

Supplemental: Accommodation and Comfort in Head-Mounted Displays

GEORGE-ALEX KOULIERIS, Inria, Université Côte d’Azur
BEE BUI, University of California, Berkeley
MARTIN S. BANKS, University of California, Berkeley
GEORGE DRETTAKIS, Inria, Université Côte d’Azur

This document provides supplemental material for the submission "Accommodation and Comfort in Head-Mounted Displays".

CCS Concepts: • **Computing methodologies** → **Perception**; *Virtual reality*;

Additional Key Words and Phrases: head-mounted displays, perception, vergence-accommodation conflict

ACM Reference format:

George-Alex Koulieris, Bee Bui, Martin S. Banks, and George Drettakis. 2017. Supplemental: Accommodation and Comfort in Head-Mounted Displays. *ACM Trans. Graph.* 36, 4, Article 87 (July 2017), 5 pages. DOI: 0000001.0000001_2

1 HMD AND MEASUREMENT DEVICE

1.1 Focus-adjustable-lens Calibration

To calibrate the focus-adjustable lenses, we placed a camera behind them in the HMD, and manually focused the camera such that a Maltese cross on the display was sharpest for each of several electric current values. Then, without altering the camera focus, we directed the camera to a printed Maltese cross. We measured the distance from the camera to the cross that made the cross appear sharpest through the camera. This value in diopters was the estimate of the focal power associated with the input current. We did this several times for each value of the current and found that the measurements were repeatable. The relationship between current and focal power can supposedly be dependent on ambient temperature, so we performed the calibration at different temperatures. We observed no discernible effect for the range of temperatures that were likely to occur in the experiments (~25C) (Fig. 1).

1.2 Calibration of the Grand Seiko WAM-5500 Autorefractor

As mentioned in the main paper, to achieve a sharp image of the cornea on the autorefractor camera, we inserted a -0.75D offset lens in the optical train of our setup. However, the combined optical power of the lens train due to the optical offset has to be re-estimated. If not, this could contaminate the accommodation measurements. To alleviate this issue, we used an experimental subject to calibrate our setup.

Publication rights licensed to ACM. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.
© 2017 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2017/7-ART87 \$15.00
DOI: 0000001.0000001_2

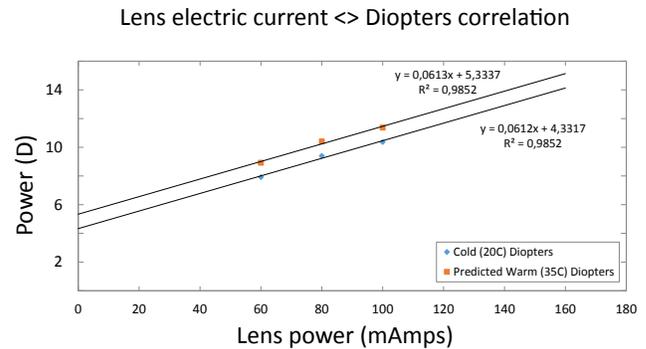


Fig. 1. Lens power prediction model based on lens current - diopters correlation for various currents and temperatures.

We applied tropicamide to the subject’s eyes and then measured his refraction without any lens in the measurement setup for different refraction powers using additional lenses (Figure 2). Then, we inserted the -0.75D offset lens and re-measured the subject using the same offsets. This allowed us to obtain a clear mapping of measured values with the offset lens to real values without the offset lens (Figure 3).

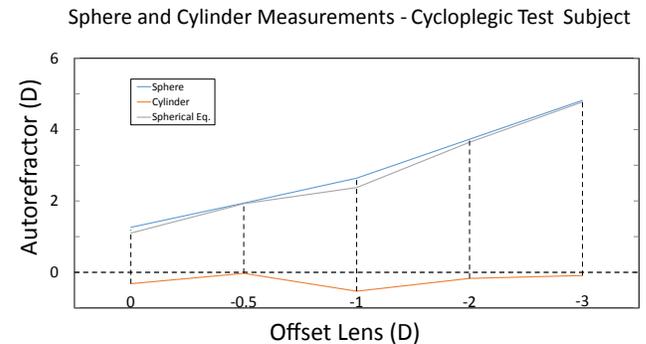


Fig. 2. Sphere and Cylinder measurements of a cycloplegic subject.

1.3 Lens Breathing

To eliminate lens magnification (“breathing”) as the lenses change powers, we had to resize the image displayed on the HMD on-the-fly by a scaling factor. To calculate the scaling factor, we define

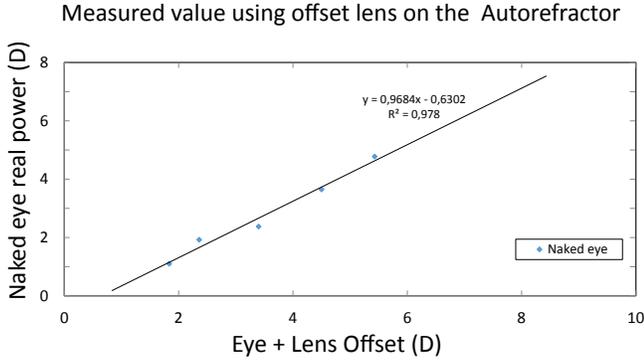


Fig. 3. Measured values of refraction using the autorefractor and the offset lens. Real values of the measured eye are obtained via fitting a curve to the data.

l_s as the diagonal length of the image for each eye on the display, d_{COP} as the display panel distance from the optical center of the lenses (the Center-of-Projection, CoP) and f as the focal length of the adjustable lenses. We re-sized the rendered image on the display based on the variable focal length of the lenses to eliminate lens breathing as follows [Cooper et al. 2012]:

For any given adjustable-lens focal length, a magnification m was estimated:

$$m = d_{COP}/f \tag{1}$$

We thus estimated a corrected diagonal image size l_s^{CORR} and re-sized the image accordingly for each frame:

$$l_s^{CORR} = m^{-1}l_s \tag{2}$$

1.4 Depth-of-Field Rendering

To perform DoF rendering, we first estimated the Circle of Confusion (CoC) due to defocus on the plane of projection. The diameter of the CoC in world coordinates is:

$$CoC = \left| A \times \frac{F(P-d)}{d(P-F)} \right| \tag{3}$$

where A is pupil diameter measured with the autorefractor, F is focal length, P is distance of plane of focus, d is distance of blurred object from the center of projection in world coordinates, and I is distance of plane of projection from the center of projection in world coordinates (Fig. 4).

We get object distance d by linearizing z values in the z -buffer:

$$d = \frac{-z_{far} \times z_{near}}{z \times (z_{far} - z_{near}) - z_{far}} \tag{4}$$

Step 1: Render a sharp pinhole camera image of the scene and scene depths in the z -buffer.

Step 2: For each pixel, blur the previously rendered sharp scene by varying amounts per pixel according to its CoC . Because CoC is in world coordinates, we map it to pixels depending on display buffer resolution.

We then sample the original image using a disc filter [Nguyen 2007; Potmesil and Chakravarty 1982], a high-performance approach similar to the one used in [Konrad et al. 2016].

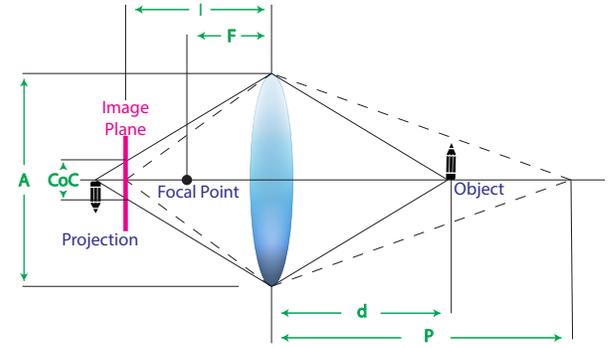


Fig. 4. Circle of confusion (CoC) in DoF rendering. A , pupil diameter measured with the autorefractor, F , focal length, P , distance of plane of focus, d , distance of blurred object from the center of projection in world coordinates, I , distance of plane of projection from the center of projection in world coordinates.

2 OPTOTUNE LENSES CUSTOM COMMUNICATION PROTOCOL

The dynamic lenses can be controlled via a serial communication protocol. Optotune provides the serial communication protocol of the lens controller along with a Labview implementation and a C# GUI-based application. However, using the lenses from inside Unity3D™ is not directly possible. Unity3D is based on the Mono implementation of the .NET framework which offers a poorly implemented SerialPort assembly. When controlling the lenses using the Mono Assembly directly from inside Unity3D, the application crashes. To avoid this issue we wrote a custom driver implementation for the lenses that overcomes the SerialPort *read* issues by ignoring messages sent from the lens controllers. Our protocol only transfers commands to the lenses by encoding power values and CRC checks without requiring responses to be received.

3 MEASUREMENTS - TRACES

In this section example autorefractor traces and gain data will be presented.

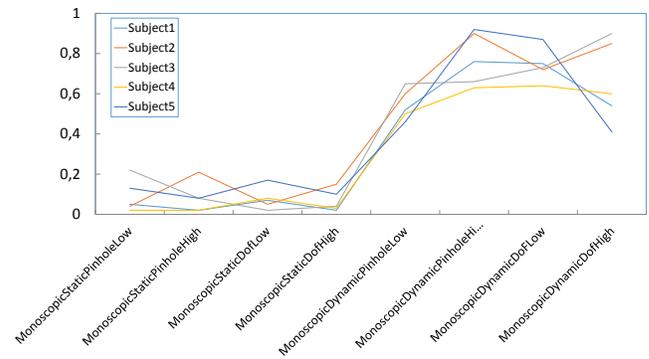


Fig. 5. Per-subject gain in all monocular conditions.

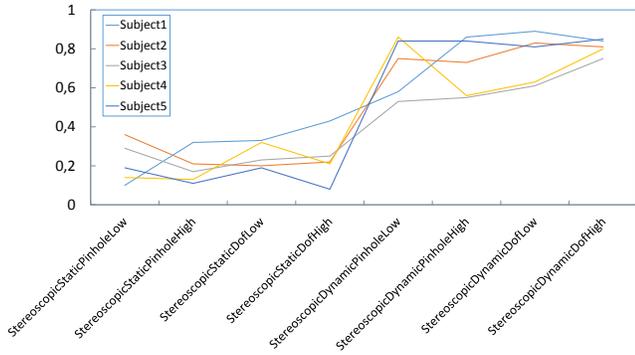


Fig. 6. Per-subject gain in all binocular conditions.

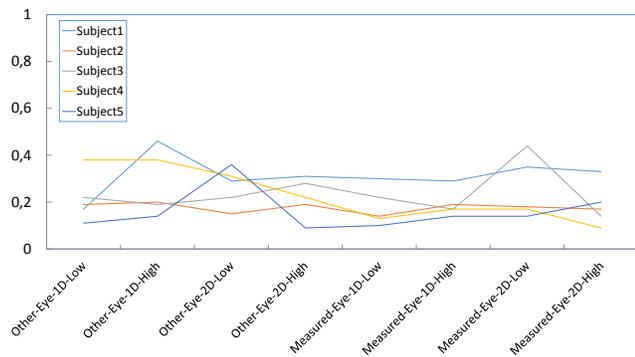


Fig. 7. Per-subject gain in all monovision conditions.

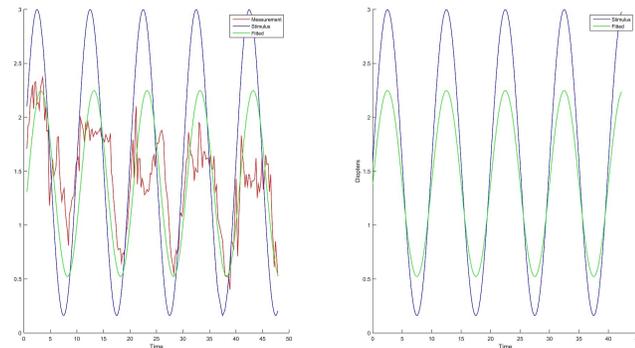


Fig. 8. Stimulus, raw data and fitted sine wave for subject JH in the monoscopic, dynamic lens, pinhole, low speed condition.

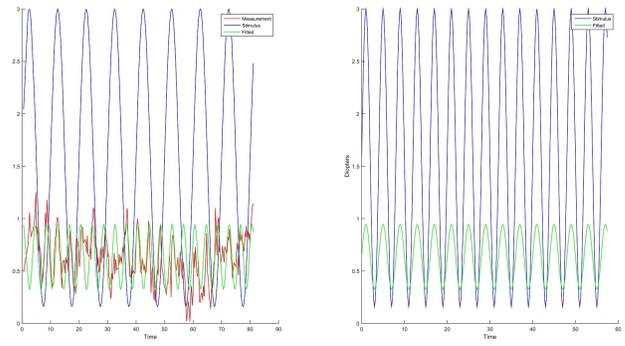


Fig. 9. Stimulus, raw data and fitted sine wave for subject JH in the monoscopic, static lens, pinhole, high speed condition. Notice the low gain of the response sine wave when not using the dynamic lenses.

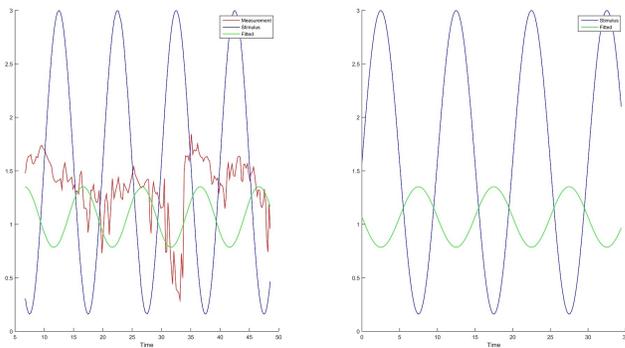


Fig. 10. Stimulus, raw data and fitted sine wave for subject WY in the stereoscopic, static lens, depth-of-field, low speed condition. Notice the low gain of the response sine wave from vergence alone.

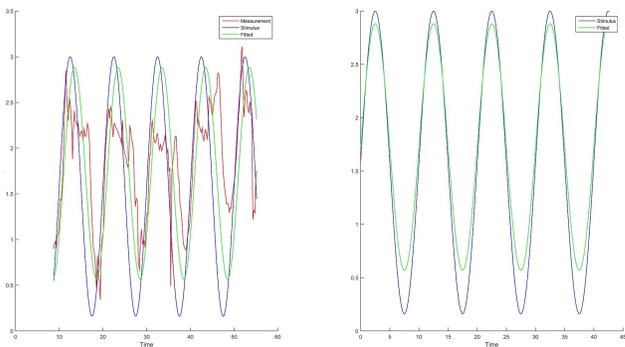


Fig. 11. Stimulus, raw data and fitted sine wave for subject WY in the stereoscopic, dynamic lens, depth-of-field, low speed condition. Notice the higher gain of the response sine wave from dynamic lenses when compared with the static lenses (previous figure).

4 PREDICTING DISCOMFORT FROM ACCOMMODATION DATA IN THE MONOVISION CONDITIONS

In Section 4.4 of the main paper, we describe a predictor of discomfort based on accommodation data. This predictor was necessary since the gain estimates alone cannot calculate the accumulated VA conflict in the case of monovision, since the VA conflict is different for each eye. We thus estimate the average accumulated VA conflict

for the monovision conditions, by estimating the VA conflict for each eye separately and then averaging the values. In this appendix,

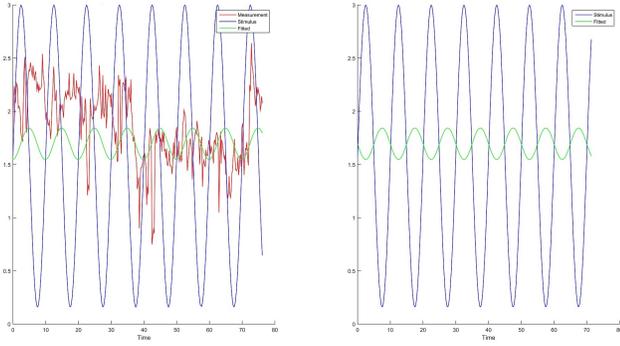


Fig. 12. Stimulus, raw data and fitted sine wave for subject WY in the monovision, 1D offset on the measured eye, low speed condition. Notice the step-like behavior in accommodation gain. The subject switched between two levels of accommodation; a near and a far level.

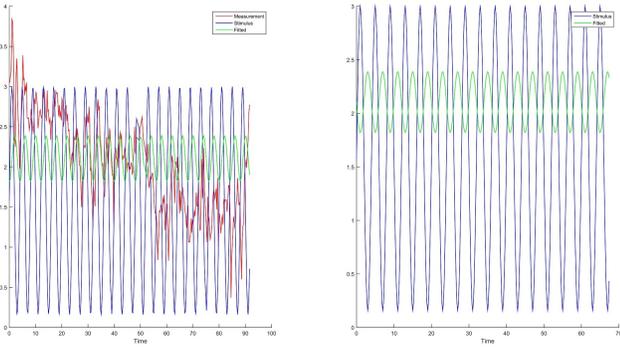


Fig. 13. Stimulus, raw data and fitted sine wave for subject JVV in the monovision, 1D offset on the non-measured eye, high speed condition. Notice the step-like behavior in accommodation gain. The subject seems to exhibit three levels of accommodation; a near and a far level, and a third, possibly driven from vergence or blur.

we present a different method to estimate the accumulated VA conflict for each eye in the monovision conditions that yields similar results. The difference with the method presented in the main paper, is that the one presented here hypothesizes that the brain is switching percepts to minimize blur at each viewing distance. As such the estimated error is slightly smaller.

The accumulated VA conflict in the Monovision conditions. We measured the accommodative response and we know the vergence stimulus over time. We assume that the vergence response is equal to the vergence stimulus since the subjects were fixating on the target. To compensate for errors introduced due to synchronization or phase we first align the vergence stimulus to the accommodation measurements. We do this by estimating the phase offset (\pm half cycle) that minimizes the accumulated VA conflict for all conditions; this procedure removes the effect of phase on the accumulated VA conflict.

We then calculate the moment-to-moment vergence-accommodation conflict (VA conflict). For each sample the conflict is the absolute value of the difference in accommodation response and vergence stimulus. We estimate the mean value of those differences and have a metric that may be able to predict discomfort.

In our experiment the vergence stimulus ranged from 0.17 – 3D for all conditions. The blur stimulus is set to 0.77D for the far eye in monovision conditions and 1.77D or 2.77D for the near eye depending on the condition.

In the monovision conditions each eye needs to be treated separately since the blur stimulus is different for each eye. There is always going to be one eye whose accommodation is closer to vergence than the other eye. In this calculation we hypothesize that the brain will be switching the percept from one eye to the other depending on target distance to obtain the sharpest image and minimize the VA conflict. We thus need to analyze the moment-to-moment VA conflict for each eye separately as a function of target distance. However, we can only measure the mean value of the average VA conflict for the two eyes, since accommodation is yoked between the eyes and as such a measurement is only able to get a single, mean value. There is no way to know during our measurements which eye is active, i.e. currently defining the percept.

An 1D offset example. The minimum mean value of the VA conflict measured for both eyes was 0.72D. Considering the 1D offset conditions (this can be done similarly for the 2D conditions) let us derive the exact VA conflict perceived by each eye in our HMD.

The far eye (set to 0.77D) will have a conflict of 0.60D at the furthest stimulus distance ($|0.17D - 0.77D|$) and a conflict of 2.23D at the nearest stimulus distance ($|3D - 0.77D|$). Zero conflict will be at 0.77D ($|0.77D - 0.77D|$). As such, for the far eye, the VA conflict error function is:

$$VA_{conflictFAR} = \begin{cases} x - 0.77 & x > 0.77 \\ -x + 0.77 & x \leq 0.77 \end{cases}$$

The near eye (set to 1.77D) will have a conflict of 1.6D at the furthest stimulus distance ($|0.17D - 1.77D|$) and a conflict of 1.23D at the nearest stimulus distance ($|3D - 1.77D|$). Zero conflict will be at 1.77D ($|1.77 - 1.77|$). For the near eye, the VA conflict error function is:

$$VA_{conflictNEAR} = \begin{cases} x - 1.77 & x > 1.77 \\ -x + 1.77 & x \leq 1.77 \end{cases}$$

At the intersection of those error functions ($x = 1.27$) the brain is expected to select the path of the minimum VA conflict and thus will opt to switch eyes to get the image from the sharper eye. For example the brain is expected to switch to the far eye if the target is moving further than 1.27D (target distance $< 1.27D$) and to the near eye if the target is moving closer than 1.27D (target distance $> 1.27D$).

By plotting the functions (Fig. 14) we can now observe that when the target distance is less than 0.77D or more than 1.77D the error in accommodation is constantly 1D more for the eye that is not been used. Between 0.77D and 1.77D the VA conflict difference has a different value at each distance that we must estimate.

We perform the calculations for each eye separately, by identifying which eye is expected to be active depending on target distance

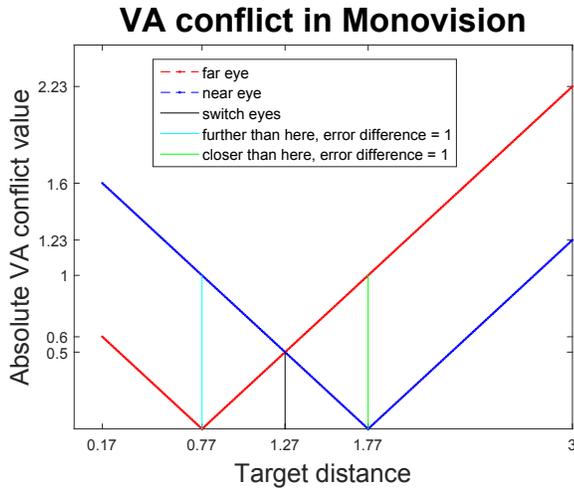


Fig. 14. Perceived VA conflict in monovision conditions for the near and far eye.

and then estimating the exact VA conflict for the eye not in use. It is this conflict disparity between the two eyes that is hypothesized to cause fatigue.

Far eye active. When the far eye's percept is selected (target distance $< 1.27D$) the target range from that switching point to the furthest distance is $1.1D$ [$1.27D - 0.17D$]. The VA conflict is eliminated at the $0.77D$ distance for that eye. Given that the mean monovision VA conflict from the data was $0.72D$ this entails that for the near eye and for the target range [$0.77D - 0.17D$] the VA conflict fraction will be $1D$ more than the far eye:

$$\frac{0.77D - 0.17D}{1.27D - 0.17D} \times (0.72D + 1D) = 0.93D \quad (5)$$

However, for the rest of that target range [$1.27D - 0.77D$] the average VA conflict is $0.5D$ (identity function ranging from 1 to $0D$) and as such

$$\frac{1.27D - 0.77D}{1.27D - 0.17D} \times (0.72 + 0.5D) = 0.55D \quad (6)$$

If we add those fractions together, we find that the near eye had an average VA conflict of $1.48D$ when the far eye was active which in turn had an average $0.72D$ of VA conflict.

Near eye active. When the near eye's percept is selected (target distance $> 1.27D$) the target range from the switching point to the closest distance is $1.73D$ ($3D - 1.27D$). The VA conflict is eliminated at the $1.77D$ distance for that eye. Given that the mean monovision VA conflict from the data was $0.72D$ this entails that for the far eye and for that target range [$3D - 1.77D$] the VA conflict fraction will be $1D$ more than the far eye which is:

$$\frac{3D - 1.77D}{3D - 1.27D} \times (0.72D + 1D) = 1.22D \quad (7)$$

However, for the rest of the range [$1.77D - 1.27D$] the average VA conflict is $0.5D$ (identity function ranging from 1 to $0D$) and as

such:

$$\frac{1.77D - 1.27D}{3D - 1.27D} \times (0.72D + 0.5D) = 0.35D \quad (8)$$

If we add those fractions together, we find that the far eye had an average VA conflict of $1.57D$ when the near eye was active which in turn had an average VA conflict of $0.72D$.

The two eyes perceive a different accommodation error at most target distances (except the $1.27D$ target distance where the conflicts measure equally $0.5D$). We expect that monovision may cause even more discomfort than the other conditions because of the difference in errors.

Consider for example an inactive near eye. While the stimulus vergence distance may be the same as for the active far eye, the focal power needed for the inactive near eye to accommodate is more when compared to the far eye. As a result we hypothesize that the near eye may actively attempt to force accommodation to the distance that it sees clearly and since accommodation is yoked between the eyes that is what may induce visual fatigue.

REFERENCES

- Emily A Cooper, Elise A Piazza, and Martin S Banks. 2012. The perceptual basis of common photographic practice. *Journal of vision* 12, 5 (2012).
- Robert Konrad, Emily A Cooper, and Gordon Wetzstein. 2016. Novel Optical Configurations for Virtual Reality: Evaluating User Preference and Performance with Focus-tunable and Monovision Near-eye Displays. In *SIGCHI 2016*. ACM, 1211–1220.
- Hubert Nguyen. 2007. *Gpu Gems 3*. Addison-Wesley Professional.
- Michael Potmesil and Indranil Chakravarty. 1982. Synthetic Image Generation with a Lens and Aperture Camera Model. *ACM Transactions on Graphics (TOG)* 1, 2 (1982), 85–108.