 seconds each, and the last trial showing 10 numbers between 1 and 20 at 2 seconds each. The range of numbers increased by 5 and the speed at which numbers changed increased by 1 second for every subsequent trial. In other words, the rate of change of the numbers increased with every trial, making subsequent trials more difficult.

5.2 Results

From Experiment 1, we had many substantial results, including the replication of several state-of-the-art methods for diagnosing patients, detection of ocular tremor using the Fast Fourier Transform, and generally good subjective responses from patients.

Our first goal was to ensure that it was possible to both evoke and detect a number of ocular abnormalities in our patient group. We targeted three abnormalities in particular, including 1) Ocular tremor (oscillations), 2) square wave jerks, and 3) abnormal smooth pursuit.

5.2.1 Ocular tremor

Ocular tremor, or small amplitude oscillation, has been shown to occur in patients with parkinsonism by Gitchel et al. [7] using a 500 Hz camera. We sought to determine whether fixation tasks in a VR environment could also evoke this tremor, as well as to see if our near-eye camera with a lower frame rate (30 FPS) could also detect such movement. Since this type of tremor has a frequency of between 4 and 7 Hz, we should theoretically be able to detect the oscillation with a less than 20 Hz sampling rate.

For detection, we used a three-step process, including visual inspection of the videos, visual inspection of the resulting eye tracking x, y coordinates, and a fast-Fourier transform (FFT) to measure the difference in amplitude of the 4-7 Hz frequency components in patients vs. controls. This oscillation was visible in the plots of the x, y eye position in screen coordinates and in the results of the FFT, which are both shown for several participants in the left and right columns in Figure 7, respectively.

The y-component of Figure 7 clearly shows several windows of ~6 Hz tremor, which are accompanied by increased microsaccades in the x-component. The overall FFT analysis shown in Figure 8 represents the difference in frequencies in patients vs. controls of all 32-sample windows in all videos for both groups in the x-direction. Clear differences are evident in the 5 and 6 Hz range. The average number of frames for which the 5 or 6 Hz components were stronger were greater in PD (44.01%) than controls (34.14%), though it may be more practical to identify particular regions that are likely to contain tremor for remote visual inspection. Alternatively, FFT could be used to identify and output video segments in which tremor may be present. In practice, it is likely that visual confirmation of plot data will also be necessary for diagnosis.

On the other hand, unlike Gitchel et al.'s work, we did not observe higher fixation instability in the patient group than in controls. One potential reason for this is that the virtual icons were coupled to head movement to help prevent vestibular reflex. In turn, it is possible that fixation stability was actually better since the icons moved with patients' heads. Physicians could re-engage head tracking for the icons if they were more interested in vestibular reflex. Additionally, this shows that ocular tremor can be detected with a frame rate of 30Hz, much lower than Gitchel's 500Hz, though higher sampling may help eliminate false positives.

5.2.2 Abnormal Smooth Pursuit

The smooth pursuit tasks also elicited abnormal movements in PD patients. As the virtual icon moves back and forth, the eye should smoothly follow except for changes direction, where we expect to see one or two corrective saccades followed by smooth movement. Figure 9 shows position and velocity data from a PD patient exhibiting strong saccadic intrusions and a stable control for one full cycle of the virtual icon (left / right at 40 degrees per second). We can clearly see that the PD patient exhibited several high-amplitude abnormal saccadic intrusions. Though controls also sometimes exhibit intrusions, frequency is higher in parkinsonism.

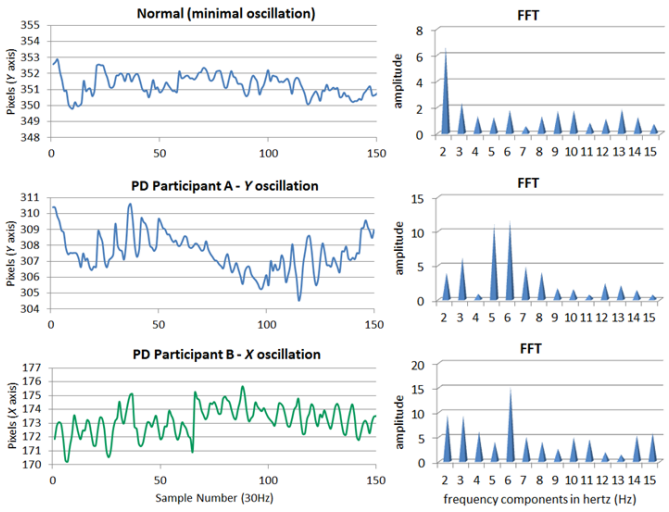


Fig 7. Graphs in the left column show typical (top) and very obvious oscillatory (middle and bottom) fixation behavior. The right column shows the resulting FFT plot for a series of 32 representative samples for that participant. PD patients had consistent oscillation, with amplitudes similar to those found by Gitchel et al. [7].

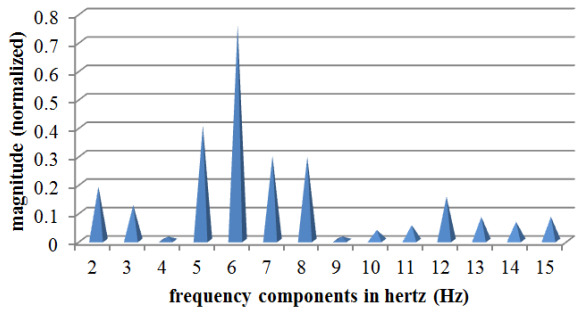


Fig 8. Graph showing average spectral differences in frequency components at 1 Hz intervals for all valid samples of all PD participants vs. controls. Differences are evident at 5 and 6 Hz.

Moreover, horizontal intrusions in vertical smooth pursuit have not yet been observed in parkinsonism to our knowledge, though it was apparent in several videos. It is possible that VR can elicit responses not yet discovered in neurodegeneration.

5.2.3 Square wave jerks

Increased square wave jerks (SWJs), or small, conjugated saccades which take the eyes away from a fixation position and return to the origin, are often a sign of neurodegeneration, especially for diseases like Progressive Supranuclear Palsy and other parkinsonism. During our virtual fixation task, we observed a number of SWJs across a number of patients, several of which were confirmed in follow-up experiments with physicians. As an objective measure, we used an algorithm to count SWJs for each video. For fixation tasks, average SWJ frequency between PD and control was 15.8 and 8.0, respectively. Analysis of variance (ANOVA) confirmed a significant difference between the two groups ($F_{(1,15)} = 4.84$, $P < .05$).

5.2.4 Evoked Pupillary Response

In this experiment we also confirmed that it is possible to generate cognitive evoked pupillary response [2], which can be used to measure cognitive load for the purposes of studying neurodegeneration [20][32]. This is typically difficult when conducting tasks in an office space or outdoors since lighting can severely affect pupil diameter. Though changes to light response occur in Parkinson's patients when compared with controls [32], these changes do not give a good representation of the patient's cognitive workload, which can be revealing of how far the disease

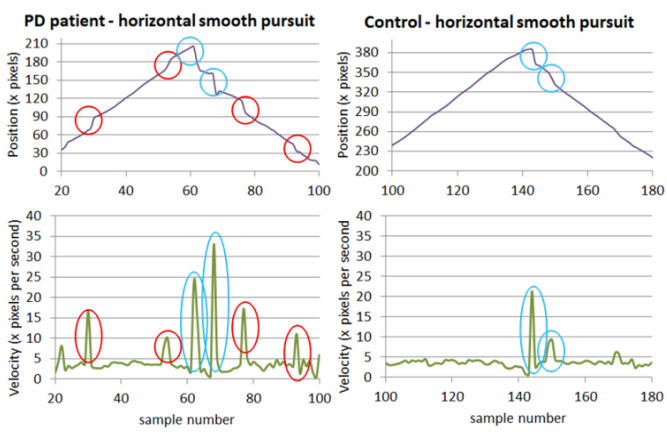


Fig 9. Graphs showing position (top row) and velocity (bottom row) for a horizontal smooth pursuit task. The left column is data taken from a PD patient, and abnormal saccadic intrusions (circled in red) are visible, though pursuit should normally be smooth. The peak position is the point at which the icon changes directions, which is followed by one or two normal corrective saccades (circled in blue). Data from a control participant is on the right for reference.

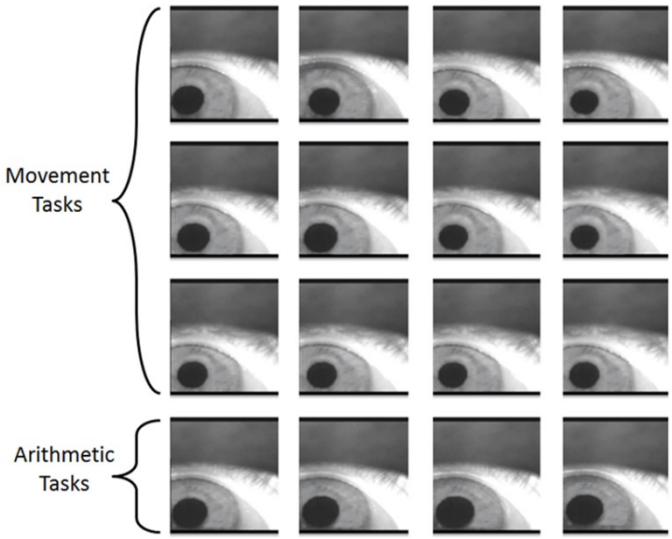


Fig 10. One frame taken from each of 12 movement task videos (top three rows), and from the 4 arithmetic tasks (bottom three rows). This shows that a VR interface can also be used to generate evoked pupil response in a short period of time. *Note: Iris masked for privacy.*

has progressed. Algorithms such as that of Marshall et al. can be used to separate pupil response from lighting changes, but this can take up to ten minutes, time that physicians usually don't have.

Because lighting in VR can be explicitly controlled, our arithmetic tasks were able to generate a pupil response in less than 20 seconds. Example images are shown in Figure 10, where differences in pupil dilation are visible for the arithmetic tasks. Average pupil dilation for all arithmetic tasks was 136.37 vs. 123.9 for static tasks (including saccade), which is shown in Figure 11 for each participant. A paired t-test revealed a significant effect of task on pupil size (t_{stat} = -4.42, P < .001). The short duration with which we can evoke this response will prove invaluable to physicians and researchers who want to study cognition for tasks with a short duration.

5.2.5 Subjective Feedback from Patients

To get a general idea of whether such VR technology would be acceptable in practice, we listened for any comments participants had throughout the experience in addition to monitoring their progress. Out of all experiment participants, only one mentioned a small amount of dizziness after the smooth pursuit tasks. Outside of this incidence, no other discomfort was observed. We also noticed that one participant started to fall asleep during several of the tasks in the middle of the experiment. Sound cues or more engaging visual targets may be able to mitigate this tendency in the future.

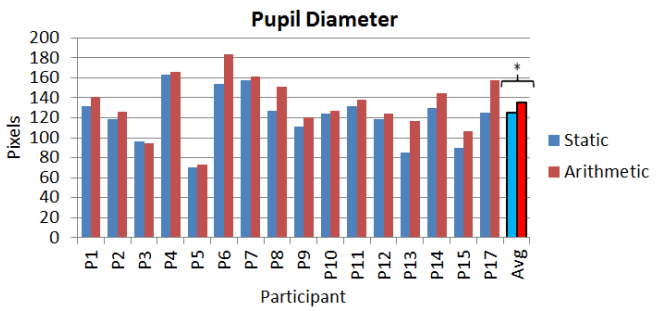


Fig 11. Graph showing average pupil diameter of each participant (with accumulated average at the end) for both the static (blue) and arithmetic (red) tasks. Except for one participant, diameter was higher during arithmetic tasks.

5.3 Clinician Pilot Study: Evaluation of Data and Visualizations

In this experiment, we had two primary goals: 1) to have physicians confirm the presence of abnormalities from the data generated by the VR experiment, and 2) to evaluate if video enhancements would be useful for making a decision and hence valuable to include in follow-up clinical experiments.

5.3.1 Setup

As an initial evaluation of our data, we set up a video-based questionnaire and asked two specialists in movement disorders and eye disease, one male and one female, to participate. Each had more than five years of experience in medical practice. The questionnaire first included a training phase that presented the physicians with eye videos with known abnormalities, including:

Normal:	Abnormal:
• saccade	• square-wave jerk
• microsaccade	• abnormal smooth pursuits
• horizontal smooth pursuit	• ocular tremor
• vertical smooth pursuit	

This training phase lasted approximately 10 minutes, and also provided physicians with videos that showed which virtual tasks were used to elicit which movements.

After the training phase finished, the physicians began the evaluation phase. This included a series of 20 videos, 10 taken from the PD group and 10 from controls. Each set of 10 videos included 5 original videos and 5 videos with visualizations applied (original, pupil overlay, TSUB, and EVM, presented side-by-side), similar to the layout in Figure 6. Video selections from all of the 476 recordings were made with two requirements in mind. The first requirement was that we had videos with a range of clear cut to borderline cases for evaluation. The clear cut cases (such as a normal saccade) would give us a good baseline for agreement since nearly all physicians should be able to identify a basic saccade. Several videos that were suspect of abnormalities, but also very borderline (according to an eye movement specialist that did not participate in the experiment) were also included so that several cases would have a wider range of answers with which we could study variance.

A number of probable, but not clear cut, cases were also chosen, so that we had a good range of data for evaluation. The second requirement was that each video had a "twin" video that exhibited similar movement in the data set of the same patient. This allowed us to randomly select one video from each pair for visualization and one video without, so that we had a relatively fair comparison between conditions. The two physicians then rated each of these videos on a 7 point Likert scale with respect to the probability of presence of a particular abnormality in that video. A rating of 1 corresponded to not likely, whereas a rating of 7 corresponded to likely.

The physicians also rated the confidence of all their answers for that particular video. Orders of all of the videos and multiple choice answers were randomized to prevent ordering effects. A final set of two videos (shown after the 20 evaluations) were presented that asked the physicians to rate each of the visualizations on overall usefulness as well as helpfulness in visualizing particular movement abnormalities. Though subjective, this would also give us an idea of the value of each of the visualizations from a doctor's perspective, and help us determine which were appropriate for a follow-up study.

5.4 Results

5.4.1 Confirmation of Abnormalities

In all of the videos from PD patients with high or medium probability abnormalities, at least one of the two physicians was able to correctly identify each abnormality, as shown in Figure 12. Ocular tremor and abnormal vertical smooth pursuit were correctly identified by both physicians, whereas horizontal smooth pursuit and other abnormalities were identified by at least one. Overall, physician ratings for likelihood met a majority of the time, defined as being within 2 points of each other (i.e., a 6 and 7 would be agreement whereas a 1 and 4 would not). These initial results suggest that physicians can already agree on many of the abnormalities present in the videos. A more in-depth follow-up study is already underway to test agreement among a larger number of physicians.

5.4.2 Visualization Preferences

Though both physicians rated the visualizations as helpful, unfortunately no clear tendency was evident as to which enhancements would be the most useful. All visualizations received either a 6 or 7, so a follow-up study would be necessary to reveal any tendencies. Moreover, physician agreements and confidence levels were almost identical between the visualization and non-visualization conditions, so we were unable to conclude whether a particular visualization would be appropriate for future testing. Although a follow-up study is already underway, it will likely be more important to explore the interface itself, rather than visualizations, to make it easier for physicians to gain real-time video feedback.

5.5 Discussion

One positive outcome of this study is that patients and other elderly individuals were very willing to participate and accepting of VR as a clinical technology. A resistance to, or at least slow acceptance of, new technologies is usually present in the medical field, though we did not find this to be the case in our first experiment. Since physicians have already begun to use smartphones as a means for conducting some types of tests, VR may not be such a far stretch since the nature of the virtual tasks are similar.

One example is the stripe test, which displays moving vertical stripes on a phone screen and requires patients to constantly follow and track new stripes appearing from the top of the screen. Such simulations may also work well in VR, especially since larger virtual objects that more accurately represent 3D physical obstacles can be displayed. Much like the study presented by Marx et al. [19], we would also like to test whether an "in-the-wild" test is possible with AR. If successful, this type of test for neurodegeneration may become as commonplace or easy as the Ishihara color blindness tests, which are already available as smartphone and web applications. Still, our interface is just as well suited to the study of abnormal eye movements as it is to diagnosis.

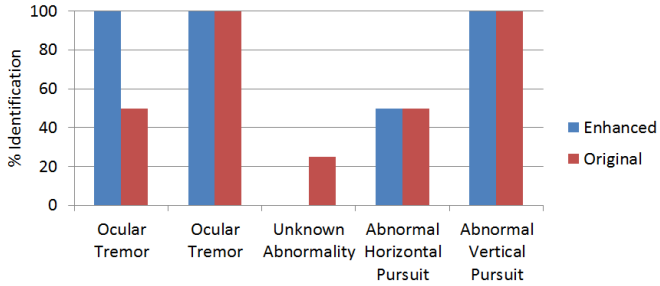


Fig 12. Graph showing the percentage of identifications for each abnormality category. No clear tendencies were visible for the videos with enhancements (visualizations).

The combination of visualizations with VR adds yet another way for physicians as well as researchers to make more effective observations about the eye. Finding horizontal saccadic intrusions in vertical smooth pursuit tasks is a good example of how we may be able to more easily find new movements or biomarkers for diseases. To our knowledge, a correlation between ocular tremor and microsaccades has not yet been observed in parkinsonism, and is currently under further investigation as future work.

Although our results were mostly positive, we also learned much about what needs to be improved. Camera placement and consistent lighting will likely be key for more accurate, automatic processing of results. Several patients did not initially wear the display correctly, eye images sometimes went out of focus, and excessive mascara interfered with eye tracking for some participants. This rendered several of our video samples unusable, which is also likely a source of error for the clinical experiment. An easily adjustable camera, software adjustable focus, and automated calibration [26] will be invaluable to future iterations of such technology. We also learned that showing participants their own eye during the introduction screen was quite helpful. Participants got an idea of how display adjustment affected eye position and were interested to see their eye in IR light.

Because the Rift's lenses are interchangeable, the infrared camera-lens system we used can easily be swapped with other displays or shipped to patients' homes as a small, standalone unit. This is important from a cost-savings perspective since eye-tracked VR is still expensive. This infrastructure may also provide another step towards inexpensive or personal performance monitoring for disabled or impaired individuals, much like systems currently used for training and simulation [16]. In comparison with other studies, our interface can facilitate diagnosis with conventional eye tracking cameras, further lowering the cost to physicians, and consequently to patients. One thing to be aware of is that accommodation mismatch is still a problem in many commercial VR displays [14]. Though we still had significant detection results from our experiment, this drawback should be taken into account by future researchers and physicians when developing more complex 3D tasks for diagnosis.

One next step we would like to explore as future work is the introduction of depth-based tasks. Since we largely have control of vergence in VR, it will be interesting to study whether any abnormalities, and possibly new biomarkers, exist for points moving at varying depths from the user. Following an object traveling in spherical or random 3D patterns may also be a viable way to test oculomotor performance. Moreover, based on the success of our evoked pupil response tasks, we also intend to study the differences in cognitive workload in PD, and explore this as a means for measuring disease progression.

6 Conclusion

In this paper, we introduce a VR system used to evaluate eye movements associated with neurodegenerative conditions through a combination of physician task emulation, eye tracking, and video enhancement. After testing a prototype device with patients and controls in a clinical setting, we found that participants were able to complete tasks just as they would with their regular physician. We were then able to replicate a number of current methods for diagnosis of oculomotor abnormalities, which were subsequently confirmed in a follow-up study with practicing physicians.

The results presented here show that we can use virtual tasks to evoke abnormalities associated with neurodegenerative disease, which represents a significant step towards making eye tracked VR practical for use in clinical settings. Emulating tasks commonly used by clinicians can be an important stepping stone towards making remote diagnosis a reality, especially considering patients' acceptance of VR in our clinical experiment.

We also hope that this system and the resulting experiments will inspire others to build more effective low-cost virtual interfaces for diagnosing neurological conditions.

ACKNOWLEDGEMENTS

We would like to thank all of the volunteers that contributed their time and feedback during the experiment. We would also like to thank the continued collaborative efforts of Osaka University, Augusta University, and Kansas University Medical center. This work was supported in part by the Japan Society for the Promotion of Science, Grant #A15J030230.

REFERENCES

[1] Armstrong, R. A. (2015). Oculo-Visual Dysfunction in Parkinson's Disease. *Journal of Parkinson's disease, 5(4)*, 715-726.

[2] Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin, 91(2)*, 276.

[3] Bradski, G., & Kaehler, A. (2008). Learning OpenCV: Computer vision with the OpenCV library. " O'Reilly Media, Inc.".

[4] Cushman, L. A., Stein, K., & Duffy, C. J. (2008). Detecting navigational deficits in cognitive aging and Alzheimer disease using virtual reality. *Neurology, 71(12)*, 888-895.

[5] Dietz, J., Bradley, M. M., Okun, M. S., & Bowers, D. (2011). *Emotion and ocular responses in Parkinson's disease. Neuropsychologia, 49(12),* 3247-3253.

[6] Fotiou, D. F., Stergiou, V., Tsiptsios, D., Lithari, C., Nakou, M., & Karlovasitou, A. (2009). Cholinergic deficiency in Alzheimer's and Parkinson's disease: Evaluation with pupillometry. *International Journal of Psychophysiology, 73(2),* 143-149.

[7] Gitchel, G. T., Wetzel, P. A., & Baron, M. S. (2012). Pervasive ocular tremor in patients with Parkinson disease. *Archives of neurology, 69(8),* 1011-1017.

[8] Gitchel, G. T., Wetzel, P. A., & Baron, M. S. (2013). Slowed saccades and increased square wave jerks in essential tremor. *Tremor and Other Hyperkinetic Movements, 3*.

[9] Gupta, V., Knott, B. A., Kodgi, S., & Lathan, C. E. (2000). Using the "VREye" system for the assessment of unilateral visual neglect: two case reports. Presence: *Teleoperators and Virtual Environments, 9(3)*, 268-286.

[10] Granholm, E., Morris, S., Galasko, D., Shults, C., Rogers, E., & Vukov, B. (2003). Tropicamide effects on pupil size and pupillary light reflexes in Alzheimer's and Parkinson's disease. *International Journal of Psychophysiology, 47(2)*, 95-115.

[11] Hacisalihzade, S. S., Kuster, F., & Albani, C. (1986). Computer-aided measuring of motor functions using pursuit tracking. *Computer methods and programs in biomedicine, 23(1),* 19-28.

[12] Kassner, M., Patera, W., & Bulling, A. (2014). Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. *In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, 1151-1160.

[13] Kon Graversen V, Jani P. Clinically useful and cost effective mobile retinal screening (2014). *In New Retina MD*.

[14] Kramida, G. (2016). Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays. *IEEE transactions on visualization and computer graphics, 22(7),* 1912-1931.

[15] Leigh, R. J., & Zee, D. S. (2015). The neurology of eye movements (Vol. 90). Oxford University Press, USA.

[16] Liston, D. B., & Stone, L. S. (2014). Oculometric assessment of dynamic visual processing. *Journal of vision, 14(14)*, 12-12.

[17] Liu, C., Torralba, A., Freeman, W. T., Durand, F., & Adelson, E. H. (2005). *Motion magnification. ACM transactions on graphics, 24(3)*, 519-526.

[18] Lowenstein, O. (1955). Pupillary reflex shapes and topical clinical diagnosis. *Neurology, 5(9)*, 631-631.

[19] Marx, S., Respondek, G., Stamelou, M., Dowiasch, S., Stoll, J., Bremmer, F., & Einhäuser, W. (2012). Validation of mobile eye-tracking as novel and efficient means for differentiating progressive supranuclear palsy from Parkinson's disease. *Frontiers in behavioral neuroscience, 6*.

[20] Marshall, S. P. (2007). Identifying cognitive state from eye metrics. *Aviation, space, and environmental medicine, 78(5),* B165-B175.

[21] Orlosky, J., Toyama, T., Sonntag, D., & Kiyokawa, K. (2014, June). Using eye-gaze and visualization to augment memory. In the *International Conference on Distributed, Ambient, and Pervasive Interactions*, 282-291.

[22] Pamplona, V. F., Mohan, A., Oliveira, M. M., & Raskar, R. (2010, July). NETRA: interactive display for estimating refractive errors and focal range. *In ACM Transactions on Graphics. 29(4)*, 77.

[23] Pan. W. QtEVM source code. <https://github.com/wzpan/QtEVM>, Accessed September 9th, 2016.

[24] Pinkhardt, E. H., Kassubek, J., Süssmuth, S., Ludolph, A. C., Becker, W., & Jürgens, R. (2009). Comparison of smooth pursuit eye movement deficits in multiple system atrophy and Parkinson's disease. *Journal of neurology, 256(9),* 1438-1446.

[25] Plancher, G., Tirard, A., Gyselinck, V., Nicolas, S., & Piolino, P. (2012). Using virtual reality to characterize episodic memory profiles in amnestic mild cognitive impairment and Alzheimer's disease: Influence of active and passive encoding. *Neuropsychologia, 50(5)*, 592-602.

[26] Plopski, A., Orlosky, J., Itoh, Y., Nitschke, C., Kiyokawa, K., & Klinker, G. (2016, September). Automated Spatial Calibration of HMD Systems with Unconstrained Eye-cameras. In *Mixed and Augmented Reality (ISMAR), 2016 IEEE International Symposium on*. 94-99. IEEE.

[27] Rizzo, A., Buckwalter, J. G., van der Zaag, C., Neumann, U., Thiébaux, M., Chua, C., ... & Larson, P. (2000). Virtual environment applications in clinical neuropsychology. *In Virtual Reality, 2000. Proceedings. IEEE, 63-70.* IEEE.

[28] Rolland, J. P., & Fuchs, H. (2000). Optical versus video see-through head-mounted displays in medical visualization. *Presence, 9(3),* 287-309.

[29] Sárkány, A., Tősér, Z., Verő, A. L., Lőrincz, A., Toyama, T., Toosi, E. N., & Sonntag, D. (2015). Maintain and Improve Mental Health by Smart Virtual Reality Serious Games. *In Pervasive Computing Paradigms for Mental Health*, 220-229.

[30] Schultheis, M. T., Himelstein, J., & Rizzo, A. A. (2002). Virtual reality and neuropsychology: upgrading the current tools. *The Journal of head trauma rehabilitation, 17(5),* 378-394.

[31] Sonntag, D., Orlosky, J., Weber, M., Gu, Y., Sosnovsky, S., Toyama, T., & Toosi, E. N. (2015, June). Cognitive Monitoring via Eye Tracking in Virtual Reality Pedestrian Environments. *Proceedings of the 4th International Symposium on Pervasive Displays,* pp. 269-270.

[32] Stergiou, V., Fotiou, D., Tsiptsios, D., Haidich, B., Nakou, M., Giantselidis, C., & Karlovasitou, A. (2009). Pupillometric findings in patients with Parkinson's disease and cognitive disorder. *International Journal of Psychophysiology,72(2),* 97-101.

[33] Świrski, L., Bulling, A., & Dodgson, N. (2012). Robust real-time pupil tracking in highly off-axis images. *Proceedings of Symposium on Eye Tracking Research and Applications,* 173-176.

[34] Tanner, C. M., & Goldman, S. M. (1996). Epidemiology of Parkinson's disease. Neurologic clinics, 14(2), 317-335. Chicago.

[35] Trujillo, D., Martínez, F., & Romero, E. (2015, January). Characterizing the eye trajectory during the gait towards Parkinson stage identification. *In the Tenth International Symposium on Medical Information Processing and Analysis*, 92871A-92871A.

[36] Van Gerven, P. W., Paas, F., Van Merriënboer, J. J., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology, 41(2),* 167-174.

[37] Weil, R. S., Schrag, A. E., Warren, J. D., Crutch, S. J., Lees, A. J., & Morris, H. R. (2016). Visual dysfunction in Parkinson's disease. *Brain, aww175*.

[38] Wu, H. Y., Rubinstein, M., Shih, E., Guttag, J. V., Durand, F., & Freeman, W. T. (2012). Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics 31(4)*. 1–8.