# Stay Focused!
# Auto Focusing Lenses for Correcting Presbyopia

Elias Wu
Stanford University
eliwu@stanford.edu

## Abstract

*Development of active glasses for mitigating presbyopia has become an active area of research, facilitated by the miniaturization of focus-tunable lenses as well as the increased processing power of modern computers. However, current systems are still not very user friendly, or are limited in ability to solve the issue completely.*

*In our project, we implement a full system to correct for presbyopia based on where the user is looking. We first create an improved housing in order to mount to the user's head more optimally, and subsequently devise a novel method of calibration-free eyetracking for better ease of use.*

*We show in our results that the headset we produce does indeed fit better on user's heads and creates a better user experience, and that our model, while not optimal, does indeed predict a gaze position that, with fine tuning, may be able to replace the current eye tracking software.*

## 1. Introduction

Presbyopia is a very common side effect of the aging process. It describes the progressively worsening ability to focus on close objects due to the hardening of the eye's lens [9]. People with presbyopia may suffer headaches and eyestrain as side effects. It is a nearly universal condition with advancing age [15].

Currently, the easiest method of correcting this type of degeneration is through the use of reading glasses, or fixed focus lenses that refocus the image for the viewer. However, this is far from a perfect solution. Some drawbacks include that lenses at a fixed focus cannot be used in all situations, and many different glasses are needed for people with varying levels of presbyopia, requiring the user to select from a variety of different types [1].

## 2. Problem

As noted, fixed focus lenses for people with presbyiopia is not a perfect solution due to the fact that different people have differing levels of presbyopia, and need to choose glasses to fit their needs, as well as the fact that these reading glasses cannot be used to view farther objects. We thus need a method of automatically focusing the lens to whatever the users eyes are pointed at. In this paper, we update a design of auto-focusing lenses for people with presbyopia and propose novel methods of performing the focus calculation for easier use.

### 2.1. Literature review

The cure for presbyopia has been long sought - there have been numerous pharmacological methods to control presbyopia. Some of these include eyedrops that change optical qualities, eyedrops that change material properties of the lens [18], lenses that go inside the eye to correct out presbyopia [8], and of course, reading glasses. This project thus aims to avoid the hassle and potential side effects of pharmacological methods as well as avoid the common problems of using standard reading glasses.

One of the technologies we use is focus-tunable lenses. These are not a new idea - commercial versions have existed for a while [2] [5], and there have been several research developments in the area: Ren et al. explore liquid crystal tunable lenses [17], while Santiago-Alvarado et al. explore the difference between solid lenses and focus tunable lenses [19].

Focus-tunable lenses have enjoyed a prominent presence in Virtual Reality and Augmented Reality. While light-field displays to introduce focal cues for virtual reality are being researched [12], focus-tunable lenses to provide focus cues have recently gained traction. Konrad et al. explore methods of using focus-tunable lenses in conjunction with image processing methods to reduce the effect of the vergence accommodation conflict in near-eye displays [13]. Padmanaban et al. explore methods of providing changes in focus that accompany natural vision by using focus-tunable

Figure 1. A view of the system that Nazmul et al. created.

lenses, gaze tracking, and actuated displays [14].

Outside of VR and AR, tunable-focus lenses for glasses are not a new idea (they were even discussed in Dune!). However, auto-focusing lenses for presbyopia are a relatively new idea; there has only been one similar project with similar end goals, from the University of Utah.

Nazmul et al. created a 3d-printed set of auto-focusing glasses that use an IR rangefinder to focus onto where the lenses are pointed [16]. This IR rangefinder was placed in the middle of the glasses' bridge, between the eyes. Thus, the glasses would not refocus depending on the gaze of the user, but rather the direction of the head. Ultimately, however, this project's main focus was not to create these glasses, but to test out the focus-tunable lens [10].

# 3. Methods

## 3.1. Implementation

We implement our system with a Intel Realsense depth camera [4] for rangefinding and gaze estimation, and two Pupil Labs eyetracking cameras [6] in order to determine where the user is looking.

At a high level, the Pupil Labs eyetrackers first determine the gaze point of the user in the scene that the user is looking at. Subsequently, the RealSense camera takes a measurement of the distance to the gaze point. Finally, software adjusts the lenses to focus according to the distance recorded by the RealSense camera.

This projects specific goal was two fold: build on the AutoFocals 1 frame, and explore methods of calibration-free eyetracking to remove the hassle of eye-tracking camera calibration for every new user.

## 3.2. Physical Hardware

The original AutoFocals frame was originally built on an optometrist's test frame. This configuration was for easy adjustability at the expense of rigidity (as the frame was originally designed to fit lightweight optics without wires), and comfort. Numerous add-on mounts were required to mount the additional hardware, and lack of inherent cable management meant many suboptimal design choices had to be made for the user to wear the glasses properly.

Figure 2. An early ski goggle model with a intruding nose bridge region.

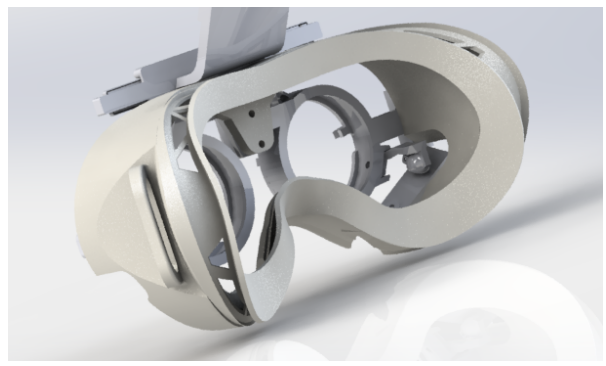

Figure 3. A view of the back of the final headset, showcasing a deeper nosebridge.

There were thus three subgoals of the new frame: stability, adjustability, and better cable management.

### 3.2.1 Stability

The new frame had to mount securely to the users head given quick motions specifically, the eye tracking cameras had to have the same view of the users eyes no matter the users motion. 3.

To do this, we designed a frame from scratch, utilizing plans from an old Bolle ski goggle. We created several of these prototypes with slightly differing nose bridges and sizes. Ultimately, we ended up with a design that, with extra foam padding, comfortably fit an average person's head.

Strap mounts were then added to allow a tight elastic strap to be tightened around the user's head, securing the headset to the user even under heavy movement.

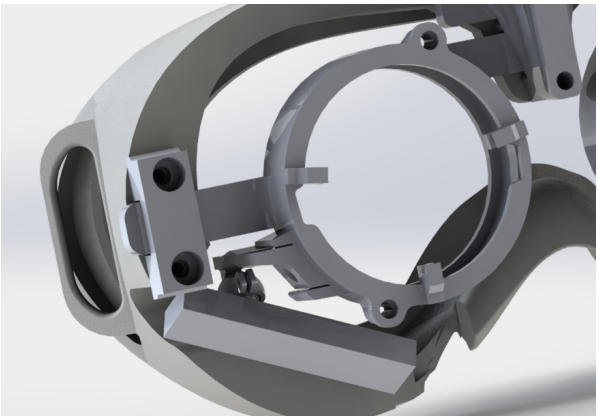

Figure 4. A closeup of the strap mount, the adjustable lens holders, and the adjustable eye tracking mount

#### 3.2.2 Adjustability

The frame had to have adjustable IPD to fit different people from 55mm to 75mm. Additionally, the eye cameras had to be on an adjustable ball mount to properly see the users eyes. A sliding-bar design was used to allow for infinite variability of IPD. the clamps on the sides are tightened in order to secure the location of the lens holders.

The eye tracking cameras were attached to the lens holders, as they are usually adjusted in tandem with the lens. This allows for easier adjusting of the entire setup for optimal performance. A ball mount ensures variability in the eye tracking cameras in relation to the lens for fine tuning.

In order to account for prescription lenses, tabs are placed on the lens holder to place corrective lenses.

Finally, the headstrap is adjustable for tightness in order to accommodate different head sizes and comfort levels.

#### 3.2.3 Cable Management

An important issue of the previous design for AutoFocals was the lack of proper cable management, hindering the use of the headset. In the revision, all cables are routed into one cable on the headset itself, simplifying cable placement on the head and overall increasing ease of use. This is done by unifying all cables to the top of the ”launchboard”. The cables are then secured to the back of the headstrap, creating a sort of ”pulley effect” that resists the natural droop of the headset due to weight.

### 3.3. Software: TrackingNet

We then explored a novel method of calibration-free eye-tracking to streamline the process of using the auto focusing lenses.

TrackingNet was created to take in data from the eye tracking cameras as well as Pupil Labs' pupil-tracking software (which does not need calibration), and output an x-y coordinate on the RealSense Camera's front view camera picture, much in the same way the Pupil Labs' default software works.

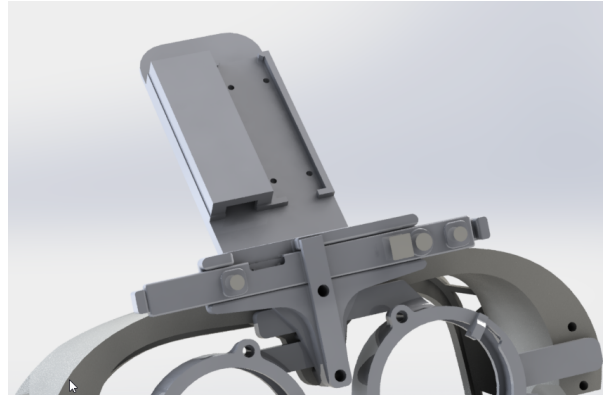

Figure 5. A closeup of the ”launchboard” that holds the lens drivers and unifies all cables.

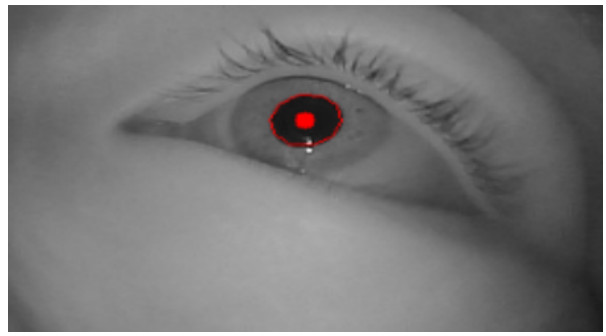

Figure 6. Raw eye camera image with pupil location and center.

Of course, the difference is that TrackingNet does not require a new calibration for new users. At a high level, TrackingNet is a deep convolutional neural network that utilizes pretrained models as well as a 5-layer fully-connected module in order to make 2-length predictions (X and Y of gaze position).

#### 3.3.1 Data Gathering

In order to train our model, we used the headset we created, calibrated the pupil labs' software on 2 people, and recorded the eye camera images, the calculated pupil statistics (including pupil center, and eyeball shape), as well as the output statistics (including X, y, and z predictions). This data was conditioned to two arrays - input (images, pupil statistics), and output (X,Y). A view of the data produced by the eyetrackers as input is produced in 6.

#### 3.3.2 Network Architecture

We use PyTorch [7] for the network implementation, running on a Google Cloud Platform [3] instance. This instance includes a 2-core CPU with a Tesla K80 GPU for

faster training and inference. After several tests of different networks, we settled on a particular network structure that worked the best. This consisted of first using ResNet-50 features [11] as the convolutional part of the network to handle the eye camera images.

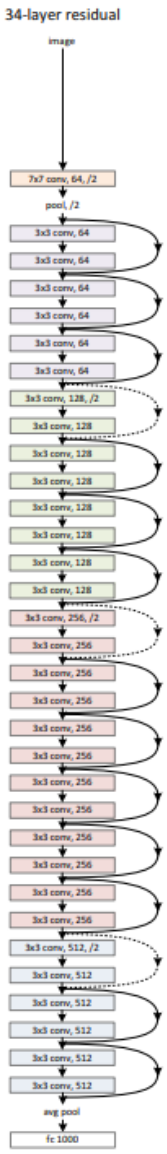


Figure 7. A smaller version (34 instead of 50) of the resnet we used in our network.

The output of the Resnet-50 without its last layer (output size of 2048) was fed into 5 fully connected layers each "squeezing" down in size. Additionally, two dropout layers are introduced between the first and last fully connected layers in order to prevent overfit. A figure representing our final network is presented in 8.

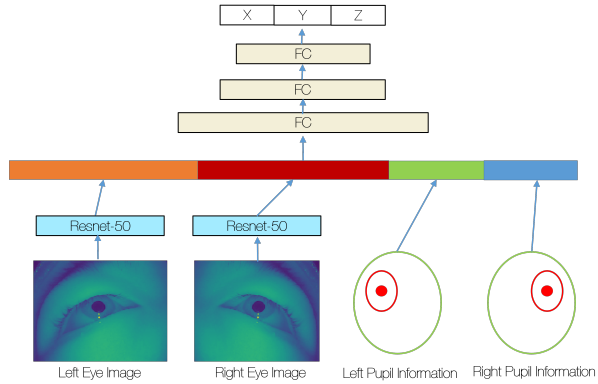


Figure 8. A graph representation of our final network.

The loss layer was an MSE Loss, which generally approximates L2 loss. This is because we want the network to be providing X and Y predictions which we want to be close in Euclidean distance to the correct coordinate, rather than one-hot class predictions.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2.$$

Figure 9. MSE Loss.

# 4. Experiments and results

## 4.1. Headset Results

In testing our system, we found that overall, the headset did in fact satisfy the requirements set out - it was much more stable than the old system on the user's head no matter the motion, the lenses and eye tracking cameras were just as adjustable as the old system, and cable management was much better, as it no longer got tangled with the user's appendages as the user moved around. Aesthetically, it was a little more bulky and less minimalistic, at the benefit of a better user experience.

## 4.2. TrackingNet Results

In order to produce a metric for the trackingnet results, we devised a Euclidean distance metric presented in 1.

$$Loss = 1/(\exp(||X_{target} - X_{predicted}||)) \quad (1)$$

$$(2)$$

This metric returns a value from 0 to 1, exponentially increasing as distance becomes smaller.

Our final model produced a score of 0.49, and loss reduced from 200 to 55. A graph of the training can be seen in 10

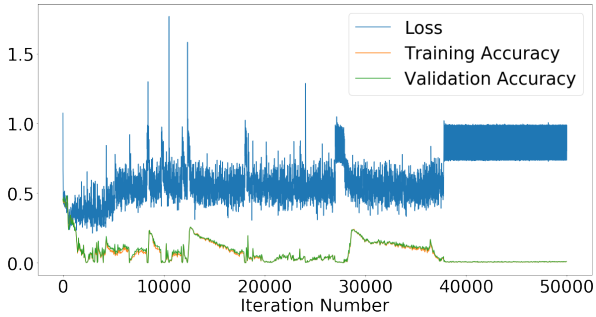


Figure 10. The training curve of our final model.

## 5. Discussion

A cursory glance at several test cases of our system reveals that although the predicted points are generally within the same "ballpark" as the target numbers, they are not close enough the replace the calibration-required system. More data and more training is needed in order for more accurate predictions.

Furthermore, in our testing, we see that using the $z$ prediction lowers prediction accuracy for $x$ and $y$ greatly - we attribute this to the fact that the $z$ numbers were consistently much higher than the $x$ and $y$, and were generally stable (low variance). Thus, if the network predicted $z$ more accurately, its score would be generally higher regardless of how accurate the $x$ and $y$ predictions were.

We can also see form the training curve that training seems to not learn much after the first few iterations, and has wild jumps once in a while. This may be attribute to the fact that MSE loss may not be the optimal loss function for this type of task. A simple L2 loss may do better. We use ADAM optimization for training, and this may not be as well suited to this task as it is suited to classification as well.

### 5.1. Future work

We have several different ideas for future work and extensions; these pertain to both increasing the usability and aesthetics of the headset as well as increasing accuracy of the model in order to replace parts of the current eye tracking software.

First, the frame needs to be sleeker and lighter. Although the frame fits well on the user's head, there is still a marked weight hanging in both the front and rear of the head, causing slight discomfort. A better designed headset would have less mass, and would also be more minimalistic, minimizing the visual impact on the user's head.

Cables should be lightened - the current cables are very heavy, to the point where they can act as a counterbalance to the main headset. In the future, lighter cables should be used; a good way to do this may be to implement a hub on the headset itself so there is only one cable coming from the headset itself. The end goal of this device would be to have it self-powered so no cables would protrude. As such, however, this goal is not easily achievable in a small form factor due to power and processing requirements of the depth sensing camera.

Different models for calibration-free eyetracking should be explored. The biggest flaw appears to be the use of ResNet-50 features as the convolutional layers of the network. ResNet-50 is optimized for classification, which means it may not be optimal for the task of finding the gaze position from images of the eyes. In particular, the convolution layers would ideally learn the relation between the position of the eyes in the image and the pupil in the image, to take over the "calibration" part of the equation. ResNet-50 features may not be able to do this adequately, leaving the work to the less-sophisticated fully connected layers.

Finally, different lenses should be explored for the final headset. The current lenses are bulky by themselves. Additionally, the driver boards for the lenses are quite bulky. Fluid-filled lenses also have the issue of distortion at the edges, as well as distortion due to gravity. An example of this distortion is presented in 11

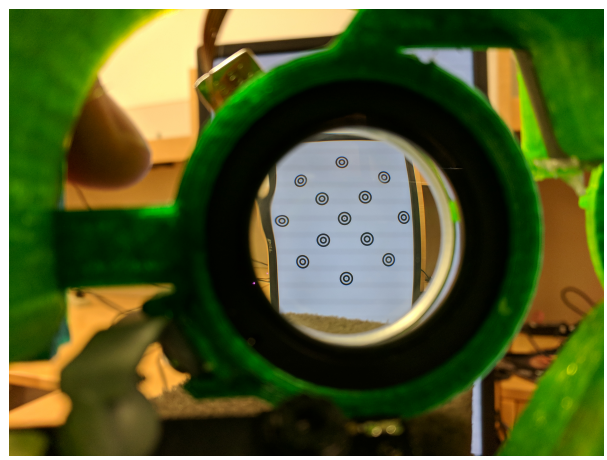

Figure 11. visible distortion at the edges of the lens

**Acknowledgments**

The author would like to thank the teaching staff of EE367 for offering a captivating and organized course. We especially enjoyed the demo sessions and appreciate the feedback we received.

## References

[1] Drugstore reading glasses: How to pick a perfect pair, 2018.
[2] Focus tunable lenses, 2018.
[3] Google cloud platform, 2018.
[4] Intel realsense, 2018.
[5] Optosigma focus tunable lenses, 2018.
[6] Pupil labs eyetracker, 2018.

[7] Pytorch, 2018.

[8] A. T. Epitropoulos. Visual and refractive outcomes of a toric presbyopia-correcting intraocular lens. In *Journal of ophthalmology*, 2016.

[9] A. Glasser and M. C. Campbell. Presbyopia and the optical changes in the human crystalline lens with age. *Vision Research*, 38(2):209 – 229, 1998.

[10] N. Hasan, A. Banerjee, H. Kim, and C. H. Mastrangelo. Tunable-focus lens for adaptive eyeglasses. *Optics Express*, 25(2):1221 – 1233, 2017.

[11] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.

[12] F. Huang, K. Chen, and G. Wetzstein. The Light Field Stereoscope: Immersive Computer Graphics via Factored Near-Eye Light Field Displays with Focus Cues. *ACM Trans. Graph. (SIGGRAPH)*, (4), 2015.

[13] R. Konrad, N. Padmanaban, K. Molner, E. Cooper, and G. Wetzstein. Accommodation-invariant Computational Near-eye Displays. *ACM Trans. Graph. (SIGGRAPH)*, (4), 2017.

[14] N. Padmanaban, R. Konrad, T. Stramer, E. A. Cooper, and G. Wetzstein. Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays. *Proceedings of the National Academy of Sciences*, 2017.

[15] I. Patel and S. K. West. Presbyopia: prevalence, impact, and interventions. *Community Eye Health*, 20(63):40 – 41, 2007.

[16] A. Potenza. These smart glasses automatically focus on what you're looking at., 2017.

[17] H. Ren, Y.-H. Fan, S. Gauza, and S.-T. Wu. Tunable-focus flat liquid crystal spherical lens. *Applied Physics Letters*, 84(23):4789–4791, 2004.

[18] A. Renna, J. L. Alió, and L. F. Vejarano. Pharmacological treatments of presbyopia: a review of modern perspectives. *Eye and Vision*, 4(1):3, 2017.

[19] A. Santiago-Alvarado, S. Vazquez-Montiel, J. Munoz-Lpez, V. M. Cruz-Martnez, G. Daz-Gonzlez, and M. Campos-Garca. Comparison between liquid and solid tunable focus lenses. *Journal of Physics: Conference Series*, 274(1):012101, 2011.