Nystagmus

Simulating and Compensating for Eye Motion Blur with Eye Tribe

Joe Landry
Stanford University
Department of Electrical Engineering
350 Serra Mall
Stanford, CA
jlandry@stanford.edu

Philip Lee
Stanford University
Department of Electrical Engineering
350 Serra Mall
Stanford, CA
philee@stanford.edu

Maneeshika Madduri
Stanford University
Department of Electrical Engineering
350 Serra Mall
Stanford, CA
mmadduri@stanford.edu

I. INTRODUCTION

Nystagmus is a periodic, involuntary eye condition. It is characterized by the smooth pursuit of the eye away from the target location followed by a quick saccade back to the target location. Nystagmus is also called "dancing eyes" as the eye movements appear pendular or jerky to an observer. Everyone has nystagmus. Trying to fixate on objects far in your periphery will produce nystagmus-like movements, as will some drugs. Sobriety tests often examine the gaze of one's eyes, looking for alcohol-induced nystagmus.

Nystagmus is either physiological or pathological [1]. The former case involves gaze-evoked nystagmus or optokinetic nystagmus (which is induced in the observer). The latter includes both acquired and congenital nystagmus. Interestingly, those with congenital nystagmus generally adapt to the condition and, depending on the severity of the nystagmus, the vision may not be affected greatly. However, a person with acquired nystagmus can suffer from oscillopsia, which is the apparent feeling that the world is spinning; this consequence of nystagmus can lead to further complications, such as vertigo [2]. It is thought that 1 in 1,000-2,000 people are affected by congenital and acquired nystagmus [3].

Those with nystagmus generally develop compensatory strategies to reduce the blurring and other consequences of their eye motions. These patients commonly tilt their head in an orientation that puts their eyes in a position that minimizes the effect of their nystagmus [1]. Some patient's nystagmus disappears altogether in specific positions, but for most people (especially those with acquired nystagmus), the consequences persist. There is no definitive cure for nystagmus. It is possible to treat it with surgery, but results vary from patient to patient. In many situations, people with this condition could benefit greatly from enhanced visual acuity.

A temporary solution involving head-mounted displays (HMDs) and eye tracking for sensitive situations such as reading or writing may be possible. By actively monitoring the gaze of the eyes in real time, a head mounted display could react by shifting the displayed image of the real world to counteract the effects of nystagmus. This would result in an image that appears stabilized to the observer, thus increasing the individual's visual acuity.

A necessary requirement in developing such a visual compensation system is access to people exhibiting nystagmus. Fortunately, it is possible to induce optokinetic nystagmus in a person with normal vision using various graphics on displays [4]. In one such experiment, a 1D spatially varying sinusoidal pattern is projected onto a screen opposite an observer (see Fig. 1). The pattern moves either left tor right at a constant angular velocity, so a certain number of different colors / shades pass a fixed point on the screen in a given amount of time. The observer is then asked to count the number of times a particular color passes the center of the screen. The motion of the eye that results is the same motion characteristic of nystagmus. The eye follows a smooth pursuit as it follows the peak of the sinusoid across the wall, and then saccades to the next peak after it passes through the center of the screen. In this way it is possible to generate the eye movements to be captured by our system and used for simulation and compensation of nystagmus.

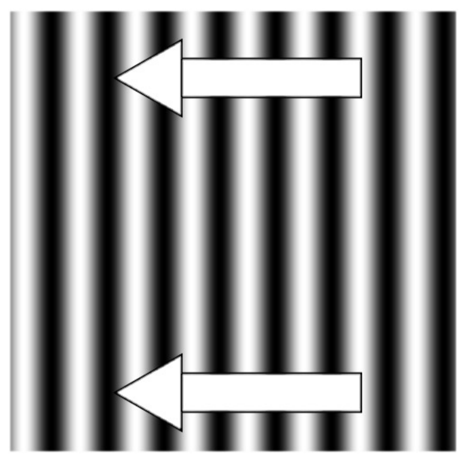

Figure 1: The sinusoidal pattern used to induce optokinetic nystagmus. The arrows indicate the direction the pattern moves. In this case, observers would be asked to count the number of black or white lines that move past.

Finally, it is important to examine the technical limitations of using HMDs for nystagmus correction. Capture/display frame rates, system latency (from the eye tracker to the HMD), and integrated HMD eye-tracking devices are important considerations. We experimentally examine the feasibility of using Eye Tribe, a 30 Hz eye tracker, against gold standard devices. In our simulations we use latency and display frames per second (FPS) as variables to better understand what the longest latency and minimum FPS are needed to stabilize the image with minimal error. We also look at another subtle effect, which is motion blur. Even if we can update the frames in the right location at each moment in time, an eye moving very quickly will still motion blur the image. For that reason, we will briefly introduce the incorporation of wiener deconvolution into our system as a predistortion mechanism.

II. Related Works

Our research methods were heavily influenced by the work of Iijima [4]. In their paper, they compare the accuracy of three different eye tracking methods for measuring the amplitude and velocity of the quick phase of nystagmus. The velocity of the quick phase of nystagmus is a key metric used for diagnosing central brain or brain stem disorders but it may not be as critical for correcting motion blur caused by clinical, acute nystagmus. Each system is tested using an optical stimulating system for evoking optokinetic nystagmus (OKN) on normal subjects. Here we summarize their methods and findings.

The clinical standards for measuring eye movements are electronystagmography (ENG) and video-oculography (VOG). ENG is the measurement of the change in the electric potential between the cornea and the retina as a result of eye movement. The system is made up of several electrodes attached to the face around the eyes. It has the disadvantage of requiring calibration and it does not record torsional or vertical nystagmus because these signals contain electrical artifacts, which confound the analysis. However, it does accurately record the quick phase of nystagmus. VOG uses a CCD camera and a computer vision algorithm to quantify eye movements. It cannot record the quick phase of nystagmus as it only samples at 30 frames per second. However, it can analyze all types of nystagmus. In this paper, the authors use a VOG system equipped with a high-speed camera to sample at 250 frames per second to accurately record the quick phase of nystagmus.

To induce OKN, a standard clinical optical stimulating system is employed. The system consists of a stimulator and a screen. The stimulator is a cylindrical rotating drum equipped with a light source and covered with a pattern of alternating

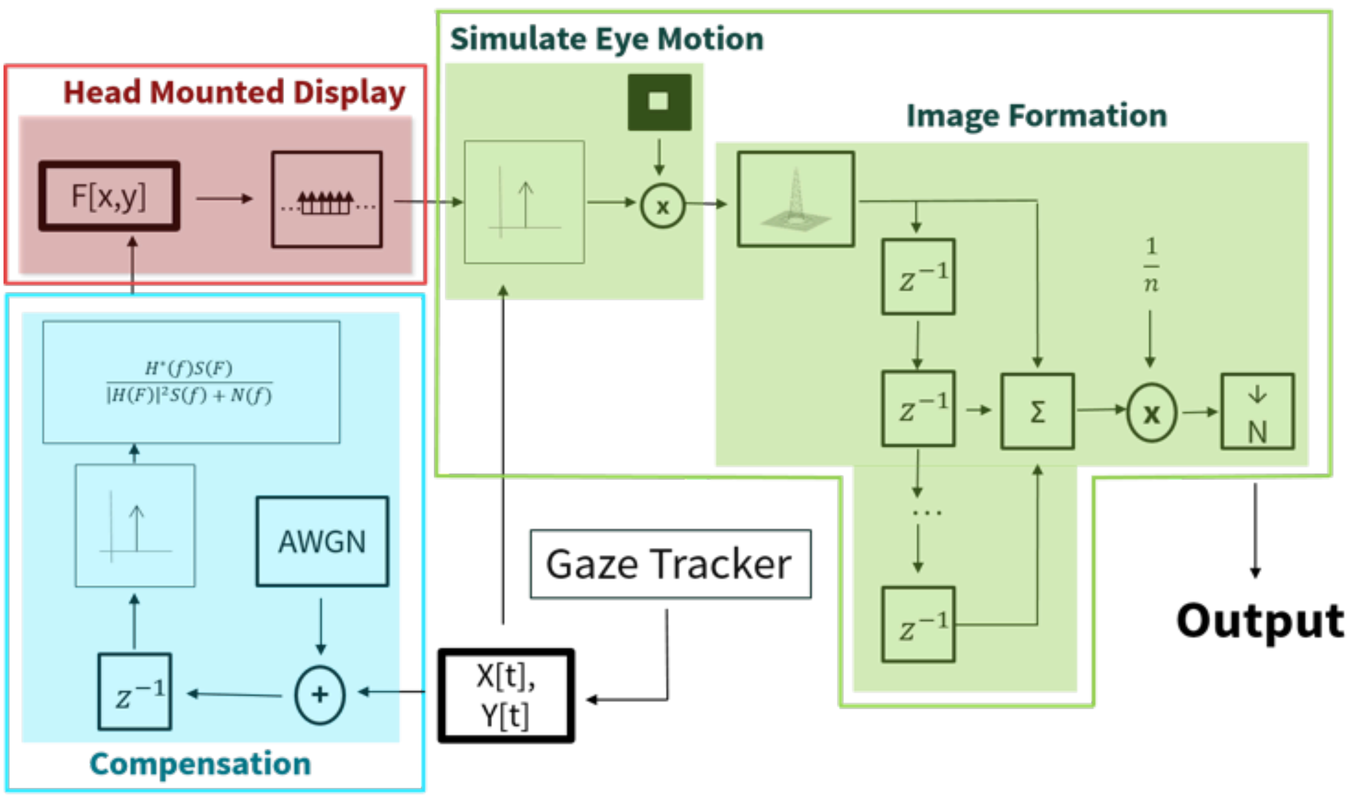

Figure 2: Linear-Time and Shift-invariant System describing the operation of our compensation methods

stripes. The stimulator rotates at a specified speed and projects a moving striped pattern onto the screen. The subject is instructed to watch the moving stripe pattern projected onto the screen in front of them while the recording equipment logs their eye movements.

To translate the high speed video images of the movements into gaze coordinates, the authors first converted the images into binary form. For each pixel on the image, they recorded the coordinate and the grey level. Then, they created a grey-scale histogram from these pixels and used it to divide the black and white sections of the image based on a threshold. With this information, they were able to calculate the coordinate of the center of the pupil. Finally, they combined all of the images and integrated the coordinates to calculate the velocity.

The results of their recordings show that the high speed camera at 250 Hz is more accurate for detecting the quick phase of nystagmus when compared against the clinical standard VOG at 30 Hz. They also show that their method is equivalent to the clinical ENG system for horizontal nystagmus and better than the clinical ENG system for vertical nystagmus. All of their comparisons are made using the least-squares correlation coefficient.

III. Approach

Our project can be divided into two distinct parts. We first used the Eyetribe gaze tracker to measure gaze coordinates in a subject with induced Nyastagmus and compared these results to a published paper that measured gaze coordinates in a controlled laboratory setting. The second part of our project is the Nyastagmus simulation and compensation system that we built in Matlab. The ultimate goal of our work is to allow people afflicted with Nyastagmus to see a stable image on a head mounted display using a gaze tracker and clever computation. We modeled this situation in Matlab as a first step towards achieving this. We used the measured coordinates from the first part of our project as inputs into the second part of the project, the simulation system. We discuss each of these two parts in the sections below.

A. Gaze Tracking

We used the data from Iijima [4], shown in Figure 3, as a reference both for creating our custom optical stimulating system and for comparing our results to the gold standards.

To take measurements using the Eyetribe gaze tracker, we developed a test to stimulate nystagmus based on the typical amplitude and velocity measures that were observed from the paper in [4]. The setup consisted of a computer monitor (1280x1024), a powerpoint presentation with our videos to be played, and a laptop running the Eyetribe API and connected to the Eyetribe gaze tracker, which was positioned underneath the monitor. The subject sat at a desk in a chair approximately 60 cm away from the Eyetribe gaze tracker. To stimulate optokinetic nystagmus, we created a video with a moving dot that travelled 15 degrees to the right at a velocity of 60 degrees/second and then immediately snapped back to the starting position where it remained stationary for 0.6 seconds before repeating the cycle. The simulated nystagmus eye-movement tracked by the Eyetribe is shown in Figure 4. We

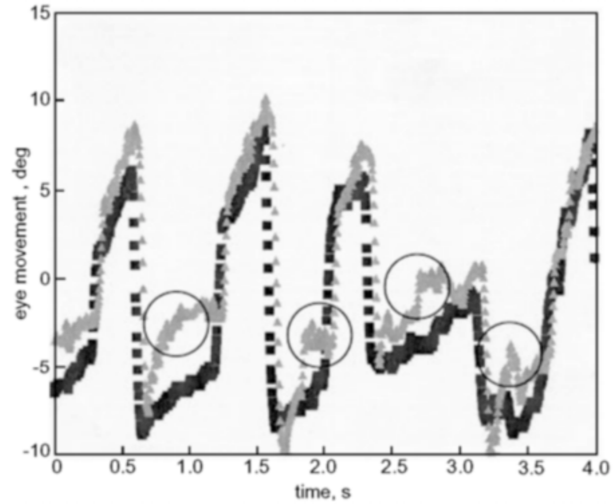

Figure 3. Horizontal Opto-kinetic Nystagmus generated in Iijima [4]

used this setup for stimulating horizontal nystagmus because we did not have the standard clinical stimulating equipment. We also hypothesized that constraining the observer to two points of focus would give us more repeatable oscillatory movements than by having the observer follow a series of striped patterns.

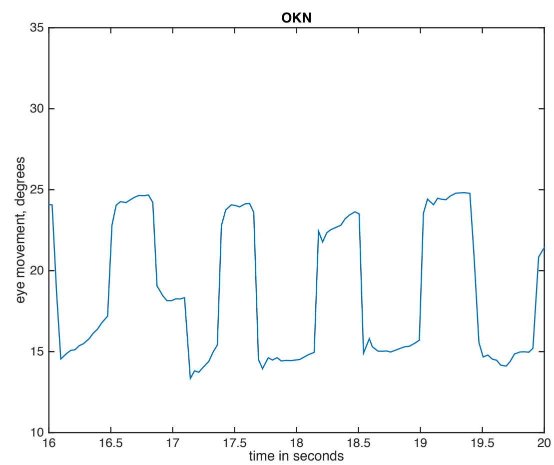

Figure 4. Project simulation of Opto-kinetic Nystagmus using EyeTribe for tracking the eye-movement

Our results show good qualitative agreement both in temporal frequency and in amplitude. We note that our test conditions and methods were not identical or to those in the paper. Therefore, we can only make qualitative comparisons between our results. A more thorough test would include validation of our method using the Eye Tribe against the gold standard clinical VOG system at 250 Hz.

B. Simulation Systems

The simulation system we built serves two purposes. First, we can use this simulation to convey how someone with nystagmus may experience the world. Second, we needed a way to test and measure our proposed compensation system. We designed the simulation system as a discrete time, linear and shift invariant system in Matlab to make it computationally feasible. Figure 2 above shows a high level block diagram and the mathematical operators of each step. The system can be grouped into three major subsystems Head-Mounted Display Input/Output (I/O) Devices, Eye Simulation and Image Formation, and Compensation System. These subsections roughly correspond to the groupings in Figure 2. The following sections will describe general details of the system and then delve deeply into each subsystem.

C. General Overview

To make the simulation computationally tractable, we modeled the devices, nystagmus, and our compensation as a discrete, linear, and shift-invariant system. We sampled the continuous system every millisecond (i.e 1000 Hz) to make our discrete approximation for the real world. Additionally, we tried to make our software design as modular as possible so that we could easily test and evaluate different specifications such as real world sampling rate, display refresh rate, gaze tracker accuracy, etc.

D. I/O Devices

There are two major components to the I/O devices. The first is the head mounted display. We use this to display a compensated image. The display has a constant framerate and can only update the image accordingly. We experimented with different framerates to see the effects on our system, but defaulted to using a framerate of 125 hz in our simulation. Mathematically, this is sampling the desired image in time (or convolving it with a Kroeneker delta train). The second input is the gaze tracker. We simulated this by pre-recording gaze coordinates of an individual with induced nystagmus as described in the Gaze Tracking section above. We interpolated these measurements to get gaze at each discrete time sample and scaled the units to account for visual angle subtended rather than image pixels. This subsystem simulates the images displayed by the head mounted display and provides gaze coordinates for the other two subsystems.

E. Eye Simulation

The next portion of our simulation was the eye motion/optic simulation. Again, we made some simplifying assumptions to make our system computationally feasible. First, we assumed that the scene of the image we wanted to show had little depth. We chose images to accommodate this either by using landscape images such that objects in view were very far away, images of flat objects such as posters, or images with small fields of view such that there was little depth present in the view. Figure 5 below is an example of this. We also assumed that the eye motion involved was a small angle to avoid worrying about spherical projections.

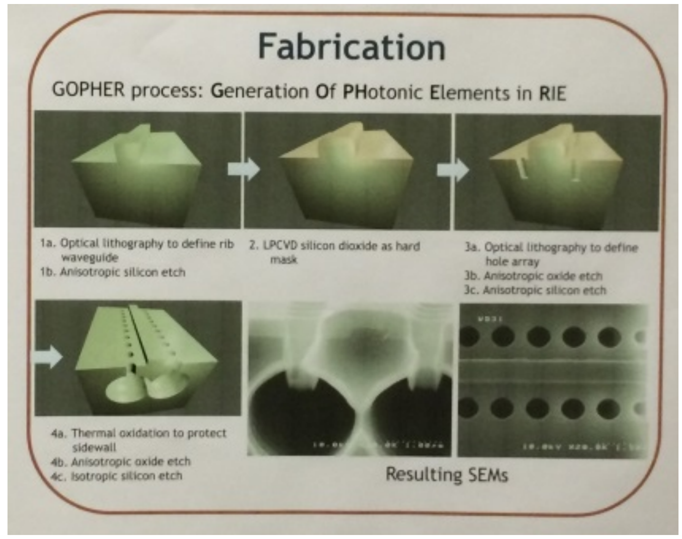

Figure 5. Example of a scene with little depth

The combination of these two assumptions allowed us to simulate small motions of the eye by small shifts in the image. In the real world, eye motion means a small rotation about the center of the eye. The image you see is a projection of the real world radially onto your retina. Items at different depths and viewing angles would be projected differently into the image you see. However, with our two assumptions, we can ignore these effects from the scene and manipulate the digital image as an approximation of eye motion. If we crop the center of the image as our field of view, we can approximate eye motion as shifting the image spatially before cropping. This is equivalent to convolving our image with a spatially shifted Kroeneker delta and then masking only the relevant pixels in our field of view. In our computational system, we arrive at the same result by shifting the coordinates of our field of view before cropping.

We then took the optics of the eye into account by utilizing the Image Systems Engineering Toolbox Biology Module (ISETBIO)[5]. The software created our image as a scene and computed the effect of the point-spread function of the pupil (airy disk). Finally, it also simulated the cone absorption under standard lighting to get to the final result. Figure A below shows an example of the output of our image after running it through ISETBIO.

The last portion of our eye simulation was image formation. We took the assumption that the eye samples images at 30 frames per second. The eye does not have a fixed sampling rate, but this is a good approximation for our purposes. Depending on the illumination conditions, the eye needs a stimulus at 16-60 Hz to appear stable [6]. Thus, we average groups of frames that span (1/30) seconds to approximate what the human eye sees. This is an analog to integrating over the exposure time of a camera sensor and is equivalent to low pass filtering and down sampling our frames in the time domain. This completes our simulation of the view of someone with active nystagmus when given their gaze coordinates.

F. Compensation

Our final subsystem is the compensation feedback system we propose. Our idea has two parts: image stabilization and pre-distortion. Each of these two subsystems is described below.

a) Image Stabilization: The first order compensation we propose is image stabilization. We track the user's gaze and shift the image displayed to match motion in the user's gaze. That is, when a user looks away from the center of the display, we shift the image so that the center of the image is matched to the center of the user's gaze.

In the interest of creating a realistic model, we also take into account as many non-idealities of the system as possible. The coordinates of the gaze tracker will have some error. The specifications of the gaze tracker we used suggest an error of about 0.5 cm, which was approximately 0.5 degrees of visual angle in our experiments. In our simulation, we took the gaze tracker coordinates to be the ground truth and simulated the noise in measurements with additive white Gaussian noise (AWGN) with standard deviation of 0.5 degrees of visual angle. Additionally, in a real system, we know there would be some latency between gaze coordinate measurements, computation, and stabilization. We simulate this by delaying the gaze coordinates of the system. This is one of the variables we tuned as we investigated the performance of our system. Results are discussed in later sections.

b) Image Pre-distortion: We also pre-distorted the image to accommodate continuous eye motion. Because we have a display that has a finite refresh rate, being viewed by a user with continuous gaze motion, we know that the final image the user sees will have some motion blur (due to eye motion on a static image). We can predict this and pre-distort the image so that when motion blur is applied by eye motion, the final image looks sharper. A standard (and computationally efficient) method of reducing blur is using a Wiener Filter. This filter is designed to reduce blur in the presence of noise (which is caused by imperfect measurements, delay, and stabilization) when an estimate of the noise is known beforehand. In our case, because the blur is applied by the motion of the eye, we are pre-distorting the image in anticipation (rather than in response to) blur. Results are shown in a later section.

IV. Evaluation

Our system combines an input image with a compensation feedback loop to produce a stabilized image for the viewer. Considering each image discretely, there is an original input image, an uncompensated image that a viewer with horizontal nystagmus would see and a compensated, stabilized image that our system outputs. A screenshot of each of these images is shown below in Figure 6.

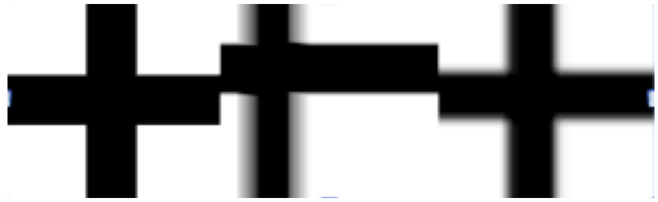

Figure 6: From left to right, this shows (1) an example of a static image, (2) an image that a viewer with nystagmus would see, (3) the compensated image that our system

The evaluation of our system was divided into three categories: comparison of pixel distance error between the stabilized image and non-stabilized image, RMS Error of the static image and the compensated image and acutance of the stabilized image. The first metric compared how effective our system was to a viewer with nystagmus, and the last two quantified and evaluated our compensation algorithm.

A. Comparison of Pixel Distance Error

The effectiveness of our compensation algorithm was considered by looking at the distance error as a function of time, as shown in Figure 7. The distance error is measured in pixels, and is a measurement of the distance between the gaze coordinates (where the observer is looking) and the ideal coordinates of the shifted image with respect to the image center. The latter set of coordinates is equal to the gaze coordinates when the latency of the system is zero, so in an ideal system the distance error of the stabilized image would be zero. By introducing some amount of expected system latency, we can get more realistic results.

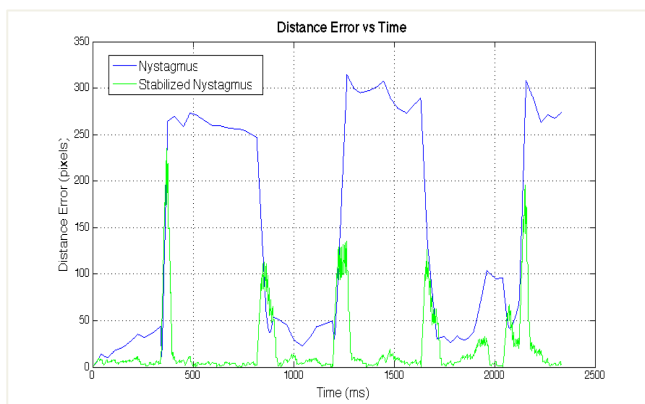

Figure 7: The distance error (measured in pixels) over times (ms) between the image seen by a viewer with nystagmus (shown as the blue line) and the stabilized version of that image (shown as the green line).

The blue line in the plot effectively shows the movement of the eye in time for an uncompensated system because the distance error simply represents the distance of the gaze from the unmoving image center. The green shows the error in our compensated system. It is evident that when the eye is relatively still our results are receiving; the distance error is always less than about 20 pixels. However, it is also apparent that during fast eye motions the compensation cannot keep up. As soon as the eye moves quickly, the distance error rises considerably for a brief time (~50-100ms) before the system can catch up. Incorporating a better predictive mechanism or a system with lower latency would mitigate the effects of fast eye movement.

B. RMS Error of Compensated Images to the Static Image

This metric evaluated the error associated with our compensation system and considered frame refresh rate and latency of the entire system. The RMS error was calculated by comparing the cropped and spatially shifted original image, shown as the first image in Figure 6, to the motion-compensated image, shown as the third image in Figure 6. The aim of this evaluation data was to quantify the error between our system and an idealized system in which a viewer with nystagmus and a viewer without nystagmus would see no

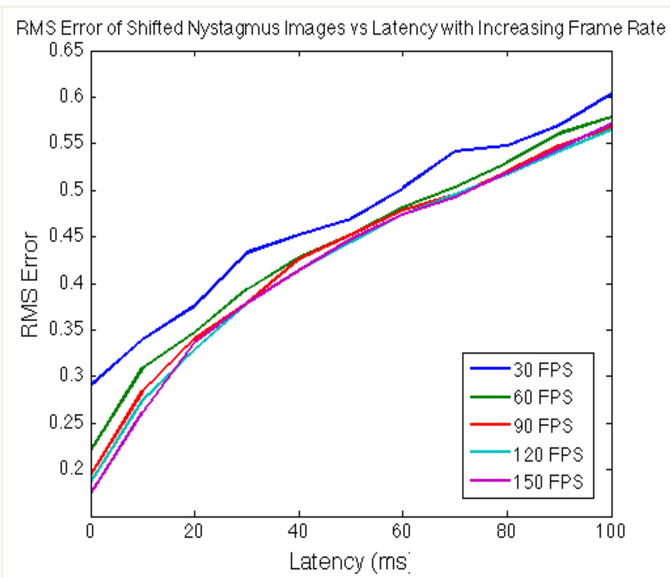

Figure 8: RMS Error of stabilized images and the original image

difference.

Additionally, this metric allowed us to consider the effects of potential latency between the eye-tracker and a headset and the frame refresh rate of the display. As seen in Figure 8, the RMS error increases monotonically with an increasing latency, which is as expected. A latency close to 0 milliseconds would produce the lowest error. The trend with the frame refresh rate is less clear. As the frame rate increases, the error decreases, which is expected. However, the error does not consistently decrease with increasing frame rates. A rate over 30 FPS (frames per second) results in a lower RMS error, but not noticeable improvements between 60-150 FPS.

C. Acutance Measurements of the Images

The last part of evaluating our system dealt with the acutance of the stabilized image with varying frame rate. We considered three different measures of acutance of the compensated image: the mean value of a Laplacian of Gaussian filtered image, the maximum value of a Laplacian of Gaussian filtered image and mean value of a Sobel filtered image [7]. Using the average acutance of the discrete, stabilized images, the acutance was compared with increasing frame rates of the display. The resulting plots are shown below in Figures 9,10, 11. This metric did not provide us with definitive information about the sharpness of the motion-compensated images. The mean value of gradient of the Sobel filtered image increased with increasing frame rates, which is

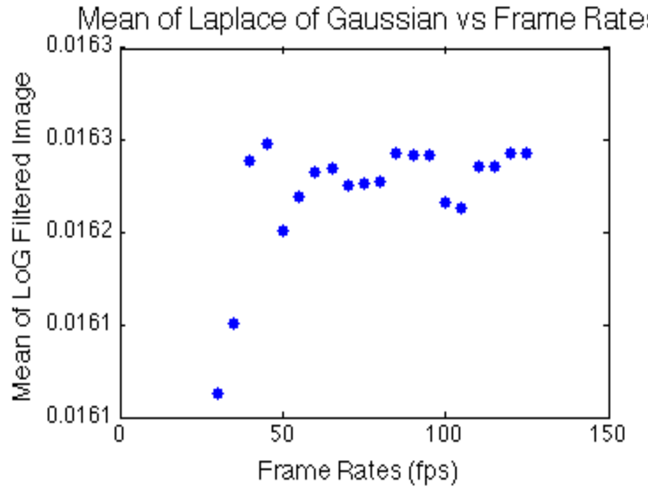

Figure 9: Acutance versus Frame Rate measured by the mean value of the array of images convolved with a Laplacian of Gaussian filter.

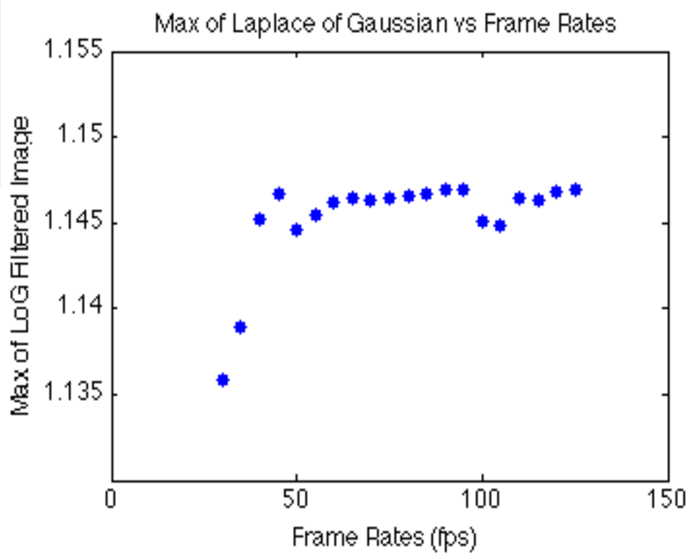

Figure 10: Acutance vs Frame Rate measured by the maximum value of the array of images convolved with a Laplacian of Gaussian filter.

as expected, showing that a faster display frame rate would provide a sharper image. However, neither of the Laplacian of Gaussian filtered images showed a clear correlation to frame rates.

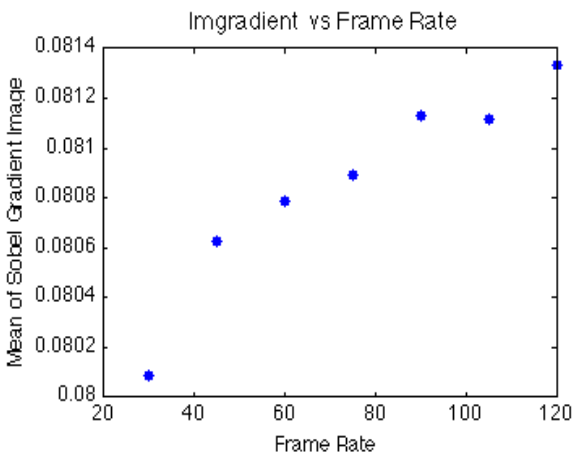

Figure 11: Acutance vs Frame Rate measured by the mean value of the array of images convolved with a Sobel kernel

V. Discussion

The major purposes of our project was to create a simulation system for nystagmus, introduce a compensation algorithm to our simulation, and measure the factor of different variables on our system. We were successful in all three aspects. Our simulation system provides a depiction of the viewpoint of someone afflicted with nystagmus. We were able to integrate compensation system with our simulation and benchmark its performance in various ways, as described in the Evaluation section. Finally, we designed our system in such a way as to be able to tune various parameters and see their effects on our simulation. Videos are available with our code for example outputs of our system. Error metrics and results have been described in previous section.

Of particular note though, is our system's performance as a function of two variables a real system would have some control over in design. These two variables are refresh rate of the screen and latency of our entire system. The RMS error as both of these variables are changed is displayed in figure X above in the evaluation system. We see that although performance seems to increase as refresh rate of the head mounted display increases, the increase is marginal after 60 fps. This makes sense, as our assumption is that the eye samples at 30 fps. Having the display refresh at twice the frequency satisfies the Shannon-Nyquist sampling theorem. As evident in Figure 8, the RMS error increases as latency increases at all refresh rates. This is expected as latency of our system largely determines how well we can compensate for eye motions. In this case, the error is roughly linear. There does not seem to be a natural threshold for acceptable latency. In subjective tests of video output, it seemed that the stability and sharpness of the image was tolerable when the latency was less than about 25 ms. However, further investigation on different use cases and a more objective metric would be required to provide a definitive recommendation.

We can conclude that at this moment, a real prototype would be infeasible with the current recommendations. Our findings suggest that a system would need to display images at 60 fps or greater and have a total latency of 25 ms or less. As discussed in the overview of head mounted displays, there are many options that currently refresh at a rate faster than the desired 60 fps. However, the current solutions for gaze tracking are too slow or invasive to be feasible. The Eyetribe Gazetracker we investigated only took measurements at 30 fps. This means, the measurement alone would take 33ms before any computation in the system has begun. The gold standard for gaze tracking, VOG, operates at 250 Hz, but requires invasive pads stuck to your eyelids in a precise fashion. This would leave enough time for computation, but would not be conducive to an integrated system.

VI. Future Work

We showed that our solution can accurately simulate Nyastagmus and our proposed compensation system can be effective. Future work in this area would involve some improvements to our metrics, the system, and a wider variety of test cases. The metrics we are currently using are very dependent on the input image. Some metrics could include smarter image-invariant blur metrics or perceptual differences such as the Just-Noticeable Difference.

A small part of our compensation system, predistortion using wiener deconvolution, can be improved by accommodating for motion blur of the eye between frames, particularly during the saccade, when it isn't possible to display frames fast enough to reduce motion blur. However, the way this predistortion is implemented can use improvement. Currently, the last two eye positions are used to predict the location of the next location. Machine learning or adaptive signal processing could greatly benefit the deconvolution by more accurately predicting where the eye will be.

An important feature we may want to add to a physical product is the ability to distinguish between involuntary and voluntary eye motions. Because nystagmus is generally an oscillatory motion, adaptive signal processing could also be applied to differentiate between oscillatory (potentially involuntary) and normal (voluntary) eye motion.

The future of HMDs for visual correction is bright. A requirement of making head mounted displays is the need to take into account many different aspects of vision and allowing them to be tuned to the individual. Extending these aspects to vision impairment will be part of the natural progression of HMDs, whether the process involves forming images correctly on the retina or erratic eye movements such as nystagmus. While these particular corrections require active use of a HMD, other visual impairments can be improved through prior training with HMDs. Already there is software in development (e.g. Diplopia) intended to offer vision therapy to people with amblyopia (lazy eye) and strabismus (crossed eye).

DIVISION OF LABOR

Joe is responsible for: capturing images and nystagmus eyes video, working with Eye Tribe, doing background research, and writing wiener deconvolution scripts.
Phil is responsible for: writing the base code for the system, including most of the system in the above document, and scripts to write out movie files
Maneeshika is responsible for: RMS error scripts, acutance scripts, background research, and additional eye tracking data

Each member participated equally in the preparation of the presentation, poster, and this paper.

Rea Rostosky collected data from the Eye Tribe for our Psych 221 project, which is used in our code here.

ACKNOWLEDGMENT

We would like to thank Prof. Wetzstein, and our TA Issac Kauvar for their help coming up with a project idea and its implementation. We would also like to thank our Psych 221 instructors, Prof. Wandell and Joyce Farrell, and our TA, Haomiao Jiang for their advice and input.

REFERENCES

[1] Schiavi, C., Fresina M. Nystagmus - A Brief Review. *European Ophthalmic Review*, 2(1):53-4 (2009)

[2] Straube, A., Bronstein, A., & Straumann, D. Nystagmus and oscillopsia. *European Journal of Neurology*, 19(1), 6–14 (2012)

[3] Nystagmus, *Royal National Institute of Blid People (RNIB)*, 2010. Accessed 3/12/15. https://www.rnib.org.uk/eye-health-eye-conditions-z-eye-conditions/nystagmus

[4] A. Iijima, H. Minamitani, and N. Ishikawa, "Image analysis of quick phase eye movements in nystagmus with high-speed video system.," *Med. Biol. Eng. Comput.*, vol. 39, pp. 2–7, 2001.

[5] Image Systems Engineering Toolbox Biology Module, http://isetbio.org/

[6] Morgan, M. J. & Benton, S. Motion-deblurring in human vision. *Nature* **340**, 385-386 (1989).

[7] "Is there a way to detect if an image is blurry?" Stackoverflow. http://stackoverflow.com/questions/7765810/is-there-a-way-to-detect-if-an-image-is-blurry