

Visuomotor Influence of Attached Robotic Neck Augmentation

Lichao Shen

Keio University

Graudate School of Media Design

shenlc@kmd.keio.ac.jp

MHD Yamen Saraiji

Keio University

Graudate School of Media Design

yamen@kmd.keio.ac.jp

Kai Kunze

Keio University

Graudate School of Media Design

kai@kmd.keio.ac.jp

Kouta Minamizawa

Keio University

Graudate School of Media Design

kouta@kmd.keio.ac.jp

Roshan L Peiris

Rochester Institute of Technology

School of Information

roshan.peiris@rit.edu

ABSTRACT

The combination of eye and head movements plays a major part in our visual process. The neck provides mobility for the head motion and also limits the range of visual motion in space. In this paper, a robotic neck augmentation system is designed to surmount the physical limitations of the neck. It applies in essential a visuomotor modification to the vision-neck relationship. We conducted two experiments to measure and verify the performance of the neck alternation. The multiple test results indicate the system has a positive effect to augment the motions of vision. Specifically, the robotic neck can enlarge the range of neck motion to 200%, and influence the response motion, by overall 22% less in time and 28% faster in speed.

CCS CONCEPTS

- Human-centered computing → Displays and imagers.

KEYWORDS

visuomotor coordination, visual expansion, spatial perception, robotic neck, human augmentation, wearable

ACM Reference Format:

Lichao Shen, MHD Yamen Saraiji, Kai Kunze, Kouta Minamizawa, and Roshan L Peiris. 2020. Visuomotor Influence of Attached Robotic Neck Augmentation. In *Symposium on Spatial User Interaction (SUI '20), October 31-November 1, 2020, Virtual Event, Canada*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3385959.3418460>

1 INTRODUCTION

Our visual system plays an important role in perceiving our surroundings, not only imagery information but also spatial information, and the type of reactions we perform are based on the complex sensory information we perceive. Our eyes capture a limited stereo visual field of view (FoV) of approximately 120° horizontally and 120° vertically. This field can be significantly expanded to cover 360°

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SUI '20, October 31-November 1, 2020, Virtual Event, Canada

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-7943-4/20/10...\$15.00

<https://doi.org/10.1145/3385959.3418460>



Figure 1: Unconstrained Neck System Overview:
the artificial neck system substitutes human neck motion limits, resulting a wider range of motion.

by combining our visuomotor system that consists of the eye-gaze, neck, torso, and overall body posture movements. The speed to observe a specific object in our surroundings depends on a variety of factors including the number of the involved visuomotor elements listed prior. If an object locates within the visual field, a simple eye movement is sufficient to see the object, however, if it locates behind us, then a full-body posture change is required to be performed requiring more time to look at it. Many scenarios demand rapid response and wide vision, thus achieving a wider and faster visuomotor system can be beneficial for various application scenarios, such as surveillance, sports, and entertainment.

To augment the spatial range of visual system, either 1) the field of view of the eyes can be expanded using optical methods (such as wide FoV lenses), or 2) the limited motion range of the body and the neck can be altered to accommodate a new visual range. In this paper, we developed a visuomotor alteration system named *Unconstrained Neck* [27], using an artificial neck (Figure 1) that attempts to expand the motor range as in the second approach.

Unconstrained Neck surpass the mechanical limits of the human neck and expands the spatial range of the visual system by remapping the original neck motion and by mechatronic modification to visuomotor coordination. In our user study, we demonstrate that our prototype speeds up locating objects surrounding the user (covering 360°) without the need to move the entire body. We also

report the efficiency of using our hands to point at the surrounding objects while using the system. Thus, our main contributions are:

- We present the design and implementation of a neck augmentation system that allows us to redefine the mapping of neck motion covering 360°, while maintaining visual consistency for the eyes.
- We report a user study evaluating the effectiveness of the proposed system in terms of the usability as well as the time to locate targets in space around the user.
- A second user study is presented that explores the augmentation of the movement of the hands with the operational parameters of the system.

2 RELATED WORK

In an attempt to expand the spatial range of visual perception, the most direct approach is to extend the original visual field. The visual field is the observable extent at any given moment. This visual field can be expanded either by getting the image from a camera or a lens with a wider field of view (FoV), or by stacking the images from multiple camera array to construct wider FoV [21, 22, 24]. Such approaches of direct alteration of the visual system typically requires training and adaptation by the user, which is a widespread phenomenon in the visual system. Adaptation occurs on multiple time-scales, ranging from seconds to hours, according to the type of the change involved and its duration [14, 17].

A conventional solution is to use wide-angle camera, for example a 360-degree panoramic camera. In the project FlyVIZ, J. Arduin et al. uses a panoramic image acquisition system to obtain omni-directional vision and then squeezes the whole image into the visual field of head mounted displays (HMD) [1]. The system achieves a visual field 360° horizontal and 80° vertical. However, using such approaches requires the user to perform long periods of training to adapt to using it. In addition, it can typically eliminate the stereoscopic vision and thus depth perception when using it. Works such as [3, 13, 23] proposed the use of wide lenses cameras with video see-through HMD to expand the perceived visual field. Similarly, wide angle lenses introduces large visual distortion which as a result reduces depth perception and skewed direction estimation compared with binocular field [2, 15].

E. Schoop et al. proposed a environment awareness system [26], which utilized deep neural network to locate and recognize objects around the user through a head-worn panoramic camera. Yet, it converted the visual information to semantic audio content, and hence the original format was not preserved. Gustafson et al. [10] and Gruenfeld et al. [9] developed methods to view off-screen objects: using a wedge, and a HMD version of it respectively. In addition, Kuhl et. al. and Li et. al. explored the effects of minification and artificially reducing the field of view on the perception of distance judgements [16, 19, 20].

K. Fan et al. developed the SpiderVision, a handset extends human visual field to mainly augment user's awareness in the back [7]. The system enables user to focus on front view by the video-see-through HMD, but it blends the back-view video in only when the system detects dynamic visual change. It keeps monitoring the surrounding environment and sets an intellectual trigger to activate the augmented content. This system requires part of the visual field

to display the back vision, which inevitably reduces the quality of normal vision: the image become blurred or narrowed. Its visual augmentation based on peripheral monitoring and the trigger for the visual blend is passive for the user so that the reliability of automatic detection limits the whole system performance.

To address the previous limitations, we propose to use an alternative neck representation, *Unconstrained Neck*, which enhances the range of motion our biological neck can achieve. This system maintains the visual properties of our eyes (FoV and binocular vision) in order to reduce the adaptation time, and to sustain depth perception when using this system.

3 DESIGN CONSIDERATIONS: VISION-NECK RELATIONSHIP

As seen from the related work, an established body of work exist for visual system augmentation using optical solutions. Given the advantages and drawbacks of such methods, we adopted a different approach by addressing the visuomotor system limitations. In this section, we describe the design considerations and implementation of the proposed system.

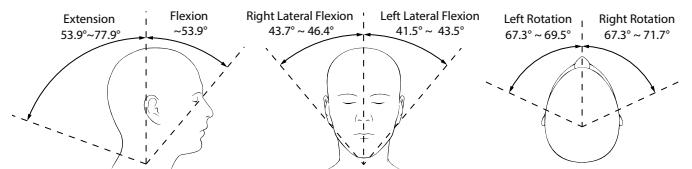


Figure 2: Estimates of Normal Cervical Active Range of Motion (ROM), at the age of 30 years.

The visual sense contributes to spatial perception, and thus involve visuomotor coordination (i.e., human vision and motion works together). Human eyes have relatively narrow field of view compared with animals such as bird or rabbit. The monocular (one-eyed) visual field of a healthy human measures approximately 100° temporally, 60° nasally, 60° superiorly and 75° inferiorly of each eye [25, 29], and the cyclopean vision (total) is approximately 200° wide and 135° tall, with a region of binocular (two-eyed) overlap of approximately 120° wide [4, 5], which is crucial for stereo vision and depth perception. Hence, our eyes need assistance from the rest of body, to expand the spatial range of eyesight, due to the narrowness of visual field.

The neck influences the visual sense though it is not a visual organ. It plays an important role in scanning our surroundings. Scanning is defined as the action shifting the visual focus in a larger scale, and with the essential purpose to expand the visual field. The motion of one or some body parts including eyeball, neck, and torso, are likely to involve in scanning the environment. When scanning, the eye gaze can be considered as the easiest to manipulate, but the ocularmotor range (OMR) of the eye gaze is restricted compared to the torso and the neck, typically not exceeding 40° to 45° [8]. The torso (trunk) is able to deform (twist and bend), and thus have larger ROM. But the torso bears the weight of the whole upper body, and its motions involve several groups of muscles, so the torso is most

difficult to action. The neck is at a balanced point: it only bears and drives the head; its ROM is relatively wide. The neck can perform six types of movements, and the cervical active range of motion (ROM) varies with the age [30] and the approximate value is as shown in Figure 2. The head motion is directly driven by the neck, and therefore is limited by the range of the neck movements.

Existing research has shown that the behavior of the neck is dependent on the target location when scanning [18]. The eye-only range (EOR), where no head motion tends to happen, is approx. $46.7^\circ (\pm 23^\circ)$ [28]. The customary ocular motor range (COMR), where the eye movements tend to happen with a probability of 90%, is approx. $57.4^\circ (\pm 29^\circ)$ [18]. As shown in Table 1, the type of motions which is very likely to happen, depends on the different final target position relative to the neutral position. The eye movements only have a small probability to occur if the intended final position is outside the COMR, that is, the human is likely to use only neck movements when the intend rotation is relatively large.

Table 1: Different Motion Types Depend on Target Position

Target Position (T)	$T < 23^\circ$	$23^\circ < T < 29^\circ$	$29^\circ < T$
Motion Type	Eyes Only	Eye and Head	Head Only

Thus, to reduce the chain of visuomotor motion, we are focusing on augmenting the neck with our system. The proposed artificial neck would alter the ROM of human neck by remapping each axis to a new motion range as shown in Figure 6. Here our goal is to enhance the visual field by reducing the effort and time required to reach targets beyond our typical visual field.

4 UNCONSTRAINED NECK SYSTEM

4.1 Implementation

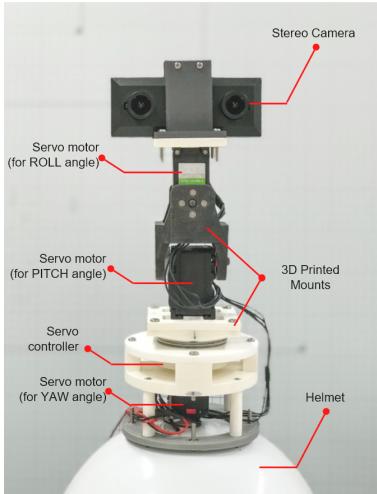


Figure 3: Unconstrained Neck system consists of 3 servo motors that control the yaw, pitch, and roll angles, and a stereo camera mounted atop a helmet

In this paper we present our prototype *Unconstrained Neck* (Figure 3). The prototype contains a 3-DoF (degree of freedom) robotic neck holding a stereo camera. The robotic neck consists of 3 Dongbu Herkulex DRS-0201 servo motors, 3D-printed plastic supports / connectors, and other necessary mechanical components. Each motor drives the rotation about one axis (i.e., yaw, pitch, and roll). A Spark-Fun Pro Micro-controller is embedded to control 3 servos by wire. This external robotic neck consists of 3 rotational kinematic pairs; thus, it has 3 degrees of freedom. An Ovrvision Pro camera set capturing stereo videos by dual lenses and sensors, streams the videos to an Oculus DK2 / Rift HMD (head-mounted display) worn by the user, as depicted in Figure 1. The distance between the two lateral placed lens is 65 mm, close to the human interpupillary distance [6], which mean value is approximately 63 mm. The camera FoV resembles to normal human vision by adjusting the magnification so that it can create a see-through immersive experience. A laptop running Unity 3D software maintains wired connections to the camera, the micro controller and the HMD. The spatial position of the camera mounted at the endpoint of the robotic neck is controlled by a customized developed Unity 3D software and is able to operate about the 3-axis rotations: yaw, pitch, and roll. The resulting motion of rotation is thus a combination of the movements of human neck and robotic neck. The robotic neck is attached to helmet to allow the user to wear the prototype with ease. Due to the height of the robotic neck, the location of the camera set has a vertical offset of approx. 450 mm from the user's eye when wearing this prototype. The overall system stands at approx. 400 mm when worn (from the top of the head to the top of the camera) and weighs approx. 500 g with the helmet, and 300 g without the helmet.

The *Unconstrained Neck* system can be described as a substitute neck system, that is, it controls the orientation of vision¹ with an alterable and programmable mechanism. The input of the system is the original vision with its inherent spatial information, and the output is the modified vision with new spatial information (Figure 4 (left)). The input vision is typically sourced from human eyes on the head, so its orientation is equivalent to the head orientation. The head rotation is driven by the human neck, and thus the input vision and the human neck can be studied by tracking the head. The modified vision output is sourced from the camera, so its orientation is equivalent to the camera orientation. The camera rotation is driven by the *Unconstrained Neck* system (the human neck plus the robotic neck), and thus the output vision and the whole system can be studied by tracking the camera motion. In brief, head orientation as input and (camera) vision orientation as output.

4.2 Mapping

The internal mechanism are defined in the controlling software (Figure 4 (right)), mainly focusing on the orientations of vision, while the imagery content of vision remains unmodified. The spatial orientation of a rigid body can be parametrized by three independent coordinates, and MEMS gyroscopes usually provide three values of yaw, pitch and roll. All the orientation and components are represented with respect to the body frame. Thus orientation mappings

¹In this paper, the orientation of vision, eyesight, or visual orientation refers to the orientation of the outward normal vector of human visual field, i.e., the gaze direction when looking straight forward.

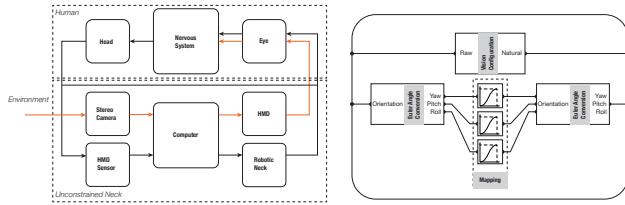


Figure 4: (left) System I/O and (right) Software Mechanism

about 3 axes can be built respectively, between the orientation of output vision (derived from camera position, and driven by the neck substitution system) and the corresponding orientation of input vision (derived from head position, and driven by the human neck). A rotation is the change of orientation, so an angular displacement can be defined as the difference in orientation, or angular position, from the origin to the current position. If the origin of the reference orientation was set as the neutral position when facing straight forward, the general mappings can be denoted by Function Group 1:

$$\begin{cases} \text{roll: } \phi_O = f_1(\phi_I) \\ \text{pitch: } \theta_O = f_2(\theta_I) \\ \text{yaw: } \psi_O = f_3(\psi_I) \end{cases} \quad \begin{cases} \text{roll: } \phi_O = 2\phi_I \\ \text{pitch: } \theta_O = 2\theta_I \\ \text{yaw: } \psi_O = 2\psi_I \end{cases}$$

where the subscript of I denotes the input and O denotes the output. The ϕ_I, θ_I, ψ_I are the input orientations of vision, identical to the motion displacements of human head in value, which the human neck acts; the ϕ_O, θ_O, ψ_O are the output orientations, identical to the resulting motion displacements of camera vision in value, which the *Unconstrained Neck* system actuates. By $\phi_O = \phi_I, \theta_O = \theta_I, \psi_O = \psi_I$, it defines the innate original mapping (i.e. no augmentation).

The prime feature of *Unconstrained Neck* is the controllable and programmable vision-neck relationship. One possible beneficial application is visual expansion. When the system is worn by a user, the maximum range of motion of the output vision is -180° to 180° about each axis (limited by the servo motors and the mechanical structure), which is almost twice the range of human neck motion. Therefore, it can be utilized to amplify or enlarge the neck's range of motion by twice or more, and the human eyesight can hence achieve omni-directional observation or simply 360° vision (i.e., a scanning range of 360° tall and 360° wide). For the scope of this paper, we will concentrate on this application scenario. One of the most typical and basic augmented mapping can be *Linear ×2* (Function Group 2, Mapping Graph Figure 6): $\phi_O = 2\phi_I, \theta_O = 2\theta_I, \psi_O = 2\psi_I$. For example, as shown in Figure 5 (a) and (b), when the user's head (input) is at a yaw angle of Ψ , the resulting camera vision (output) is at a yaw angle equivalent to 2Ψ due to the combined movement of the human neck and robotic neck. In this case, the eyesight can fully cover a task space of S^2 , with only neck movements.

In this paper, the neck mapping refers to mapping of the whole system: head rotation as input, system (camera) vision as output; the hand mapping was defined similarly: physical hand position as input, virtual hand position as output. All the following experiments were conducted with the augmented mapping of *Linear ×2* applied, that is, in plain languages, the rotational angles the camera (output) travelled is always twice ($\times 2$) as large as the rotational angles the head (input) travelled, in both Study 1 and Study 2.

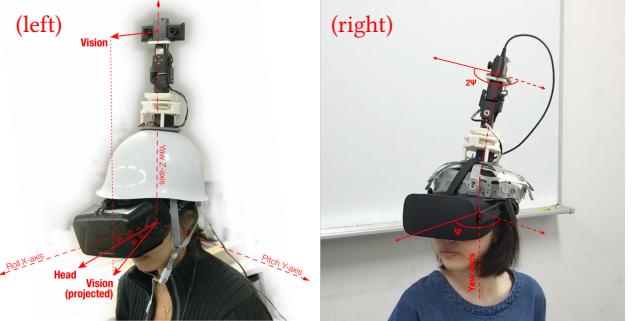


Figure 5: (left) Parameter and axis definitions; (right) the system in use with augmentation of *Linear ×2* mapping ($[Camera Rotation] = 2 \times [Head Rotation]$)

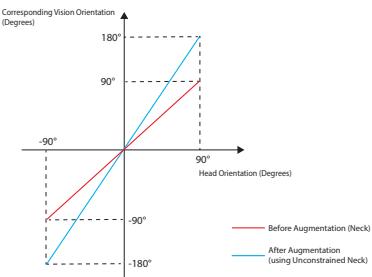


Figure 6: Graph of Original/Augmented Mapping (*Linear ×2*) for each axis, (x: Input, Orientation of Human Head, y: Output, Corresponding Orientation of (Camera) Vision)

5 STUDY 1: USER EVALUATION OF UNCONSTRAINED NECK SYSTEM

The main goal of Study 1 was to evaluate whether the users can scan their surrounding space more efficiently using the *Unconstrained Neck*, and to observe its influences on the users. The hypothesis was that the system could reduce the response time and boost the response speed for scanning motion with neck involved.

5.1 Method

We took the scanning as a representation of visuomotor practice, and then measure the perform of the response reaction to evaluate the effectiveness of the system. We presented several virtual targets around the subject's spatial range as shown in Figure 7. As the main task, the subject was required to locate the presented target as quickly as possible. The targets were generated and presented in augmented virtual environment through the study software referenced from an external global point in space, to measure the head rotations and present the target location. Virtual target was designed to be distinct and distinguishable from the surrounding environment using color of yellow (Figure 8). As depicted in Figure 7, we presented a total of 20 targets distributed evenly around the origin of the subject's head. The spatial location of the target was denoted in the form of (latitude, longitude): 0 latitude is the transverse plane (horizontal), and 0 longitude is the sagittal plane (vertical); the original point of the coordinate is (0,0) at the right



Figure 7: Experiment Setup: (top) the subject sat in the experimental field, and (bottom) the distribution of targets along three different latitudes / tilts: $-30^\circ, 0^\circ, +30^\circ$

front of human head which is also the intersecting line of transverse plane and sagittal plane;

In Study 1, we presented 20 virtual targets in total, of which the locations are $(0, 30), (0, 60), (0, 90), (0, 120), (0, 150), (0, 180), (\pm 30, 0), (\pm 30, 30), (\pm 30, 60), (\pm 30, 90), (\pm 30, 120), (\pm 30, 150)$ and $(\pm 30, 180)$, as per Figure 7, where lat. $+30^\circ$ refers to the upper plane, and lat. -30° refers to the lower plane. The target was a virtual sphere with a radius of 50 mm, and located at a distance of 500 mm from the origin subject's head.

5.1.1 Study Design. The study was a within-subject design and consisted of 2 main conditions (independent variables of the study), as shown in Table 2. The condition when the subject wore the system but with robotic neck deactivated is Condition 1 (denoted as C1, with original neck mapping). Under C1, the scanning motion of camera was only as a result of the subjects' human neck rotations; The condition when the subject wore the system with robotic activated is Condition 2 (denoted as C2, with augmented neck mapping). In C2, the scanning motion of camera was as a result of the rotation of the subject's human neck rotations plus the robotic neck rotations as described in the previous section, and hence the augmentation effect was on (enhanced by two times, $Linear \times 2$). As for the measurements (dependent variables), the response time to locate and select the target was recorded. Lastly, the participant answered the simulator sickness questionnaire [12] after completing each condition. Each target was repeated 4 times; thus, each subject faced a total of 160 task trails ($20 \text{ targets} \times 4 \text{ repetitions} \times 2 \text{ conditions}$).

5.1.2 Subjects. 14 participants volunteered to participate in study 1 (9 females and 5 males, age range: 22 to 35 years, mean: 25.6, SD: 3.59). The subjects were recruited from the university where the experiment took place, and consented for the study. All subjects had prior experience with using HMD and virtual reality.

5.1.3 Apparatus. This study utilized the *Unconstrained Neck* system as the main apparatus of the study. Plus, two remote controllers

Table 2: Condition Specification: Setup and Mapping / Motion (\dagger robotic neck deactivated; \ddagger robotic neck activated.)

	Abbr.	Setup	Mapping / Resulting Motion	
			System Drive	(Camera) Vision
Study 1	C1	Human Neck †	Original	n/a
	C2	Human Neck + Robotic Neck ‡	Augmented (x2)	n/a
Study 2	C3	Human Neck †	Original	Original
	C4	Human Neck + Robotic Neck ‡	Augmented (x2)	Original
	C5	Human Neck + Robotic Neck ‡	Augmented (x2)	Augmented (x2)

(model: Oculus Touch) were introduced to track the hand locations and optimize the experimental interface. The experimental setup was as shown in the Figure 7(top).

5.1.4 Procedure. After collecting the informed consent and the biographical data, each subject was briefed on the study. In addition, for safety reasons, the subjects were all queried through a questionnaire to confirm that their physical condition fulfilled all required criteria. The criteria were as follows: (1) no current neck pain, (2) no history of any neck medication, (3) no current eyesight illness, (4) no severe virtual reality related sickness, (5) unambiguous communication with the experimenter. Information of virtual reality and motion sickness was introduced and explained for the subjects prior to conducting the study.

During the study, the subject sat on a rigid chair, wearing the robotic neck and HMD; the subject held two remote controllers with action keys. The subject could only move and twist his/her upper body, neck or torso to scan. At the beginning of task, the subject placed his/her neck and torso at the neutral position, and kept aiming at the origin point. Before the recording started, the subject had time to practice and understand the task. Once the task began, a 'Find' command was displayed on the HMD screen instructing the subject to initiate localizing the target and finish aiming. After each condition, a break of 5 minutes was provided for compensation. Specific procedures of the conditions are as follows. The experiment took approximately 45 minutes per participant.

5.1.5 Task Procedure. As shown in Figure 8, the sight (aiming mark) consisted of a cross and a circle in the center of visual field, and the virtual target was a yellow 3D sphere anchored in external coordinate. During the experiment, the participant was required to move the field of view until the virtual target entered the circle mark. A text indicator showed the progress of the experiment and the status of current target. 'Find' indicated the presence of a new target, while in the meantime ' $<<$ ' or ' $>>$ ' below indicated that the target located in the left or right hemisphere. 'Back' meant the target had been destroyed, and the subject should resume his/her position to the original position.

The subject pressed any action key to start the procedure. Once the action key was pressed, a target appeared and the text displayed 'Find'. Then the subject should move the field of view, to search for the target and aim at it. To aim was to overlap the central circle with the target. Once the aiming was finished, the subject pressed the action key, and instantly the existing target would disappear and the text displayed 'Back'. Then the subject could resume his/her position of head and torso to the origin, and should aim at the

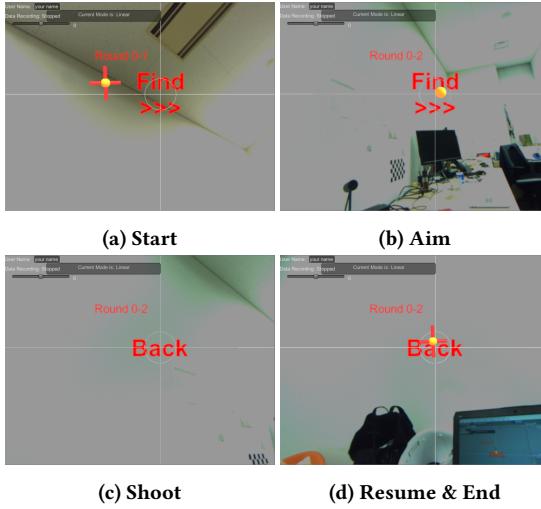


Figure 8: Procedure of Aim at the Virtual Targets: the subject (a) start at the initial point, was asked to find the target, (b) searched for the target and aimed, (c) shoot by pressing a button, then was asked to return to the origin, and (d) ended one target, then prepared for a next target

original point. When ready, the subject could press the action key again, and instantly a new target would appear. Then the subject continued to next target. For this part, a subject was presented with 4 repetitions of the 20 targets, at a distance of 500 mm. This study presented conditions of C1 and C2 in a randomized order. The subject answered the simulator sickness questionnaire (SSQ) at the end of each condition.

5.2 Results

Study 1 consisted of 20 virtual targets that were distributed in the space around the user, at varying angular distances from the initial starting location. Therefore, to analyse the results, we initially observed the results of the conditions (C1 and C2) as a whole using the following definitions. Every participant completed 4 repetitions (R) under each condition (C), and 20 different targets (θ) under each repetition (R). “Target Time” (t_θ) is the time costed to complete aim at one specific target from the initial location, that is, the interval between the moment one target appears and the moment the next target appears. “Condition time” (t_C) is defined as the total time costed to complete all 4 repetitions for all 20 targets and is defined by the Equation 3. In addition, “Condition Speed” is defined as the average head rotational speed of the whole condition, and is similarly calculated by using the Equation 4 ($\Delta\theta$ is the angular displacement for each target from the initial location).

$$t_C = \sum_{r=1}^4 \sum_{n=1}^{20} t_{\theta_{n,r}} \quad (3)$$

$$\bar{\omega}_C = \frac{\sum_{r=1}^4 \sum_{n=1}^{20} \Delta\theta_{n,r}}{t_C} \quad (4)$$

The overall results of C1 and C2 from Study 1 are depicted in Figure 9. Here, the Condition Time for C1 is 219.72s (SD: 57.997s) and

for C2 is 171.99s (SD: 40.762s). It indicates the participants were able to aim at the virtual targets significantly faster in C2 (augmented neck mapping) than in C1 (original neck mapping). Similarly, the Condition Speed for C1 is 36.732°/s and for C2 is 46.344°/s. Results are further analyzed using a paired t-test with $\alpha = 0.05$. The t-test results of condition time, $t(14) = 3.51, p = .00384$, and condition speed, $t(14) = -4.27, p < .001$, show statistically significant differences between two conditions, this is, time decreased and speed increased.

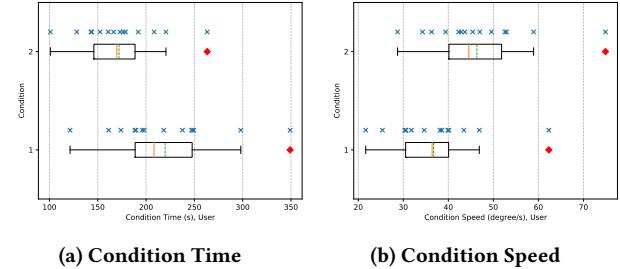


Figure 9: (C1,C2) Box Plot: Mean Condition Time / Speed for 14 Subjects. Mean of 4 Repetitions. (blue cross: scattered point of each sample, red diamond: outlier, orange solid line: median, green dash line: mean)

Figure 10 indicates the time and speed values for each target. Here, the “Target Time” is t_θ as defined previously. The “Target Speed” was defined as $\bar{\omega}_\theta = \frac{\Delta\theta}{t_\theta}$. For clarity, the Figure 10 results are further presented separately by 3 latitude values. The results are further analyzed using a two-way repeated measure ANOVA. A significant main effect is found across condition (C1,C2) and across target, from both Target Time, $p_{\text{condition}} < .001, p_{\text{target}} < .001$, and Target Speed, $p_{\text{condition}} < .001, p_{\text{target}} < .001$. Post-hoc pairwise comparison with Bonferroni and FDR (False Discovery Rate) correction between each pairs of the targets (C1,C2) reveal the unevenness of the effect as shown in Table 3. A overall efficacy ratio index (R) is calculated to quantify the degree of effectiveness, $R_{\text{time}, C2/C1} = t_{C2}/t_{C1} = 0.782 = 78\%$ and $R_{\text{speed}, C2/C1} = \bar{\omega}_{C2}/\bar{\omega}_{C1} = t_{C1}/t_{C2} = 1.28 = 128\%$.

Table 3: (C1,C2) P-values of Posthoc Pairwise T-test for 20 Targets / 3 Latitudes / 7 Longitudes (* naïve significance, ** FDR significance, * Bonferroni significance)**

Lon.	(a) Target Time				(b) Target Speed			
	-30	0	30	Average	-30	0	30	Average
0	0.199	n/a	4.07 E - 2	6.48 E - 2	0	7.61 E - 2	n/a	3.80 E - 2
30	0.351	0.597	0.623	0.334	30	4.47 E - 2	0.518	0.835
60	0.447	0.721	9.62 E - 2	0.864	60	0.931	0.708	0.112
90	0.506	0.500	0.154	0.213	90	0.369	0.336	0.293
120	0.537	7.75 E - 2	0.124	0.119	120	0.162	3.81 E - 2	6.07 E - 2
150	8.06 E - 3	9.23 E - 3	3.10 E - 3	2.20 E - 3	150	1.13 E - 2	1.15 E - 3	1.81 E - 3
180	1.16 E - 2	9.51 E - 3	6.99 E - 4	4.35 E - 4	180	6.30 E - 3	3.98 E - 3	2.95 E - 4
Avg.	8.81 E - 3	1.16 E - 2	4.13 E - 3	3.84 E - 3	Avg.	1.59 E - 2	3.18 E - 3	3.99 E - 4
	***	***	***	***		***	***	***

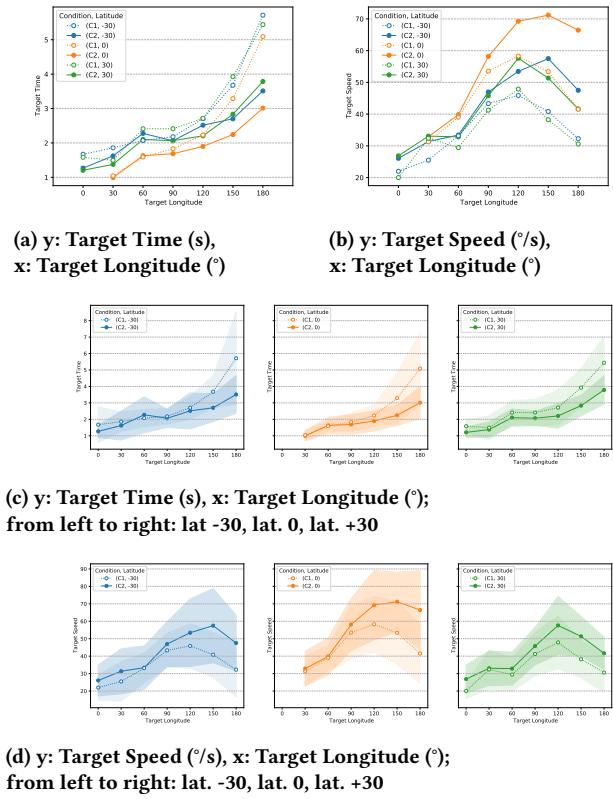


Figure 10: (C1,C2) Mean Target Time / Speed for 20 Targets. Mean of 4 Repetitions × 14 Subjects. (color area: the error bar with standard deviation)

Table 4: (C1,C2) Sickness Questionnaire for 10 Subjects

	SSQ Score		SSQ T-Test		Neck Score		Neck T-Test	
	Mean	SD	Stats.	P-value	Mean	SD	Stats.	P-value
C1	430	338	-1.08	0.309	3.50	3.06	2.86	0.0187*
C2	518	332			1.70	2.54		

We conducted surveys after C1 and C2 using the simulator sickness questionnaire (SSQ). The data summary is shown in Table 4, and the higher a score was, the severer a sickness was. The result of t-test shows no statistically significant differences of sickness index existed between two conditions, but significant difference of neck load existed. Thus, the modification of vision-neck relationship, tended to cause no significant VR / simulator sickness extra. However, as for the influence on neck, the score declined significantly, so the augmentation was likely to reduce and relief the neck load, despite of the extra weight and height brought by the system.

5.3 Discussion

The hypothesis was proved. These results clearly indicated that in C2 (augmented neck mapping) where the *Unconstrained Neck*

system provided enhanced neck movements resulted in faster target acquisition times (t_θ) than in C1 (original neck mapping), from the condition perspective. Further individual target-wise analysis showed those target locations which contributed to the significant differences, concentrated around longitude 150° and 180°. As for the target speed, its result showed more significant locations than the target time. This notion can be considered as an interesting observation from Study 1. We identify that targets around longitude 150° to 180° are those that are typically beyond our visual field. Hence, those targets would typically require neck and torso movements to look at them, and in C2 were aimed significantly faster compared to C1 where there was no enhancement from the system. However, in C2 the rest targets which were not in this range although were helped from the system did not yield significant effects and were achieved with similar effort and performance to C1.

In summary, in Study 1, all the results implied the existence of positive influence was of high probability. This system and its augmentation, generally reduced the response time and correspondingly boosted the speed of vision shifting, while in the meantime relief the neck motion and expand the visual field. The augmentation effect was uneven across targets, and it applied significant benefit mostly for the motion in rear hemisphere.

6 STUDY 2: USER EVALUATION OF HAND INVOLVEMENT

The main goal of this study was to observe the influences of hand-eye coordination in the surrounding space when using the *Unconstrained Neck* system. The hypothesis was that the neck augmentation would similarly have a positive influence on the performance of tasks involved with the hand-eye coordination.

6.1 Method

We conducted this study similar to Study 1 with a few adjustments. Similar to previous study, we presented several virtual targets around the subject's spatial range as presented in Figure 7. As the main task, the subject was required to locate and then additionally touch the presented target as quickly as possible using the two virtual hands. The virtual hand was represented by a virtual sphere with a radius of 50 mm, same size as the virtual target, and it was positioned by the spatial position data collecting from the remote controller. We only presented targets at the 0° latitude, and at a closer distance of 300 mm so that the participant could reach the targets with minimal effort.

In Study 2, we presented 6 virtual targets in total, and the locations are (0, 30), (0, 60), (0, 90), (0, 120), (0, 150), (0, 180); The target was a virtual sphere with a radius of 50 mm, and located at a distance of 300 mm from the origin subject's head.

6.1.1 Study Design. The study was a within-subject design and consisted of 3 main conditions (independent variables). In the system, the resulting neck motion outputs two levels, original neck mapping and augmented neck mapping, and the resulting virtual hand motion outputs two levels, original hand mapping and augmented hand mapping, so the combination conditions are defined as C3 (original neck mapping, original hand mapping) C4 (augmented neck mapping, original hand mapping) and C5 (augmented neck mapping, augmented hand mapping), as shown in Table 2. Similar

to the previous study, under C3, the robotic neck was deactivated, and under C4 and C5, the robotic neck was activated and hence the vision was enhanced. In C3, C4 and C5, hands were presented in the visual system as described in Section 6.1. As for the measurements (dependent variables), the response time to locate and select the target was recorded. Lastly, the participant answered the NASA Task Load Index questionnaire [11]. Each target was repeated 4 times; thus, each subject faced a total of 72 task trials ($6 \text{ targets} \times 4 \text{ repetitions} \times 3 \text{ conditions}$).

6.1.2 Subjects. 6 participants volunteered to participate in study 2 (6 males, age range: 24 to 36 years, mean: 29.8, SD: 4.96).

6.1.3 Apparatus. The experiment utilized the same *Unconstrained Neck* system as the Study 1. Thus, the experimental setup was the same as indicated in Figure 7(top).

6.1.4 Procedure. The subject selection process was similar to the Study 1 and the initial procedure (pre-study questionnaires, etc) is the same as described in Study 1 in Section 5.1.4.

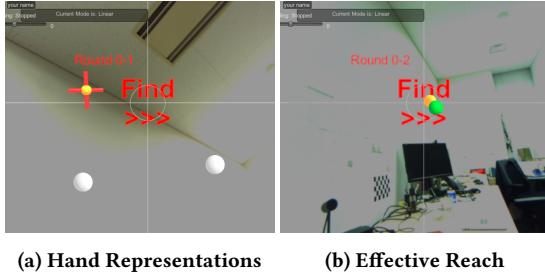


Figure 11: (a) The locations of virtual hands in AR were represented by two white spheres. (b) The subject was required to overlap the spheres so to “reach” the target, then the hand sphere would turn green

6.1.5 Task Procedure. The task followed the same procedure as shown in Figure 8 and described in Section 5.1.5. Besides, two virtual hands (white spheres) represented the hand location. The specific locating mechanism could be set to normal or augmented, i.e., the movements of virtual hands (white spheres), relative to the body, were twice the movements of real hands following the augmentation principle of the neck. In addition, after aiming at the target, the subject was asked to control either hand to reach the target. Once the target was reached, the white sphere turned green, to confirm an effective target acquisition, as shown in Figure 11. Then the subject could press a key to destroy the target. The rest of the procedures was similar to Study 1. In this part, a subject performed 4 repetitions of the 6 targets in the horizontal plane. This study presented the C3, C4, and C5 in a randomized order. At the end of each condition, the participant answered the NASA Task Load Index (TLX) questionnaire.

6.2 Results

The overall results of the C3, C4, and C5 from Study 2 are depicted in Figure 12. Here, the Condition Time for C3 is 48.623s (SD: 5.8320s), for C4 is 50.459s (SD: 5.6717s) and for C5 is 40.210s (SD: 4.8023s).

Results are further analyzed using a Repeated Measures ANOVA reveals significant effects between the conditions, $F(2, 15) = 7.89, p = .00454$. Post-hoc pairwise tests reveals significant effects between comparison groups of (C3,C5), $t(6) = 6.88, p = .00801$, and (C4,C5), $t(6) = 4.26, p < .001$, except (C3,C4), $t(6) = -1.02, p = .354$. Similarly, Condition Speed for C3 is $52.395^\circ/\text{s}$, for C4 is $50.485^\circ/\text{s}$ and, for C5 is $63.418^\circ/\text{s}$.

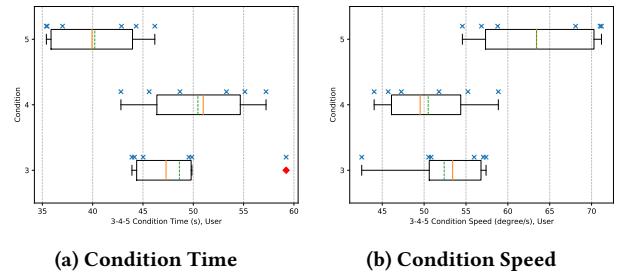


Figure 12: (C3,C4,C5) Box Plot: Average Condition Time / Speed for 6 Subjects. Mean of 4 Repetitions. (blue cross: scattered point of each sample, red diamond: outlier, orange solid line: median, green dash line: mean)

Figure 13 indicates the time and speed values for each target. Here, the “Target Time” and “Target Speed” are defined similar to Study 1, when the participant acquires the target. For clarity, the Figure 13 results are further presented separately in three comparison groups. The results are further analyzed using a series of two-way repeated measure ANOVA tests. A significant main effect was found across conditions and across targets, from both Target Time and Target Speed. Post-hoc pairwise comparison with Bonferroni and FDR correction between each pairs of the targets revealed the uneven effects as shown in Table 5. The result indicates that in the comparison group of (C3,C4), every location did not have significant differences, but in the group of (C3,C5) and (C4,C5), significant differences occurred in the latitude range from 120 to 180. This is similar to the observation of (C1,C2) in Study 1, where the rear hemisphere yielded significant effects. As for the efficacy index, $R_{time, C4/C3} = 1.038$, $R_{time, C5/C3} = 0.827$, $R_{time, C5/C4} = 0.797$.

We conducted surveys using the NASA TLX questionnaire after each condition. The result in Table 6 showed that the task load of C3 and C4 seems not to be statically different, and of C5 were likely to have the smallest load. One-way ANOVA of 3 conditions and individual pairwise t-test showed similar results.

6.3 Discussion

Hypothesis was partly proved when synchronizing the hand augmentation with the vision augmentation. Observing the result in Figures 12 and 13, in C5 (augmented neck mapping, augmented hand mapping) the participants responded best when the virtual hands with augmentation followed the same neck augmentation mapping (*Linear* \times 2 in this scenario). However, when comparing C3 (original neck mapping, original hand mapping) with C4 (augmented neck mapping, original hand mapping), the differences

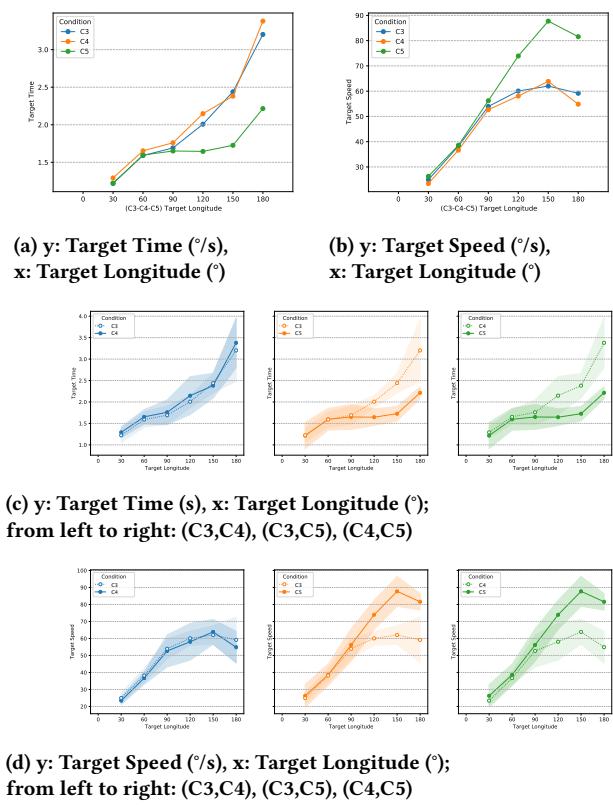


Figure 13: (C3,C4,C5) Mean Target Speed / Time for 6 Targets. Mean of 4 Repetitions × 6 Subjects. (color area: the error-bar with standard deviation)

Table 5: (C3,C4,C5) P-Values of Posthoc Pairwise T-test for 6 Targets × 3 Comparison Groups (* naïve significance, ** FDR significance, * Bonferroni significance)**

(a) Target Speed			(b) Target Time				
Lon.	Comparison Group		Lon.	Comparison Group			
	C3,C4	C3,C5		C3,C4	C3,C5		
30	0.415	0.568	0.305	30	0.482	0.948	0.533
60	0.452	0.858	0.390	60	0.470	0.943	0.473
90	0.799	0.609	0.591	90	0.657	0.786	0.551
120	0.720	2.16 E -2	6.20 E -3	120	0.546	1.80 E -2	1.61 E -2
150	0.633	1.90 E -3	2.90 E -3	150	0.699	2.02 E -3	4.29 E -3
180	0.258	7.75 E -3	4.26 E -4	180	0.389	2.14 E -2	3.55 E -3
Avg.	0.374	1.37 E -2	1.23 E -3	Avg.	0.370	6.82 E -3	7.98 E -4
	***	***		***	***	***	

between 2 conditions were not statistically significant at any levels, which mean the involvement of normal hands without augmentation nullified the influence brought by the augmented neck. The reason may be the proprioception conflict between neck and hand. Comparing C3 with C5, and C4 with C5, the intervention of augmented hand improved the performance, and then reached a

Table 6: (C3,C4,C5) Task Load Index for 6 Subjects

	Task Load Index		TLX T-Test		
	Mean	SD	H_0	T-Stats.	P-value
C3	26.0	6.42	$\mu_{C3} = \mu_{C4}$	0.250	0.812
C4	25.2	10.4	$\mu_{C4} = \mu_{C5}$	3.42	0.0188*
C5	12.8	6.71	$\mu_{C3} = \mu_{C5}$	3.32	0.0209*

significant promotion. When the hand and neck was augmented synchronous, the overall effect seems to return. Yet, we are unable to decide exactly which part, neck or hand, played a prior role in this compounded phenomenon and mixed situation.

7 DISCUSSION AND CONCLUSION

In conclusion, the *Unconstrained Neck* was proven to enlarge the spatial range of vision, and augmented the response motion. To be specific, the enlarged range was 200% as wide as before (100% wider); the speed of augmented response motion was 128% as fast as before (28% faster); the time of augmented response motion was 78% as long as before (22% less). One possible factor that inhibited the response time from decreasing greater to an ideal efficacy ratio index of 0.5 (50% less) may be the human visual system, which itself had a upper limit of tolerance of vision shifting. Shifting too fast may exceed the visual system capability to receive information and lead to discomfort. The effect was uneven: target location and user influenced the efficiency of the augmentation. The side effect of neck load and VR sickness, resulting in fatigue and dizziness, may be responsible for the sporadic outlier of negative influence. But on average, VR sickness did not increase, and neck load decreased.

For the hand-involved tasks, however, with a normal virtual hand, the overall influence of the *Unconstrained Neck* was weakened to an insignificant level; by an augmented virtual hand, the overall positive influence was recalled. Possible constructions are put as follows: first, the hand always had a higher priority to the head / neck in the spatial perception process of body; second, the disorder caused by a conflict between two spatial reference objects, diminished the influence.

Feedback from participants spotlighted the general experience was smooth and immersive, and none felt severe sickness. Usually users could accommodate to this new visual augmentation experience within a short time. The anisotropy of experience existed: the users commented that extension and flexion were likely to cause more dizziness than any other motion. We suggested it was catalysed by the complete flip of sky-ground reference (i.e., everything was upside down) which totally conflicted with the common sense. Different patterns of mapping were tried: linear and nonlinear. The users were able to recognize and be aware of the difference between two mappings, but usually felt more uncomfortable about the nonlinear mapping. That may because the nonlinear mapping produced relatively unpredictable behaviors, and the nonlinear visuomotor integration was alien to human.

The *Unconstrained Neck* efficiently contrives vision augmentation by modifying the vision-neck relationship and altering the visuomotor coordination as a bridge. The vision-neck, vision-neck-hand, hand-eye relationship are typical examples of visuomotor

coordination or integration, and furthermore belong to sensorimotor coordination, the ability to operate different parts of the somatic body together compatibly and effectively, both input and output. The coordination is flexible but fragile, and could be interrupted by the alternation from other human sense, function, or organ. Moreover, the body schema is inferred to possess the potential capacity of being modified. The human body, essentially the brain is adaptable enough to the alteration and modification. Rebuilding mapping of body schema could create a profound impact on human augmentation. Human behavior, mechanism and structure are complicated, we need to take extra care of the overall body coordination when applying alternation to a part. In some degree, coordination was more essential than augmentation.

7.1 Future Potential Applications

As future applications, we are considering a similar technique to be used to explore more novel vision-motor relationship, for instance, concept of transplanting eyes to anywhere. Besides, the technique of alternating sensorimotor coordination and its further implementation could have beneficial application in research, entertainment, sport and other scenarios. For example, researchers can make use of this programmable system in medical, bioengineering research, or assistance for rehabilitation therapy. When driving, riding and outdoors exploring, the system can be worn to gain wider field of view and quicker response action, for security and performance. As for the entertainment, the technique can bring unusual and unreal experience of vision-head relationship that is suitable for game world of exceeding the human sensation.

ACKNOWLEDGMENTS

Many thanks to all the volunteers participating in our user test, for their neck rotations of more than 6000 times. This research is supported by Embodied Media Project of Keio University.

REFERENCES

- [1] Jérôme Ardouin, Anatole Lécuyer, Maud Marchal, Clément Riant, and Eric Marchand. 2012. FlyVIZ: a novel display device to provide humans with 360 vision by coupling catadioptric camera with hmd. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology*. ACM, 41–44.
- [2] Th Brandt, Jo Dichgans, and E Koenig. 1973. Differential effects of central versus peripheral vision on egocentric and exocentric motion perception. *Experimental brain research* 16, 5 (1973), 476–491.
- [3] C Bill Chen. 2002. Wide field of view, wide spectral band off-axis helmet-mounted display optical design. In *International Optical Design Conference 2002*, Vol. 4832. International Society for Optics and Photonics, 61–67.
- [4] Robert Cubridge. 2005. *Visual fields*. Elsevier Health Sciences. 2 pages.
- [5] Gislin Dagnelie. 2011. *Visual prosthetics: physiology, bioengineering, rehabilitation*. Springer Science & Business Media.
- [6] Neil A Dodgson et al. 2004. Variation and extrema of human interpupillary distance. In *Proceedings of SPIE*, Vol. 5291. 36–46.
- [7] Kevin Fan, Jochen Huber, Suranga Nanayakkara, and Masahiko Inami. 2014. SpiderVision: extending the human field of view for augmented awareness. In *Proceedings of the 5th Augmented Human International Conference*. ACM, 49.
- [8] Edward G Freedman. 2008. Coordination of the eyes and head during visual orienting. *Experimental brain research* 190, 4 (2008), 369.
- [9] Uwe Gruenefeld, Tim Claudius Stratmann, Lars Prädel, and Wilko Heuten. 2018. MonoculAR: A Radial Light Display to Point Towards Out-of-view Objects on Augmented Reality Devices. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (Barcelona, Spain) (*MobileHCI '18*). ACM, New York, NY, USA, 16–22. <https://doi.org/10.1145/3236112.3236115>
- [10] Sean Gustafson, Patrick Baudisch, Carl Gutwin, and Pourang Irani. 2008. Wedge: Clutter-free Visualization of Off-screen Locations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) (*CHI '08*). ACM, New York, NY, USA, 787–796. <https://doi.org/10.1145/1357054.1357179>
- [11] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. 52 (1988), 139 – 183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [12] Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and M G Lilienthal. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology* 3, 3 (1993), 203–220.
- [13] Kiyoishi Kiyokawa. 2007. A wide field-of-view head mounted projective display using hyperbolic half-silvered mirrors. In *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*. IEEE, 207–210.
- [14] Adam Kohn. 2007. Visual adaptation: physiology, mechanisms, and functional benefits. *Journal of neurophysiology* 97, 5 (2007), 3155–3164.
- [15] Ernst Kruijff, J Edward Swan, and Steven Feiner. 2010. Perceptual issues in augmented reality revisited. In *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on*. IEEE, 3–12.
- [16] Scott A. Kuhl, William B. Thompson, and Sarah H. Creem-Regehr. 2009. HMD Calibration and Its Effects on Distance Judgments. *ACM Trans. Appl. Percept.* 6, 3, Article 19 (Sept. 2009), 20 pages. <https://doi.org/10.1145/1577755.1577762>
- [17] MiYoung Kwon, Gordon E Legge, Fang Fang, Allen MY Cheong, and Sheng He. 2009. Adaptive changes in visual cortex following prolonged contrast reduction. *Journal of vision* 9, 2 (2009), 20–20.
- [18] R John Leigh and David S Zee. 2015. *The neurology of eye movements*. Vol. 90. Oxford University Press, USA.
- [19] Bochao Li, Anthony Nordman, James Walker, and Scott A. Kuhl. 2016. The Effects of Artificially Reduced Field of View and Peripheral Frame Stimulation on Distance Judgments in HMDs. In *Proceedings of the ACM Symposium on Applied Perception* (Anaheim, California) (*SAP '16*). Association for Computing Machinery, New York, NY, USA, 53–56. <https://doi.org/10.1145/2931002.2931013>
- [20] Bochao Li, Ruimin Zhang, Anthony Nordman, and Scott A. Kuhl. 2015. The Effects of Minification and Display Field of View on Distance Judgments in Real and HMD-Based Environments. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception* (Tübingen, Germany) (*SAP '15*). Association for Computing Machinery, New York, NY, USA, 55–58. <https://doi.org/10.1145/2804408.2804427>
- [21] Feng Liang, Stevanus Kevin, Holger Baldauf, Kai Kunze, and Yun Suen Pai. 2020. OmniView: An Exploratory Study of 360 Degree Vision using Dynamic Distortion based on Direction of Interest. In *Proceedings of the Augmented Humans International Conference*. 1–10.
- [22] Feng Liang, Stevanus Kevin, Kai Kunze, and Yun Suen Pai. 2019. PanoFlex: Adaptive panoramic vision to accommodate 360° Field-of-view for humans. In *25th ACM Symposium on Virtual Reality Software and Technology*. 1–2.
- [23] Hajime Nagahara, Yasushi Yagi, and Masahiko Yachida. 2006. A wide-field-of-view catadioptrical head-mounted display. *Electronics and Communications in Japan (Part II: Electronics)* 89, 9 (2006), 33–43.
- [24] Yun Suen Pai, Benjamin I Outram, Benjamin Tag, Megumi Isogai, Daisuke Ochi, and Kai Kunze. 2017. GazeSphere: navigating 360-degree-video environments in VR using head rotation and eye gaze. In *ACM SIGGRAPH 2017 Posters*. 1–2.
- [25] Peter J Savino and Helen V Danesh-Meyer. 2012. *Color Atlas and Synopsis of Clinical Ophthalmology—Wills Eye Institute—Neuro-Ophthalmology*. Lippincott Williams & Wilkins.
- [26] Eldon Schoop, James Smith, and Björn Hartmann. 2018. HindSight: Enhancing Spatial Awareness by Sonifying Detected Objects in Real-Time 360-Degree Video. (2018).
- [27] Lichao Shen, Mhd Yamen Saraji, Kai Kunze, and Kouta Minamizawa. 2018. Unconstrained Neck: Omnidirectional Observation from an Extra Robotic Neck. In *Proceedings of the 9th Augmented Human International Conference* (Seoul, Republic of Korea) (*AH '18*). Association for Computing Machinery, New York, NY, USA, Article 38, 2 pages. <https://doi.org/10.1145/3174910.3174955>
- [28] John S Stahl. 2001. Without Title. *Experimental brain research* 136, 2 (2001), 200–210.
- [29] William B Trattler, Peter K Kaiser, and Neil J Friedman. 2012. *Review of Ophthalmology E-Book: Expert Consult-Online and Print*. Elsevier Health Sciences.
- [30] James W Youdas, Tom R Garrett, Vera J Suman, Connie L Bogard, Horace O Hallman, and James R Carey. 1992. Normal range of motion of the cervical spine: an initial goniometric study. *Physical therapy* 72, 11 (1992), 770–780.