EyesOnMe: Investigating Haptic and Visual User Guidance for Near-Eye Positioning of Mobile Phones for Self-Eye-Examinations

Luca-Maxim Meinhardt
luca.meinhardt@uni-ulm.de
Institute of Media Informatics,
Ulm University
Germany

Kristof Van Laerhoven
kvl@eti.uni-siegen.de
Ubiquitous Computing,
University of Siegen
Germany

David Dobbelstein
david.dobbelstein@zeiss.com
Carl Zeiss AG
Oberkochen, Germany

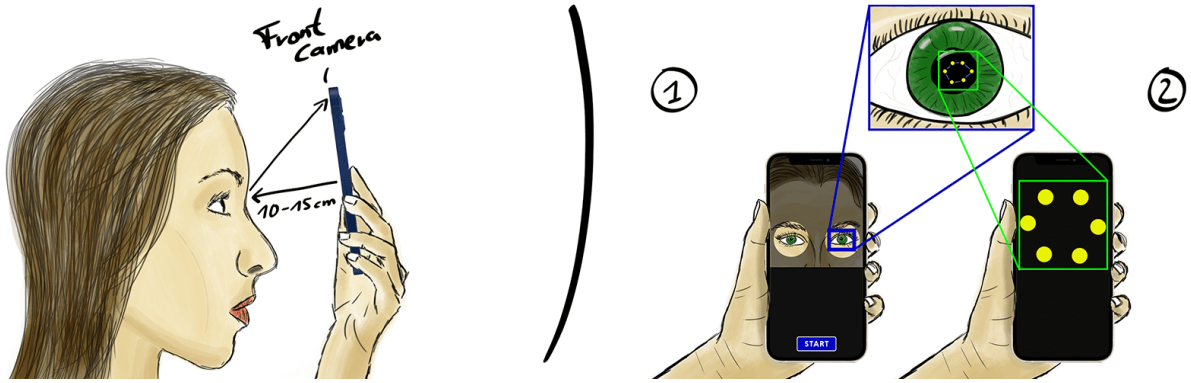


Figure 1: For self-eye-examinations, the user has to position the phone close to their eyes to capture a photo. In this work the procedure is divided into two steps: (1) The users are guided to position the phone at a near-eye position of 15cm utilizing dynamic distance rings and haptic feedback. (2) Subsequently, a close-up picture is taken from the user's eyes, including the reflection of the pattern displayed on the phone's screen.

ABSTRACT

The scarcity of professional ophthalmic equipment in rural areas and during exceptional situations such as the COVID-19 pandemic highlights the need for tele-ophthalmology. This late-breaking work presents a novel method for guiding users to a specific pose (3D position and 3D orientation) near the eye for mobile self-eye examinations using a smartphone. The user guidance is implemented utilizing haptic and visual modalities to guide the user and subsequently capture a close-up photo of the user's eyes. In a within-subject user study (n=24), the required time, success rate, and perceived demand for the visual and haptic feedback conditions were examined. The results indicate that haptic feedback was the most efficient and least cognitively demanding in the positioning task near the eye, whereas relying on only visual feedback can be more difficult due to the near focus point or refractive errors.

CHI EA '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3585799

CCS CONCEPTS

• **Human-centered computing** → *Usability testing*; *Empirical studies in interaction design*; Haptic devices; *Empirical studies in HCI*.

KEYWORDS

user guidance, near-eye positioning, mobile self-ophthalmology, smartphone

ACM Reference Format:
Luca-Maxim Meinhardt, Kristof Van Laerhoven, and David Dobbelstein. 2023. EyesOnMe: Investigating Haptic and Visual User Guidance for Near-Eye Positioning of Mobile Phones for Self-Eye-Examinations. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3544549.3585799

1 INTRODUCTION

The human's visual sense is the most essential to perceive the surrounding environment. However, visual impairments such as refractive errors are widely spread in the population [21]. These refractive errors can be corrected comfortably with optical aids, but professional diagnostic equipment is required for their determination. Ophthalmic equipment and professionals, however, are hardly available in the world's rural areas and during exceptional situations such as the COVID-19 pandemic [4, 26]. Both circumstances reveal the need for inexpensive and accessible mobile examination methods in the field of ophthalmology [3, 11, 23]. Additionally,

it reveals the need for self-eye-examination applications that are independent of professional ophthalmologists.

There are already limited solutions for smartphone applications that can detect refractive errors via capturing a close-up photo of the patient's eye, e.g., GoCheck Kids [18]. However, for mobile eye examinations, it is important to position the phone in a precise position in front of the user's eyes. Since GoCheck Kids [18] suffers from a missing user guidance, multiple attempts were needed to position the phone at the correct position to conduct the measurement [37]. This problem is supported by Pujari [30]. He mentioned that positioning the device at the optimal position for the eye is one of the major issues for future smartphone ophthalmology. Similar challenges with positioning the device close to the eye were reported by Barry et al. [8]. In their application for self-measure pupil sizes, they stated, "*[...] that older adult users would struggle to properly position and hold the phone to record their pupil by themselves.*". To address this problem, they had to 3D-print an attachment positioning the user at the correct distance to the phone. However, for future smartphone-based eye examinations, it would be beneficial to have a method for positioning the device without the need for additional attachments. Familiarity with positioning the face in front of the phone to unlock it could be leveraged to make near-eye positioning more intuitive. However, it should be noted that it is difficult for humans to focus on objects closer than about 20cm, particularly for those over 40 years, as noted in [16]. Furthermore, individuals with refractive errors such as myopia, hyperopia, or astigmatism may require additional non-visual guidance for proper near-eye positioning.

This work introduces a novel user-guided smartphone application to guide the user to a specific near-eye position for future self-eye-examinations by capturing a close-up photo of the eye. To validate that the desired six-degrees-of-freedom (6 DoF = 3D position + 3D orientation) position is reached and the user was prepared for the photo (did not blink), we displayed a pattern on the phone's screen and detected its reflection onto the user's eye with the phone's front camera (Figure 1). Using this reflected pattern can be useful for future eye measurement applications, such as detecting refractive errors [28] or examining strabismus [29, 31] with the phone. However, more importantly, with this method, we could compare four user guidance conditions (baseline, such as used in [18], dynamic visual rings, haptic, visual rings and haptic feedback) in a user study (n=24). The three dependent variables were (a) the time the user required to reach the 6 DoF position, (b) the task load index [20], and (c) the success rate whether the reflected pattern was detected or if the user blinked during the capturing.

Our study found that the haptic modality resulted in a significantly lower task load index compared to the baseline condition. However, there were no significant differences observed in the average success rate or average required time. This may be attributed to the within-subject counterbalanced design of the study, which resulted in a learning effect among all modalities, thus neutralizing any differences in the required time. When only comparing the initial trial of the task, in which the application was unfamiliar to participants, the haptic feedback resulted in the fastest performance. These findings suggest that utilizing non-visual guidance, such as haptic cues, may be more effective than solely relying on visual guidance. This may be because visibility at close distances is impaired for many individuals. However, further research is needed to evaluate the effectiveness of different non-visual guidance methods in near-eye positioning tasks.

Contribution Statements.

- Development and design of a user-guided smartphone application to guide the user toward a near-eye 6 DoF position with haptic and visual feedback to subsequently capture a close-up photo of the eyes.
- Conducting a user study (n=24) to compare the design parameters: baseline, dynamic visual rings, haptic feedback, and the combination of visual rings and haptic feedback. The results show that haptic feedback trends to achieve the lowest workload and highest success rate. Especially for the user's first trial, haptic feedback was the fastest in reaching the desired position.

2 RELATED WORK

2.1 Smartphone applications for self-eye-examination

As a result of a quantitative analysis by Aruljyothi et al. [7], 131 smartphone applications were found to deal with ophthalmology. 32% of them address vision screening. However, it has to be distinguished between apps supporting professional ophthalmologists and applications developed for patient use to self-examine the eyes. In a meta-review of ophthalmologist-supportive applications by Akkara and Kuriakose [2], multiple apps were named, such as Peek Acuity [28] and GoCheck Kids [18], which are used to examine refractive errors. Both applications were examined during field studies in academic publications [33, 37]. Pundlik et al. [31] found a use case for applying the camera's rear flash to screen the patients' eyes. They developed a smartphone application that implements the Hirschberg test to detect strabismus by analyzing the flash's reflection relative to the iris's center.

Contrary to ophthalmologist-supportive applications, there are only limited solutions for self-eye-examination smartphone applications in the area of ophthalmology. Early work published in 2010 used a mobile phone positioned close to the user's eyes [27]. They attached a lenslet to the phone, which refracts specific sets of patterns onto the user's retina. By calculating the perceived distance between the patterns and the eye, refractive errors could be examined [27]. A similar technique is deployed in Vision Tracker app from the US company EyeQue. By attaching an additional ocular to the phone's display, the user is asked to align a red and a green line visible through the ocular. After completing several exams, the users' refractive error values are provided [13]. The app developed by Bonfanti et al. [9] uses Google's Cardboard to detect a malfunction in stereo vision. The user is asked to look at a pattern of dots through the cardboard. Due to slightly shifted dots for the right and left eye, geometric shapes appear that the user is supposed to identify Bonfanti et al. [9]. Most recent work by Barry et al. [8] has demonstrated that new smartphones' near-infrared camera technology allows for the objective measurement of absolute pupil diameter with low errors. This has the potential to pave the way

for new smartphone-based self-examination methods in the field of ophthalmology.

2.2 User guidance for positioning and orientation tasks

User guidance, as defined by Lapointe et al. [25], is a method of remotely assisting users in performing specific tasks, commonly used in augmented reality (AR) scenarios to improve understanding of the spatial relationship between virtual and real objects. It can take various forms, such as haptic, auditory, or visual. In medicine, visual user guidance has been applied in smartphones to aid in needle insertion [22]. Similarly, Rojtberg [32] used visual guidance to aid in smartphone camera calibration for AR applications. Other forms of user guidance have also been studied, such as auditory feedback [14] and haptic feedback [6].

3 DETERMINATION OF THE NEAR-EYE 6 DOF POSITION

This chapter will introduce the procedure to determine the 6 DoF position to which the smartphone is supposed to be guided. Further, it explains the algorithm of the pattern detection, which is used to define whether the taken picture was successful or not (the user is shaking the phone or blinking during capturing).

3.1 Requirements for the phone's position

To detect the pattern unambiguously inside the user's eyes, it has to be reflected in the pupil, as the contrast between yellow dots and black pupils was highest. To achieve this, three requirements for the phone position were determined (see Figure 2a):

Eye and phone camera form an isosceles triangle with the pattern's center. This condition needs to be given to centrally reflect the pattern in the user's eyes. Here, mainly the relative y coordinate of the eye (with the front camera as the coordinate's origin) is crucial. When this condition is given, it can be assumed that the pattern is always reflected in the eye. However, also x coordinate is important to ensure that the phone is positioned at the center of both eyes.

Correct distance. The distance between the phone and the user's eye must be in a specific range for the close-up photo. If the user's face is too close to the phone, there is a higher chance for the pattern not to be reflected in the user's eye. On the other hand, if the face is too far away, the computer vision algorithm might have difficulties recognizing the reflected dot pattern. A distance of approx. 15cm was chosen for the near-eye position.

Correct user's gaze orientation. The examination of a user's eye gaze in relation to a mobile device's screen can provide valuable information about the reflection of patterns in the pupil. When the gaze is oriented orthogonally to the screen, it can be deduced that the pattern is reflected in the center of the pupil. Conversely, if the user's gaze is directed upwards or downwards, it is still reflected in the eye (due to the above requirements), but it is uncertain whether it will be reflected in the pupil. Hence, it is important to note that this phenomenon is distinct from that of a mirror, where the gaze is always orthogonal to the mirror's surface when looking at one's own eyes. In the case of a smartphone's front camera, the gaze orientation may not be orthogonal but rather directed downwards (see $\hat{y}$-*offset* in Figure 2b). To achieve an orthogonal gaze, the user must look at a point close to their forehead, as illustrated in Figure 2c. This effect is due to a shifting factor caused by the distance between the front camera and the center of the phone's display.

3.2 Determination of the near-eye position

To determine the z coordinate of the user's eye relative to the phone (distance between the eye and the phone), the cross-platform framework MediaPipe Iris was used [19]. It is based on a neural network for iris landmark detection and a pinhole camera model to infer distance from the iris size. The iris landmarks and the intercept theorems were also used to determine the eye's relative x and y coordinates, which were needed to fulfill the upper requirements for the optimal eye's position (see subsection 3.1).

3.3 Pattern Detection

For the purpose of experimentation, a pattern consisting of six yellow dots arranged in a hexagonal configuration was used and detected by a Blob Detection algorithm [15, 34]. As most mobile devices are equipped with a *fixed*-focus lens for their front camera (resulting in a focus range of approx. 30-90cm), a 3D-printed mount was designed to attach a 4.0D (250mm) lens in front of the camera to ensure sharp photos of the reflected pattern in a range of approx. 13-20cm. The latest generation of phones includes *auto*-focus lenses for front cameras (e.g., iPhone 14 [5], Galaxy S22 [35]), which illustrates that in the future an additional lens will not be required anymore.

4 USER GUIDANCE TOWARDS THE 6 DOF POSITION

The following chapter will address the user guidance to fulfill and guide the user according to the three requirements for the desired near-eye position (1. forming an isosceles triangle, 2. correct distance and 3. correct user's gaze orientation). The user guidance is depicted in Figure 3.

4.1 Forming an isosceles triangle

For the eye to form an isosceles triangle with the pattern's center and the front camera, the designed solution was to mark the respective position in the video preview, similar to GoCheck Kids [18]. By using geometrical formulas, the y-offset between the upper edge of the video preview and the desired eye position on the screen was calculated. This position was highlighted as two *circular windows* on top of the preview video of the front camera.

4.2 Guiding the user to the optimal z-distance

To guide the user to the optimal z-distance, the *circular windows* introduced in the last section were extended to also represent the distance between the phone's screen and the user's eyes. Therefore, two colored rings (dynamic visual distance rings) were positioned around each circular window. The rings' sizes represent the distance between the user's eye and the optimal distance to capture the photo. Accordingly, their sizes increase and decrease depending on the distance. When the optimal z-distance is reached, the ring's size is the same as the circle. Simultaneously, the user's right iris starts

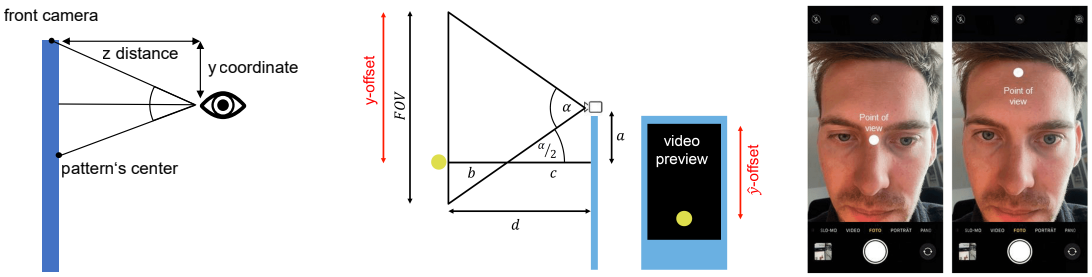


Figure 2: Requirements for the optimal eye's position: (a) 1. the eye forms an isosceles triangle with the pattern's center and the front camera. 2. Optimal z-distance. (b,c) 3. the user's gaze is orthogonal to the screen

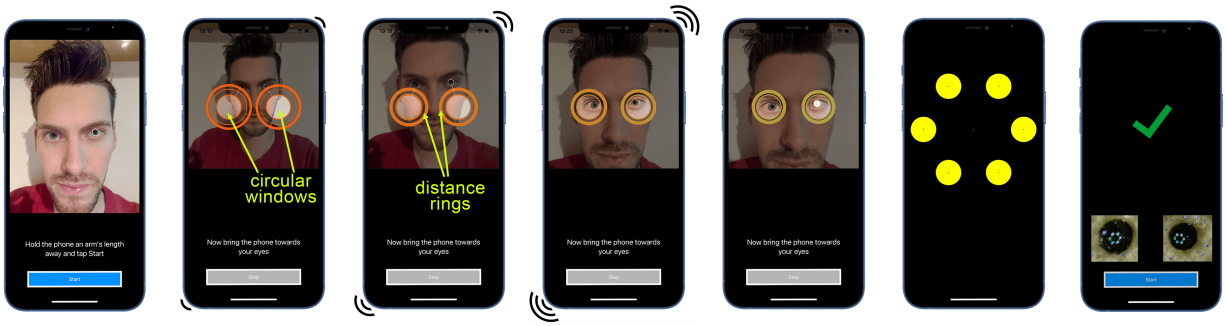


Figure 3: When starting the user guidance (a), the preview window is shifted to align for an isosceles triangle and displays circular windows for the eye positioning (b). Subsequently, animated dynamic distance rings as well as vibro-tactille feedback will provide visual and haptic guidance for positioning (c). Once correctly positioned (d), a visual stimulus will flicker to guide the gaze toward the correct orientation (e). The pattern is briefly displayed for photo capture (f) and the captured photos are displayed (g). Detailed description see Appendix C

to blink rapidly. This flickering stimulus draws the eye's attention and hence guides the eye's orientation so that the pattern will be reflected centrally. It also indicates that the photo will be captured within the next 600 milliseconds.
A 3D error range of 15mm radius around both eyes' optimal positions was introduced, in which the application will tolerate the user's eyes' positions.

4.3 Guiding the user's gaze orientation

Considering the condition of the shifting factor mentioned above (see section 3.1), the users have to look at a specific point on the upper screen to have their gaze orthogonally to the screen. Nevertheless, simultaneously the circular windows for the optimal position must be displayed in the lower part of the screen to ensure the isosceles triangle is formed. The solution to this problem was to move up the video preview and align the circular window with the desired point of view of the user. This concept neutralized the shifting factor as the video preview was moved up that the center of the video preview rests on the front camera's actual position. Therefore, a similar effect to a mirror (see section 3.1) Accordingly, this new concept assumes that the users look into their own eyes (on the screen) to position them inside the circular windows on the video preview. Hence, since the preview is shifted upwards, when they look at the windows, they are already looking straight at the screen (see Figure 2b).

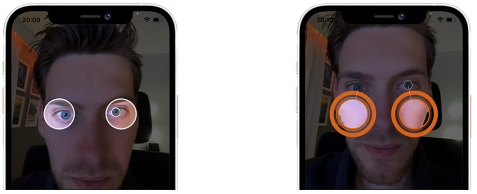
4.4 Haptic Feedback

As the users might not see the phone's screen in focus due to the eye's near accommodation point [16] or refractive errors, another modality besides visual user guidance was introduced: haptic feedback. Indeed, Ablavatski et al. [1] showed that it is possible to guide users with haptic feedback when additional visual information is available. For the haptic feedback, the Euclidean distances between each eye and the desired 6 DoF position were calculated continuously. Subsequently, these distances could be represented as vibration patterns: The closer the respective eye gets to its optimal location, the higher the vibration frequency becomes. Conversely, with increasing distance, the frequency decreases. The two individual distances of both eyes were averaged to reduce the haptic feedback to one dimension and simplify the user guidance. When both eyes reach the optimal positions, the phone's vibration stops.

This indicates that an acceptable position is reached, and the photo will be captured.

5 USER STUDY

A user study was conducted to compare the performance and preference of the four feedback modalities: "*baseline*," "*dynamic visual rings*," "*haptic*," and "*visual rings and haptic*." The baseline condition was inspired by current applications to position the eyes in the ophthalmologic context, such as GoCheck Kids [18], which uses eye icons to align the user's eyes accordingly. The four modalities represent the independent variable of the within-subject study design. The three dependent variables were (a) the time the user required to reach the 6 DoF position. (b) the raw task load index [20], and (c) the success rate, whether the reflected pattern was detected or if the user blinked during the capturing process.



(a) baseline **(b) dynamic visual rings**

(c) haptic **(d) visual rings and haptic**

Figure 4: The four conditions with varying feedback modalities. The baseline condition (a) was based on aligning the user's eyes visually with circular windows based on related work. This concept was extended with (b) dynamical visual rings highlighting distance via color and motion, (c) haptic feedback utilizing vibration intensity for distance and position, and (d) a combination of both extensions

5.1 Procedure

Before the study started, each of the n=24 participants signed the consent form. Subsequently, the four conditions (see Figure 4) were conducted in a counter-balanced order. Each condition started with a brief introduction (quick guide) on the phone (see appendix). To begin the user guidance and positioning task, participants were asked to hold the phone an arm's length away from their eyes. This procedure was performed twice per condition, resulting in eight data points for each participant. The time between the start of the user guidance and the final photo was recorded and whether the reflected pattern was detected successfully (no shaking or blinking.). After each condition, the participants were asked to complete the raw NASA TLX questionnaire (with its sub-scales mental demand, physical demand, temporal demand, performance, effort, and frustration) [20].

5.2 Results

The statistical summaries can be found in Appendix A and are depicted in Figure 5.

5.2.1 Task Load Index. For the accumulated result of the TLX, the mean of the six sub-scales was calculated. Before the statistical test, we checked the required assumptions of normally distributed data using the Shapiro-Wilk test [36]. Since the latter was not given, we used Friedman's ANOVA for repeated measures [17], resulting in $\chi^2(3)=10.81$, $p=0.012$ (see Figure 5a). The range of the TLX scores was between 0 and 100. The baseline condition had the highest task load (M=25.34, SD=15.31). The haptic feedback was rated with the lowest task load (M=17.74, SD=11.68). For the Post-hoc test, the Wilcoxon test [38] with Bonferroni correction [10] was used. Significances were found between the baseline condition and the haptic feedback ($p = 0.015$).

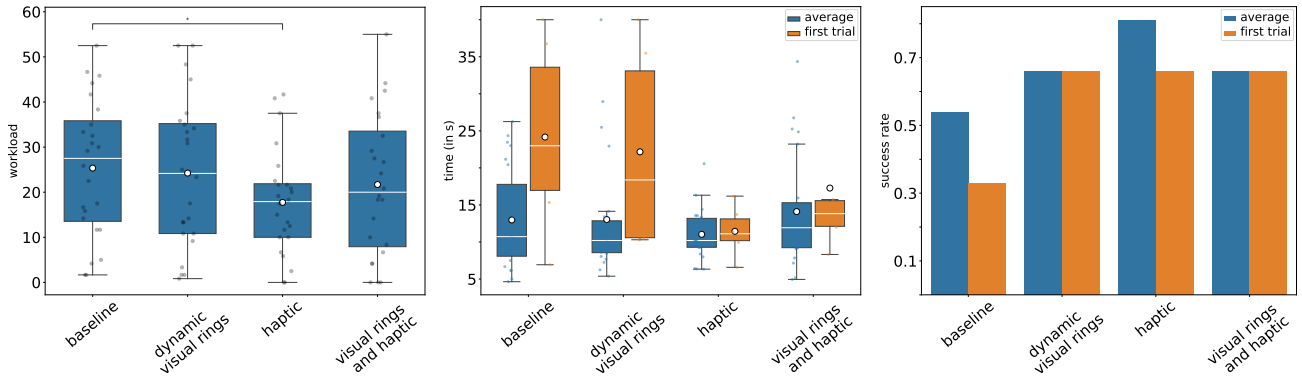
5.2.2 Success Rate. A successful measurement was defined as both eyes' reflected patterns being detectable. For the binary data Cochran's Q test [12]) was used. **Average**: $\chi^2(3)=0.320$, $p=0.852$ (see Figure 5c). The most successful condition was haptic (81.25%), and the least one was the baseline (54.16%). **First Trial**: $\chi^2(3)=2.66$, $p=0.26$ (see Figure 5c). The least successful condition was the baseline (33.33%), and the other conditions were equally successful (66.66%).

5.2.3 Required Time. Before the statistical test, we checked the required assumptions of normally distributed data using the Shapiro-Wilk test [36]. **Average**: Friedman's ANOVA repeated measures [17], resulting in $\chi^2(3)=2.7$, $p=0.432$ (see Figure 5b). The fastest condition was haptic feedback (M=11.03s, SD=3.29s), the slowest one was visual rings and haptic (M=14.10s, SD=7.44s). **First Trial**: The Kruskal Wallis test for independent measures [24], resulting in $\chi^2(3)=3.50$, $p=0.31$ (see Figure 5b). The fastest condition was haptic feedback (M=11.44s, SD=3.28s), and the slowest one was the baseline (M=24.16s, SD=12.98s).

6 DISCUSSION

This work introduces a novel user guidance for tele-ophthalmology that utilizes dynamic visual rings and haptic feedback to improve near-eye positioning tasks in smartphone-based eye examination applications. Hence, it addresses the limitation of existing tele-ophthalmology applications such as [8, 18, 31], which lack user guidance for the near-eye positioning task, particularly challenging for self-eye-examinations. A user study (n=24) was conducted to compare this guidance to a baseline condition similar to [18].

When considering the average time required to complete a measurement over all eight trials (twice per condition), no statistically significant differences were found among the different modalities. The median time for all modalities was approximately 11 seconds. This can be attributed to a strong learning effect due to the counter-balanced study design, where participants were able to position the phone faster in consecutive trials once they were familiar with the target position - independently from the modality. However, when exclusively examining the initial first trial between subjects,



(a) NASA TLX **(b) Required time** **(c) Success rate**

Figure 5: For the task load index, significant differences were found between the baseline and the haptic condition (p = 0.015). Further, the haptic component in the User Guidance, especially for the initial trial, yielded the trend to be faster than the conditions without haptic modality. The highest success rate was achieved for the haptic feedback in the averaged trials, implicating sufficient preparation for the photo to be captured.

it was observed that haptic feedback and the combination of dynamic visual rings and haptic feedback trends in faster positioning than the conditions without haptic feedback (baseline and dynamic visual rings). Nonetheless, even for the initial trial, no statistical differences were found. Regardless, this suggests that the haptic component of the user guidance provided sufficient feedback to enable users to quickly position the phone near the eye for independent measurements.

For the success rate, there is a trend (no significance) that the haptic feedback yielded the most successfully detected reflected patterns, while the baseline feedback yielded the lowest. This implies that the user guidance not only guides users to the correct position but also prepares them for the photo to be taken by ensuring they hold the phone steady and open both eyes at the right time.

In contrast, the NASA TLX results showed significant differences between haptic and baseline feedback, indicating that the visual feedback may have been overwhelming for users and could have led to confusion. Participants also indicated a preference for the haptic modality in comments based on a difficulty to understand the dynamic visual rings.

This highlights that haptic feedback in the form of phone vibrations subconsciously enriches the application with distance information, making it more intuitive and easy to understand, thus helping users to determine the position near the eye more effectively.

It is of importance to note that due to the close proximity of the phone to the eyes (15cm), the user's visual perception is impaired by the eyes' near focus point and potential refractive errors. As a result, the visual feedback was not clearly visible in these conditions. Therefore, haptic feedback showed to be adequate in providing additional cues for reaching the near-eye position.

In the future, we will investigate the effectiveness of *different* non-visual guidance methods for near-eye positioning tasks, such as other haptic feedback patterns. Further, it would be beneficial to conduct studies with larger sample sizes and diverse populations to understand the generalizability of the findings. Additionally, we want to investigate the method's applicability to different groups of visual impairments.

7 CONCLUSION

This work introduced a novel user-guided smartphone application for tele-ophthalmology, which utilizes dynamic visual rings and haptic feedback to guide the user to a specific 6 DoF near-eye position to capture a close-up photo of the eyes. Results from a user study (n=24) showed that the haptic modality was faster in initial trials, had a higher success rate, and resulted in less workload than the other modalities. This may be due to the close proximity of the phone to the eye, where the user's visual perception is limited, making haptic feedback an effective method of providing distance cues. Therefore, haptic feedback should be considered a useful modality for near-eye positioning tasks in future mobile self-eye-examination applications.

REFERENCES

[1] Artsiom Ablavatski, Andrey Vakunov, Ivan Grishchenko, Karthik Raveendran, and Matsvei Zhdanovich. 2020. Real-time Pupil Tracking from Monocular Video for Digital Puppetry. *CoRR* abs/2006.11341 (2020). arXiv:2006.11341 https://arxiv.org/abs/2006.11341

[2] JohnDavis Akkara and Anju Kuriakose. 2018. Innovative smartphone apps for ophthalmologists. *Kerala Journal of Ophthalmology* 30, 2 (2018), 138. https://doi.org/10.4103/kjo.kjo_68_18

[3] John D. Akkara and Anju Kuriakose. 2020. Commentary: Gamifying teleconsultation during COVID-19 lockdown. *Indian journal of ophthalmology* 68, 6 (2020), 1013–1014. https://doi.org/10.4103/ijo.IJO_1495_20

[4] John Davis Akkara and Anju Kuriakose. 2020. Commentary: Teleophthalmology and electronic medical records: Weighing the pros and cons of unavoidable progress. *Indian journal of ophthalmology* 68, 2 (2020), 367–368. https://doi.org/10.4103/ijo.ijo_2082_19

[5] Apple. 2023. *iPhone 14 - Technical Specifications.* Apple. Retrieved Jan 10, 2023 from https://www.apple.com/iphone-14/specs/

[6] J. C. Arbeláez, Roberto Viganò, and Gilberto Osorio-Gómez. 2019. Haptic Augmented Reality (HapticAR) for assembly guidance. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 13, 2 (2019), 673–687. https://doi.org/10.1007/s12008-019-00532-3

[7] Lokeshwari Aruljyothi, Anuja Janakiraman, B. Malligarjun, and Balasundaram Manohar Babu. 2021. Smartphone applications in ophthalmology: A

quantitative analysis. *Indian journal of ophthalmology* 69, 3 (2021), 548–553. https://doi.org/10.4103/ijo.IJO_1480_20

[8] Colin Barry, Jessica de Souza, Yinan Xuan, Jason Holden, Eric Granholm, and Edward Jay Wang. 2022. At-Home Pupillometry Using Smartphone Facial Identification Cameras. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 235, 12 pages. https://doi.org/10.1145/3491102.3502493

[9] Silvia Bonfanti, Angelo Gargantini, and Andrea Vitali. 2015. A Mobile Application for the Stereoacuity Test. Springer, Cham, 315–326. https://doi.org/10.1007/978-3-319-21070-4{_}32

[10] Carlo Bonferroni. 1936. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze* 8 (1936), 3–62.

[11] Rupert R A Bourne, Jaimie D Steinmetz, Jost B Jonas, and Theo Vos. 2021. Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. *The Lancet. Global health* 9, 2 (2021), e130–e143. https://doi.org/10.1016/S2214-109X(20)30425-3

[12] W. G. Cochran. 1950. THE COMPARISON OF PERCENTAGES IN MATCHED SAMPLES. *Biometrika* 37, 3-4 (12 1950), 256–266. https://doi.org/10.1093/biomet/37.3-4.256

[13] EyeQue Corporation. 2021. *EyeQue - Vision Test From Home.* Retrieved Dec 1, 2022 from https://www.eyeque.com/

[14] James M. Coughlan, Brandon Biggs, Marc-Aurèle Rivière, and Huiying Shen. 2020. An Audio-Based 3D Spatial Guidance AR System for Blind Users. *Computers helping people with special needs : ... International Conference, ICCHP ... : proceedings. International Conference on Computers Helping People with Special Needs* 12376 (2020), 475–484. https://doi.org/10.1007/978-3-030-58796-3_55

[15] Michael B. Dillencourt, Hanan Samet, and Markku Tamminen. 1992. A General Approach to Connected-Component Labeling for Arbitrary Image Representations. *J. ACM* 39, 2 (apr 1992), 253–280. https://doi.org/10.1145/128749.128750

[16] Alexander Duane. 1922. Studies in Monocular and Binocular Accommodation with their Clinical Applications. *American Journal of Ophthalmology* 5, 11 (1922), 865–877. https://doi.org/10.1016/S0002-9394(22)90793-7

[17] Milton Friedman. 1937. The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *J. Amer. Statist. Assoc.* 32, 200 (1937), 675–701. https://doi.org/10.1080/01621459.1937.10503522

[18] GoCheck Kids. 2021. *GoCheck Kids - Pediatric Vision Screening Solution.* Retrieved Dec 1, 2022 from https://www.gocheckkids.com/

[19] Google. 2022. *MediaPipe Iris.* Retrieved Dec 16, 2022 from https://google.github.io/mediapipe/solutions/iris.html

[20] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[21] Hassan Hashemi, Akbar Fotouhi, Abbasali Yekta, Reza Pakzad, Hadi Ostadimoghaddam, and Mehdi Khabazkhoob. 2018. Global and regional estimates of prevalence of refractive errors: Systematic review and meta-analysis. *Journal of current ophthalmology* 30, 1 (2018), 3–22. https://doi.org/10.1016/j.joco.2017.08.009

[22] Rachel Hecht, Ming Li, Quirina M. B. de Ruiter, William F. Pritchard, Xiaobai Li, Venkatesh Krishnasamy, Wael Saad, John W. Karanian, and Bradford J. Wood. 2020. Smartphone Augmented Reality CT-Based Platform for Needle Insertion Guidance: A Phantom Study. *Cardiovascular and interventional radiology* 43, 5 (2020), 756–764. https://doi.org/10.1007/s00270-019-02403-6

[23] Asim Kichloo, Michael Albosta, Kirk Dettloff, Farah Wani, Zain El-Amir, Jagmeet Singh, Michael Aljadah, Raja Chandra Chakinala, Ashok Kumar Kanugula, Shantanu Solanki, and Savneek Chugh. 2020. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. *Family medicine and community health* 8, 3 (2020). https://doi.org/10.1136/fmch-2020-000530

[24] William H. Kruskal and W. Allen Wallis. 1952. Use of Ranks in One-Criterion Variance Analysis. *J. Amer. Statist. Assoc.* 47, 260 (1952), 583–621. https://doi.org/10.1080/01621459.1952.10483441

[25] Jean-François Lapointe, Heather Molyneaux, and Mohand Saïd Allili. 2020. A Literature Review of AR-Based Remote Guidance Tasks with User Studies. In *Virtual, Augmented and Mixed Reality. Industrial and Everyday Life Applications*, Jessie Y. C. Chen and Gino Fragomeni (Eds.). Lecture Notes in Computer Science, Vol. 12191. Springer International Publishing, Cham, 111–120. https://doi.org/10.1007/978-3-030-49698-2_8

[26] Neha Misra, Rohit C. Khanna, Asha L. Mettla, Srinivas Marmamula, Varsha M. Rathi, and Anthony V. Das. 2020. Role of teleophthalmology to manage anterior segment conditions in vision centres of south India: EyeSmart study-I. *Indian journal of ophthalmology* 68, 2 (2020), 362–367. https://doi.org/10.4103/ijo.IJO{_}991{_}19

[27] Vitor F. Pamplona, Ankit Mohan, Manuel M. Oliveira, and Ramesh Raskar. 2010. NETRA. *ACM Transactions on Graphics* 29, 4 (2010), 1–8. https://doi.org/10.1145/1778765.1778814

[28] PeekVision. 2018. *Peek Vision.* Retrieved Dec 1, 2022 from https://www.peekvision.org/en_GB/peek-solutions/peek-acuity/

[29] Warachaya Phanphruk, Yingna Liu, Katharine Morley, Jacqueline Gavin, Ankoor S. Shah, and David G. Hunter. 2019. Validation of StrabisPIX, a Mobile Application for Home Measurement of Ocular Alignment. *Translational vision science & technology* 8, 2 (2019), 9. https://doi.org/10.1167/tvst.8.2.9

[30] Amar Pujari. 2021. Smartphone Ophthalmoscopy: is there a place for it? *Clinical Ophthalmology (Auckland, NZ)* 15 (2021), 4333. https://doi.org/10.2147/OPTH.S243103

[31] Shrinivas Pundlik, Matteo Tomasi, Rui Liu, Kevin Houston, and Gang Luo. 2019. Development and Preliminary Evaluation of a Smartphone App for Measuring Eye Alignment. *Translational vision science & technology* 8, 1 (2019), 19. https://doi.org/10.1167/tvst.8.1.19

[32] Pavel Rojtberg. 2019. User Guidance for Interactive Camera Calibration. Springer, Cham, 268–276. https://doi.org/10.1007/978-3-030-21607-8_21

[33] Hillary K. Rono, Andrew Bastawrous, David Macleod, Emmanuel Wanjala, Gian Luca Di Tanna, Helen A. Weiss, and Matthew J. Burton. 2018. Smartphone-based screening for visual impairment in Kenyan school children: a cluster randomised controlled trial. *The Lancet. Global health* 6, 8 (2018), e924–e932. https://doi.org/10.1016/S2214-109X(18)30244-4

[34] H. Samet and M. Tamminen. 1988. Efficient component labeling of images of arbitrary dimension represented by linear bintrees. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 10, 4 (1988), 579–586. https://doi.org/10.1109/34.3918

[35] Samsung. 2023. *Specifications - Samsung Galaxy S22.* Retrieved Jan 10, 2023 from https://www.samsung.com/global/galaxy/galaxy-s22/specs/

[36] S. S. Shapiro and M. B. Wilk. 1965. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* 52, 3/4 (1965), 591–611. http://www.jstor.org/stable/2333709

[37] Mollianna Walker, Alba Duvall, Mackenzie Daniels, Mailynh Doan, Luke E. Edmondson, Edward W. Cheeseman, M. Edward Wilson, Rupal H. Trivedi, and Mae Millicent W. Peterseim. 2020. Effectiveness of the iPhone GoCheck Kids smartphone vision screener in detecting amblyopia risk factors. *Journal of AAPOS : the official publication of the American Association for Pediatric Ophthalmology and Strabismus* 24, 1 (2020), 16.e1–16.e5. https://doi.org/10.1016/j.jaapos.2019.10.007

[38] Frank Wilcoxon. 1945. Individual Comparisons by Ranking Methods. *Biometrics Bulletin* 1, 6 (1945), 80–83. http://www.jstor.org/stable/3001968

A DESCRIPTIVE DATA

Table 1: Table of scores

Variable	Conditions	n	Mean	Median	SD	rate
NASA TLX	baseline	24	25.34	27.5	15.31	
	dynamic visual rings	24	24.23	24.16	16.53	
	haptic	24	217.74	17.91	11.68	
	visual rings and haptic	24	21.70	20.00	15.85	
Required Time - average	baseline	24	12.97s	10.77s	6.77s	
	dynamic visual rings	24	13.03s	10.20s	8.22s	
	haptic	24	11.03s	10.20s	3.29s	
	visual rings and haptic	24	14.10s	11.92s	7.44s	
Required Time - first trial	baseline	24	24.16s	22.56s	12.98s	
	dynamic visual rings	24	22.17s	18.37s	13.53s	
	haptic	24	11.44s	11.10s	3.28s	
	visual rings and haptic	24	17.27s	13.85s	11.43s	
Success rate - average	baseline	24				0.54
	dynamic visual rings	24				0.66
	haptic	24				0.81
	visual rings and haptic	24				0.66
Success rate - first trial	baseline	24				0.33
	dynamic visual rings	24				0.66
	haptic	24				0.66
	visual rings and haptic	24				0.66

B PROGRAM FLOW

(1) **Modality Choice.** The first step is to choose the desired modality of the application: "haptic," "visual," "visual and haptic," or "baseline."

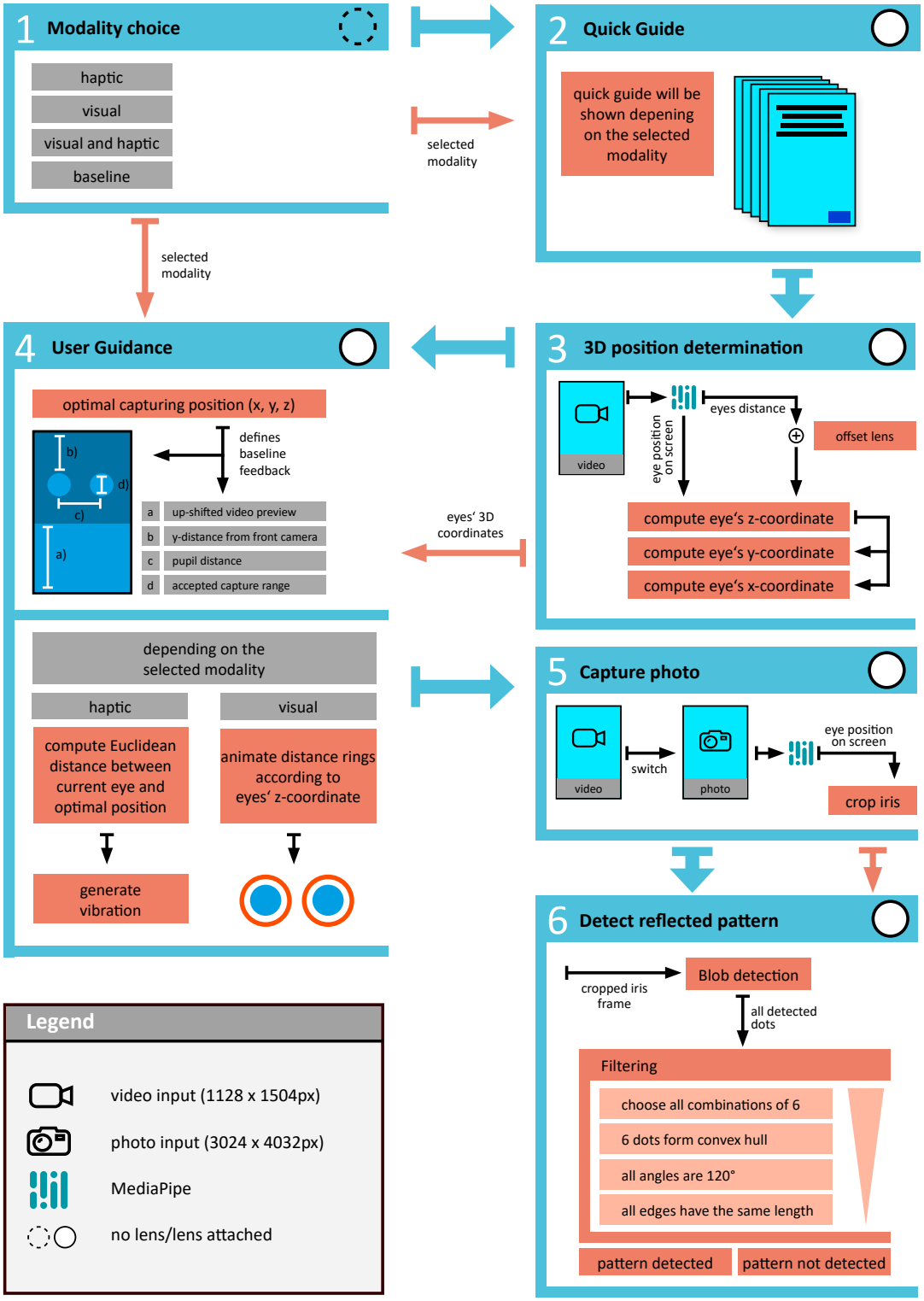


Figure 6: Program flow of the application

(2) **Quick Guide.** The Quick Guide shortly explains the upcoming interaction steps to the user. Depending on the selected modality, its content changes. Furthermore, the user is asked to attach the additional lens to the iPhone at the end of the Quick Guide.
(3) **3D position determination.** From the video input, the eye's distance is constantly determined through MediaPipe. Additionally, the magnification factor of the lens is added to the raw distance. From the eye's position on the screen and the corrected distance, the eye's z-coordinate is deduced. Subsequently, using the latter z-coordinate and the eye's screen position, x- and y-coordinates are computed.
(4) **User Guidance.** The User guidance is divided into two parts.
  (a) The chosen optimal capturing position determines the baseline feedback. The latter defines the circular window's position on the screen and the up-shifting factor of the video preview.
  (b) Depending on the selected modality, the distance rings' animation is determined by the eye's z-coordinate for the visual feedback. For the haptic feedback, the Euclidean distance between the current eye position and the optimal capturing position is computed and averaged over both eyes. This distance determines the vibration frequency.
(5) **Capture photo.** When the optimal eye's position is reached, the application automatically switches the input from the low-resolution video to the high-resolution photo mode to take the picture. Using MediaPipe, the iris of both eyes is cropped out of the photo.
(6) **Detect reflected pattern.** A Blob detection algorithm is applied to the cropped iris frame. Therefore, all possible reflected dots will be located. However, since more than the desired dots are detected, a filtering algorithm must ensure that only the correct dots will be considered. As the dots must be arranged hexagonally, all combinations of 6 are chosen and examined individually. First, it is checked if the six dots form a convex hull. Second, the angles between the six dots are examined. If they are about 120°, this combination will be analyzed in the final step to check if all edges have a similar length. When a combination of six dots remains after the filtering steps, it is most likely the reflected dot pattern.

C USER GUIDANCE

As soon as the users finish the quick guide (Figure 7), the application asks them to hold the phone an arm's length away from their face and press "Start" to activate the user guidance (Figure 8a). Subsequently, the video preview moves upwards. Additionally, two circular windows appear as an overlay to the preview, which becomes darker. In contrast, the two windows' transparency remains the same (Figure 8b). Furthermore, other graphic elements appear on the screen: two thin white lines connect the windows' rims with the user's eyes, indicating to move them inside the windows. Moreover, the right iris is marked with a white ring (Figure 8b). More noticeable, however, are the two colored distance rings that appear around the circular windows (Figure 8c). The rings' sizes enlarge and shrink depending on the user's distance to the phone. Thus, the closer the user gets, the smaller the rings become, and

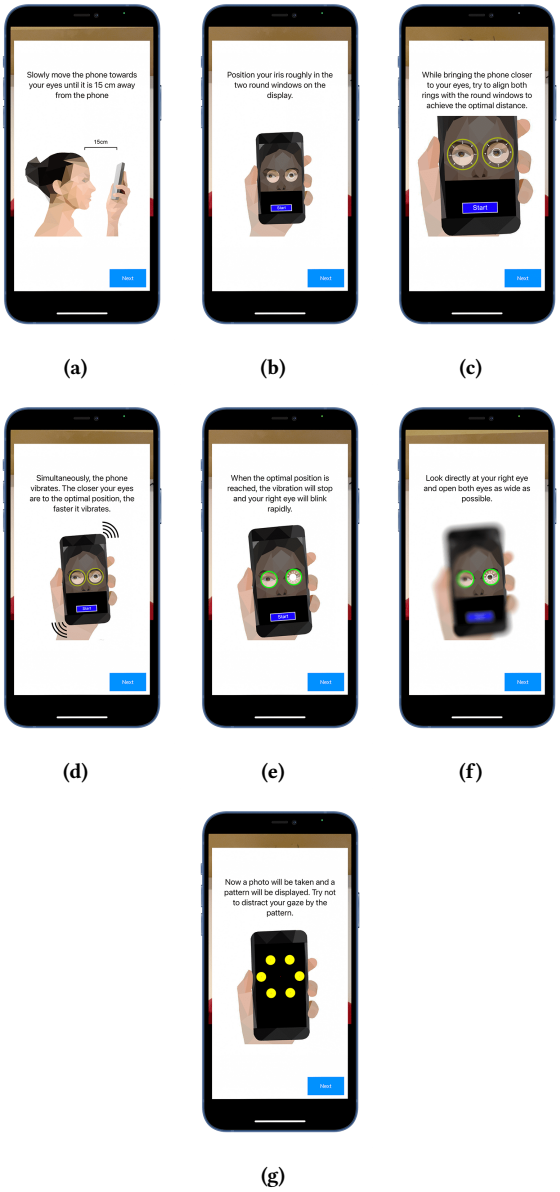


Figure 7: Quick Guide for the condition visual rings and haptic

vice versa. Additionally, their color changes accordingly: is the user far away, their color is reddish. In contrast, when the user gets closer, the color slowly changes to green (Figure 8b - Figure 8d). Similarly happens with the white iris ring mentioned above. As soon as the user moves the eyes inside the circular windows, this ring's color changes from white to the same color as the distance ring's (Figure 8d). Simultaneously to the visual feedback, the user guidance also provides haptic feedback: The phone vibrates depending on the user's distance. The further away from the user, the less frequently the phone vibrates. The closer the user gets, the higher the vibration frequency (Figure 8b - Figure 8d). Once

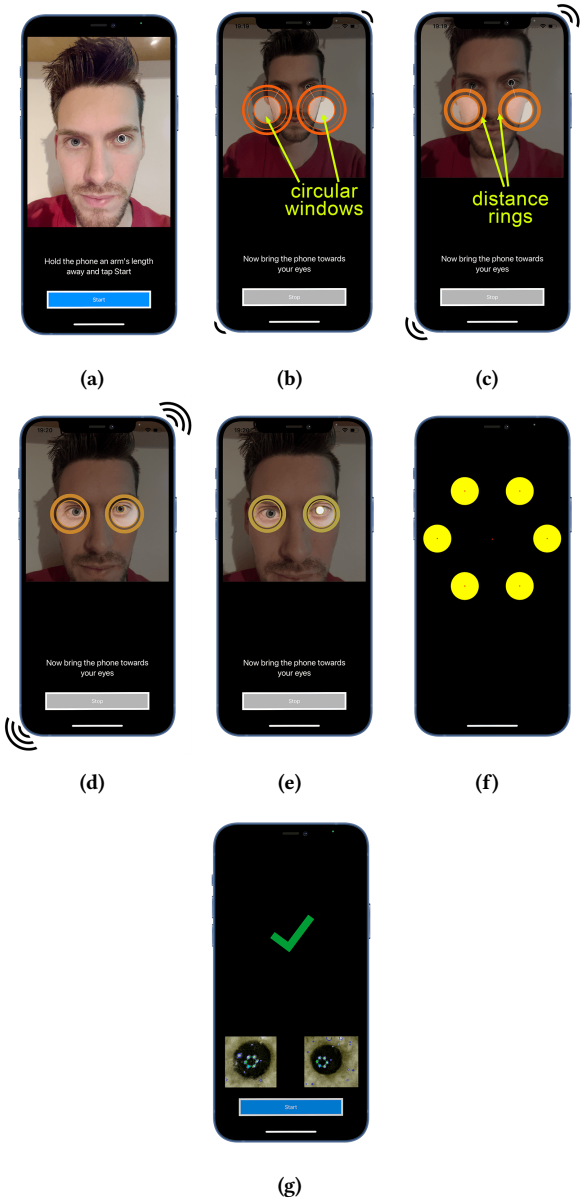


Figure 8: User Guidance for the condition visual rings and haptic feedback

the user reaches the optimal position, the user unconsciously gets informed by several indicators: the two distance rings are aligned with the circular windows' rims, and the vibration stops. Additionally, the right iris ring starts to blink rapidly for 0.6s (Figure 8e). Subsequently, the pattern gets displayed, and the photo is captured (Figure 8f). Finally, the detected pattern's reflection is displayed for both the right and left eye (Figure 8g).