



Evaluation of Spatial Directional Guidance Using Cheek Haptic Stimulation in a Virtual Environment

Fumihiko Nakamura^{1*}, Adrien Verhulst², Kuniharu Sakurada¹, Masaaki Fukuoka¹ and Maki Sugimoto¹

¹ Graduate School of Science and Technology, Keio University, Kanagawa, Japan, ² Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan

OPEN ACCESS

Edited by:

Thomas Kosch,
Darmstadt University of Technology,
Germany

Reviewed by:

Zhen Gang Xiao,
Simon Fraser University, Canada
Yomna Abdelrahman,
Munich University of the Federal
Armed Forces, Germany
David Steeven Villa Salazar,
Ludwig Maximilian University of
Munich, Germany

*Correspondence:

Fumihiko Nakamura
f.nakamura@imlab.ics.keio.ac.jp

Specialty section:

This article was submitted to
Mobile and Ubiquitous Computing,
a section of the journal
Frontiers in Computer Science

Received: 30 June 2021

Accepted: 08 April 2022

Published: 12 May 2022

Citation:

Nakamura F, Verhulst A, Sakurada K,
Fukuoka M and Sugimoto M (2022)
Evaluation of Spatial Directional
Guidance Using Cheek Haptic
Stimulation in a Virtual Environment.
Front. Comput. Sci. 4:733844.
doi: 10.3389/fcomp.2022.733844

Spatial cues play an important role in navigating people in both physical and virtual spaces. In spatial navigation, visual information with additional cues, such as haptic cues, enables effective guidance. Most haptic devices are applied to various body parts to make mechanical stimuli, while few devices stimulate a head despite the excellent sensitivity. This article presents Virtual Whiskers, a spatial directional guidance technique by cheek haptic stimulation using tiny robot arms attached to a Head-Mounted Display (HMD). The tip of the robotic arm has photo reflective sensors to detect the distance between the tip and the cheek surface. Using the robot arms, we stimulate a point on the cheek obtained by calculating an intersection between the cheek surface and the target direction. In the directional guidance experiment, we investigated how accurately participants identify the target direction provided by our guidance method. We evaluated an error between the actual target direction and the participant's pointed direction. The experimental result shows that our method achieves the average absolute directional error of 2.54° in the azimuthal plane and 6.54° in the elevation plane. We also conducted a spatial guidance experiment to evaluate task performance in a target search task. We compared the condition of visual information, visual and audio information, and visual information and cheek haptics for task completion time, System Usability Scale (SUS) score, NASA-TLX score. The averages of task completion time were $M = 6.39$ s, $SD = 3.34$ s, and $M = 5.62$ s, $SD = 3.12$ s, and $M = 4.35$ s, $SD = 2.26$ s, in visual-only condition, visual+audio condition, and visual+haptic condition, respectively. In terms of the SUS score, visual condition, visual+audio condition, and visual+haptic condition achieved $M = 55.83$, $SD = 20.40$, and $M = 47.78$, $SD = 20.09$, and $M = 80.42$, $SD = 10.99$, respectively. As for NASA-TLX score, visual condition, visual+audio condition, and visual+haptic condition resulted in $M = 75.81$, $SD = 16.89$, and $M = 67.57$, $SD = 14.96$, and $M = 38.83$, $SD = 18.52$, respectively. Statistical tests revealed significant differences in task completion time, SUS score, and NASA-TLX score between the visual and the visual+haptic condition and the visual+audio and the visual+haptic condition.

Keywords: spatial guidance, facial haptics, virtual reality, robot arm, photo reflective sensor

1. INTRODUCTION

Spatial cues are essential in Virtual Reality (VR) applications. In most cases, they rely solely on visual perception. However, VR already has much visual information to process (arguably, more than in the physical environment), and too much visual information can cause visual overload (Stokes and Wickens, 1988). To reduce the visual workload, some adopt audio/visual information, but they also affect the workload since both the audio/visual coordinates need to be translated to the body coordinate (Maier and Groh, 2009) and the head/eyes typically need to face the visual/audio target. As another approach, haptic cues are utilized (Weber et al., 2011). Haptic cues are less likely to be overloaded, are already mapped to the body coordinate, and do not require looking at the target.

Previous works already succeeded haptic-based guidance (Günther et al., 2018; Chen et al., 2019). As a haptic cue for the guidance, vibrotactile stimulation has frequently been adopted. In vibrotactile-based systems, spatial information, such as directions and distance, are encoded into vibration patterns. Most vibrotactile systems have fixed the vibrotactors to stimulation points on the specific body parts. To provide vibrotactile information in an effective manner, the tactile characteristics should be considered. However, the localization of the vibrotactile perception is based on the phantom sensation: perceiving the location of stimulus between actuators as a sensory illusion by amplitude. Therefore, vibrotactors afford limited information.

Also, in most VR haptic systems, users have to wear the haptic devices in addition to the HMD. Such additional setup reduces ease of use. Therefore, by integrating haptic devices the HMD, the VR haptic system become easier to access.

Moreover, for guidance with haptic stimulation where precision is required, not only the type of stimulation (vibration, pressure, wind, thermal, etc.) but also the stimulus position is important. Previous studies reported that, if the head is stationary, the haptic stimuli position is likely to be perceived relatively to a body-centered reference frame (Van Erp, 2005) (i.e., the body midline). On the other hand, if the head is not stationary (like in VR applications), the stimuli location is likely to be perceived relative to an eye-centered reference (Pritchett et al., 2012). Furthermore, if the haptic stimuli are provided on the torso when the head moves, the perceived stimulus position is shifted (Ho and Spence, 2007).

With facial haptic stimuli at the eye-level, we can avoid having two reference frames when the head is non-stationary. A previous study on facial haptics revealed that cheeks' facial haptic was better in localization perception compared to the forehead and to above the eye-brow (Gil et al., 2018). Several facial-based systems have been developed, such as winds (Wilberz et al., 2020) and ultrasounds (Gil et al., 2018).

With the advent of consumer immersive VR Head Mounted Display (HMD), consumer VR haptic devices have emerged in the market^{1,2}. Those devices typically stimulate hands (e.g.,

haptic gloves), waist (e.g., haptic belt), or the body (arms, legs, chest, etc.; e.g., haptic suit). Nevertheless, few devices target the user's face, despite the excellent sensitivity of the facial region (Siemionow, 2011).

In this study, we present Virtual Whiskers, a spatial directional guidance system attached to an HMD with facial haptic stimuli to cheeks (Figure 1). We develop an HMD-based facial haptic system that offers stimulation to the cheeks (Figure 2). In advance of the stimulation, we detect the cheek skin surface using proximity sensors (Figure 3). It consists of two robotic arms attached to the bottom of an HMD (c.f., Figure 4) with proximity sensors. We investigate how accurately haptic cues on the cheek offer directional information; and how effectively facial haptic cues allow spatial guidance in a virtual environment (VE). Our contributions are as follow:

- Cheek stimulation using robot arms integrated into an HMD to provide spatial guidance cues. The robot arms are moved with proximity sensing to control contact with the cheek surface;
- Investigating how cheek stimulation affects directional guidance. In our experiment on directional guidance, we found that haptic cues on the cheek provided accurate direction two cues (azimuth and elevation) in VR space, but that azimuthal angular accuracy was better than elevational one.
- Investigating how effectively cheek stimulation guides users; We experimented on how effectively haptic cues on the cheek improve task performance, usability, and workload. In the target searching task, our guidance technique achieved a shorter task completion time, high usability, low workload than the other conditions that provided only visual information and visual/audio information. Also, our technique enhanced spatial directional perception.

This work is an extended version of the previous work (Nakamura et al., 2021).

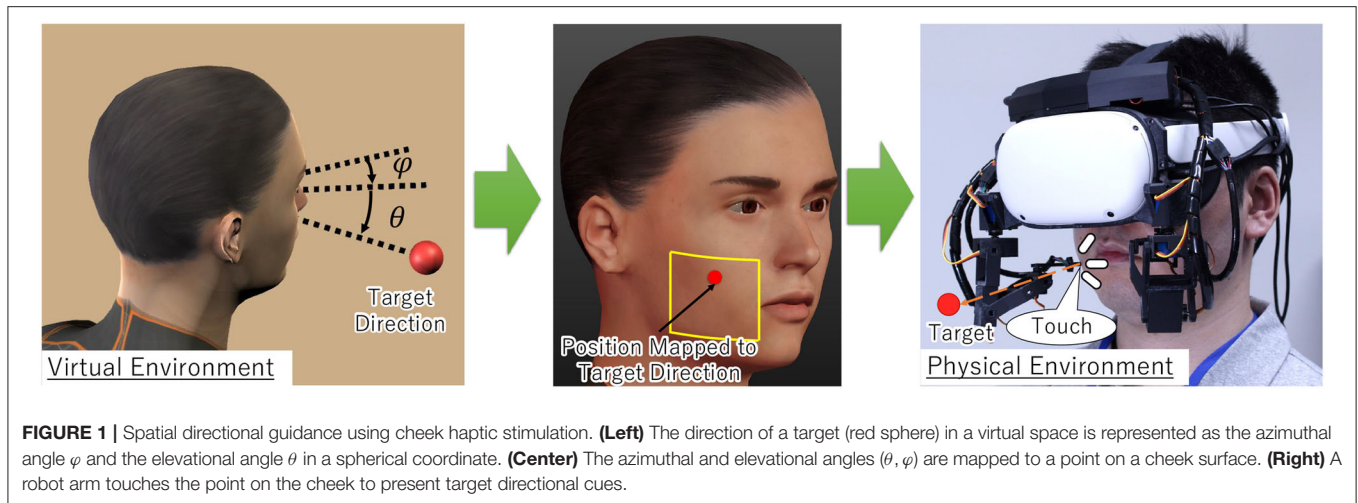
2. RELATED WORK

Haptic cues are important in terms of VR and computer-human interaction. Haptic receptors are distributed throughout the body, while receptors for other major senses: vision, audio, taste, and smell are located in specific facial locations. There are various types of haptic stimuli, such as touch, temperature, and pressure. Previous studies have offered haptic stimulation on various body parts and have integrated other modalities such as visual and audio cues in order to enrich VR experiences as a multi-/cross-modal stimulation.

Our torso has a large surface area, which allows installing haptic actuators on various places to stimulate the large area. Delazio et al. (2018) proposed a force feedback system that placed airbags inside of a vest to enrich VR experience. As one of the most sensitive parts to haptic stimuli, there are our hands. Günther et al. (2018) mapped spatial information into vibration pattern to navigate in 3D space with tactile glove embedded multiple factors. Chen et al. (2019) built a handheld pin-array

¹ bhaptics, <https://www.bhaptics.com/tactsuit/>.

² HaptX | Haptic gloves for VR training, simulation, and design, <https://haptx.com/>.



haptic display to present a direction to the palm. Ion et al. (2015) displayed simple spatial shapes on the user's skin by dragging the skin with tactors embedded into wearable devices. Tsai et al. (2021) proposed a wrist-worn force directional guidance system to provide spatial information by pulling a wristband toward the target direction. As another approach, some studies employed external devices, such as a quadcopter (Hoppe et al., 2018) and swarm robots (Suzuki et al., 2020), to render touchable surfaces. Shen et al. (2020) developed a neck augmentation system using a head-worn robotic arm. Al-Sada et al. (2020) designed a wait-worn wearable robotic arm to provide haptic feedback to an immersive HMD user. They mounted different types of haptic actuators, such as a gripper, a fan, and a brush, on the end effector of the arm to enrich the VR interaction. However, if our head was not stationary, haptic stimuli to the body were encoded in an eye-centered reference frame (Pritchett et al., 2012).

It is equally important that our head has a dense distribution of haptic receptors, which allows perceiving the haptic stimuli sensitively. So, many studies targeted the head region to present haptic cues. Cassinelli et al. built a prototype to let users perceive spatial information with haptic feedback to the head (Cassinelli et al., 2006). Berning et al. (2015) designed a head-worn pressure-based haptic device for augmenting spatial perception by translating distance between the user and surrounding objects into pressure. In terms of VR users, haptic actuators have been integrated into an HMD. Tseng et al. (2019) attached physical widgets to the HMD to interact with the virtual environment. Wang et al. (2019) deployed a skin-stretch array to an HMD facial interface to provide two-dimensional haptic feedback. Some studies tried providing three-dimensional haptic feedback. Tsai and Chen (2019) built impact devices attached to the front side of the HMD to present a 2.5D instant impact. Matsuda et al. (2020) encoded directional information into vibration patterns and presented the pattern around the neck. Kaul and Rohs (2017) installed multiple vibrotactors around the head to offer directional information even at different heights. de Jesus Oliveira et al. (2017) developed a vibrotactile-based guidance system that

provided directional information around the forehead for immersive HMD users. They mapped the azimuthal direction into vibrational position and the elevational direction into vibrational frequency. Moreover, Peiris et al. provided spatial information using thermal feedback on the forehead (Peiris et al., 2017a,b). As indicated above, facial haptics augments our ability and enriches interaction, such as spatial guidance and spatial awareness.

A cheek is one of the facial parts and is also sensitive to haptic stimuli, but few studies utilized it for interaction. In the field of brain-computer interface, Jin et al. (2020) explored the potential of cheek stimuli. A previous study indicated that when in-air ultrasonic haptic cues were provided on the cheek, users could perceive stimulus location well (Gil et al., 2018). A previous study found that cheek haptic stimulation synchronized with visual oscillation by user's footstep reduced VR sickness (Liu et al., 2019). Teo et al. (2020) presented weight sensation by combining visual stimulation with the cheek haptic stimulation and vestibular stimulation. The above approaches stimulated specific locations, while Wilberz et al. (2020) leveraged a robot arm mounted to an HMD to stimulate around the mouth in fully localizable positions. They attached some haptic actuators to provide various feedback, such as wind and ambient temperature, and showed that users could judge directions from wind cues. They also showed that the multi-modal haptic feedback to the face improved the overall VR experience.

Haptic systems enable us to augment our ability. Primarily, the cheeks and mouth are more sensitive than the other facial areas. Previous studies revealed that the cheek had a better ability to locate haptic stimuli than other parts and that localizable stimuli improved the VR experience. Previous studies on mouth and cheek stimulation presented ambient information, such as wind (Wilberz et al., 2020) and ultrasonic (Gil et al., 2018), and investigated horizontal directional cues. With the cheeks' capability of haptic stimuli, we developed a cheek haptic-based guidance system in 3D space. Lip has a thin skin layer and is susceptible to various damage. In addition, direct lip stimulation can cause discomfort. Therefore, we stimulate only the cheek,

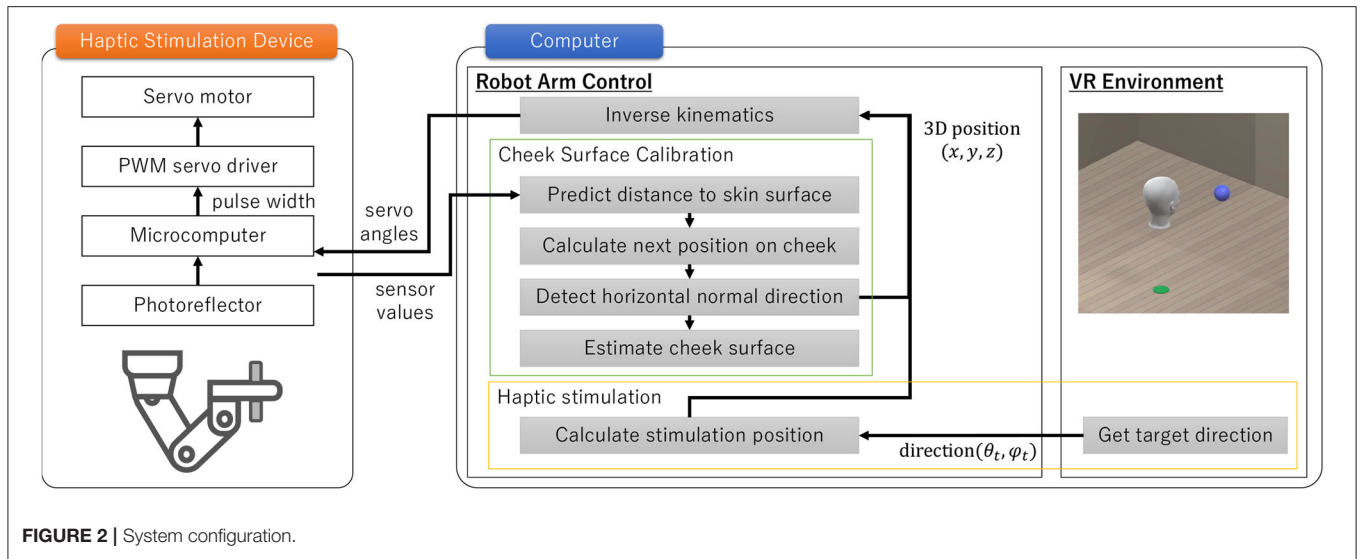


FIGURE 2 | System configuration.

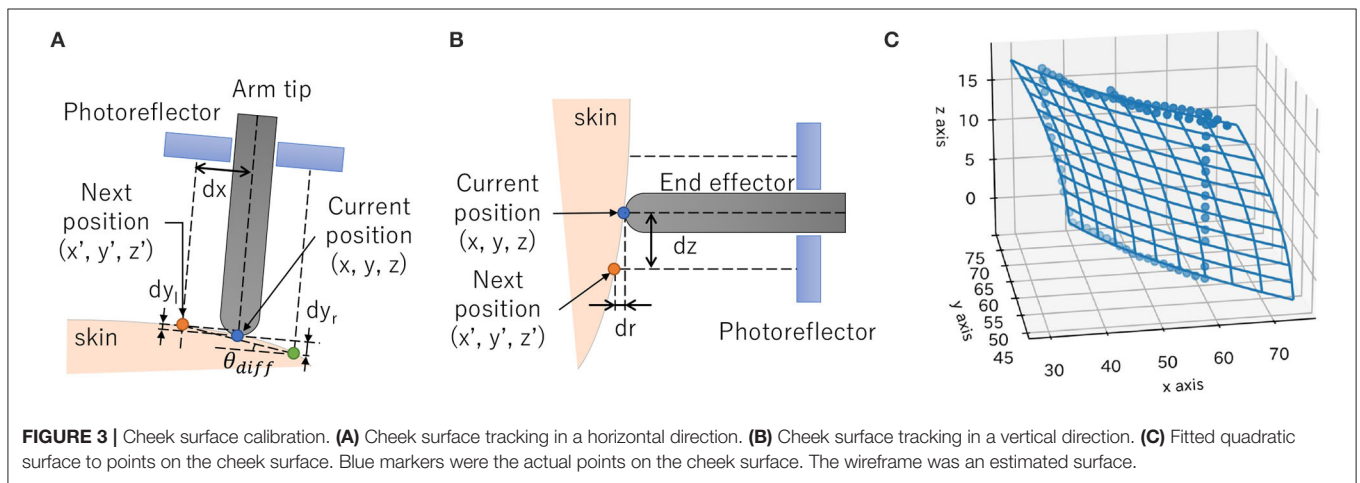


FIGURE 3 | Cheek surface calibration. (A) Cheek surface tracking in a horizontal direction. (B) Cheek surface tracking in a vertical direction. (C) Fitted quadratic surface to points on the cheek surface. Blue markers were the actual points on the cheek surface. The wireframe was an estimated surface.

excluding the lips, to explore the potential of cheek haptics. We investigated how “direct” cheek haptic stimulation could guide in 3D space. Also, we investigated how accurately vertical movements on the cheek can guide users in the elevational plane. We designed a robot arm-based system that stimulated localizable positions in order to present haptic cues on the cheek.

3. SYSTEM DESIGN

We present directional cues on a cheek to guide the user. We stimulate a cheek for presenting directional cues. Our face has a dense distribution of haptic receptors, as indicated in Penfield’s homunculus. In particular, the mouth and cheek are sensitive to mechanical stimuli so that we can recognize the stimulation position precisely. However, the lips are more susceptible to even slight stimulation due to their thin skin layer. In addition, the lip engages in essential activities such as eating and speaking. Therefore, we avoid stimulating the lip. Strong force toward the teeth can injure the mouth due to the teeth’ hardness. Therefore,

we touch both sides of the cheeks with weak to moderate force.

We attach two robotic arms to an HMD to stimulate a cheek. The cheek can detect a stimulation position, which requires a haptic device capable of stimulating the precise position. So, we adopt a robot arm with a linkage that can move to a localizable position on the cheek. The robot arms are attached to an HMD to present the stimulation even while walking around. However, if only a single robot arm is used, long linkages and high-torque motors are required for joints, which increases the overall weight of the arm. Therefore, to avoid the increase of the weight, we placed two robot arms on the left and right sides of the cheeks.

To obtain a point on the cheek skin surface, the accurate surface geometry information is essential. A camera-based approach is popular for measuring facial geometry. However, an HMD and robot arms may occlude the face. When cameras are mounted on the robot arms, the sensing would be difficult due to a motion blur. Therefore, we use photoreflectors as proximity sensors. A photorelector has a light emitting diode and a phototransistor. A light emitting diode emits the light, and

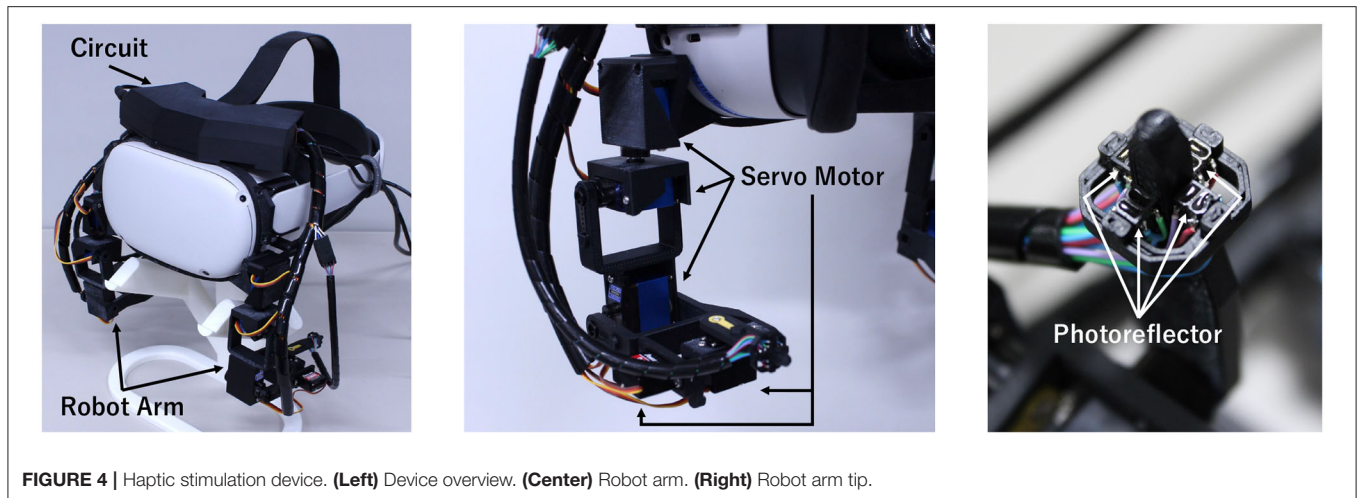


FIGURE 4 | Haptic stimulation device. **(Left)** Device overview. **(Center)** Robot arm. **(Right)** Robot arm tip.

a phototransistor detects the intensity of the light reflected from a skin surface. The light intensity varies with the distance between the photoreflexor and the skin surface. So, the phototransistor value can be converted to a distance with a power-law model. Especially, photoreflexors can detect the distance within a close range, so we attach them to the arm tips. On the four sides (top, down, left, right) of the arm tip, photoreflexors are arranged to detect the distance to the skin surface to obtain the shape around the arm tip. Photoreflexors detect points on the cheek surface, and the arm traces the cheek surface (**Figure 3**). However, the cheek has a large curvature. If the arm tip faces at a constant angle during the tracing of the cheek surface, the photoreflexors and the skin becomes too far from the cheek skin, which can cause the photoreflexors to miscalculate the distance. So, while tracing the cheek, the arm detects the horizontal normal direction to keep the distance within an appropriate range. From the arm trajectory, we collect points on the cheek surface. We assume that the local facial part can be predicted as a quadratic surface. We fit a quadratic surface to the points on the cheek to obtain the cheek surface.

We map a direction to a point on the cheek to present a haptic stimulation indicated a target direction (**Figure 5**). We get the target direction (azimuthal and elevational angles) by converting the Cartesian coordinate system to a spherical coordinate system. The azimuthal and elevational angles are transformed into a line equation. The elevational angle is mapped to a height offset; The azimuthal angle is converted to the line slope. We calculated an intersection between the quadratic cheek surface, and the line encoded the azimuthal and elevational angles. We use this intersection as a stimulation point. However, we avoid touching the lips, which are located on the center of a face. Therefore, when the target is located in the front of the user, we stimulate the left and right cheeks with two robot arms simultaneously. Thus, we present directional cues by touching on the position encoded the target's spatial information.

For the safety of users, we stop and move them away from the cheek if the end effector is at an abnormal position during calibration and cheek stimulation. When the robotic arms stimulate the cheek, the photoreflexors measure the distance

between the end effector and the cheek surface. Using the distance of the photoreflexors, we judge whether the end effector is inside the cheek. If the distance is larger than a threshold, we move the robotic arms toward the opposite direction of the user and stop actuating the robotic arms.

4. IMPLEMENTATION

We developed a system to present directional cues on a cheek (**Figure 2**). Our system consisted of a haptic stimulation device (**Figure 4**) and software. The device detected the cheek surface and presented haptic stimulation on the cheek using two robotic arms. The software controlled the robot arms, detected the cheek surface, and translated a direction to a point on the cheek. We integrated software and hardware to VE developed with Unity.

4.1. Haptic Stimulation Device

We created a haptic stimulation device by modifying an Oculus Quest 2 (**Figure 4**, Left). We attached two robotic arms and a circuit out of the HMD's built-in camera view. The robotic arms were attached to the left and right sides of the bottom of the HMD to stimulate each cheek side. We fixed the robotic arms using brackets created with a 3D printer. The circuit was mounted on the top of the HMD through the HMD's head strap.

The robotic arm was designed with 5 DoF (degree of freedom), allowing stimulating the cheek in 3D space. We used five servo motors of two types for the robot arm joints (**Figure 4**, Center). Three of five were Tower Pro MG92B servo motors that had an enough torque (0.30 N·m) to support the arm. The other two were PowerHD DSM44 servo motors that were lighter than MG92B servo motors and had a less torque (0.12 N·m). MG92B servo motors were placed at the first, second, and third joints. DSM44 servo motors were installed on the fourth and fifth joints to reduce the load of the other joints by reducing the arm tip weight. The arm's tip was rounded not to injure the cheeks. On each arm's tip, four photoreflexors are located on the right, left, up, and down sides (**Figure 4**, Right). The robotic arm could present a force within the range of 0.4 to 5.8 N. In stimulating

the cheek, the robotic arms provided the least force (around 0.4 N) not to injure the cheek.

The circuit consisted of a microcomputer (Arduino Nano) and a 16-channel pulse width modulation (PWM) servo driver (PCA9685, NXP). The microcomputer was connected to a computer *via* a USB cable. The microcomputer received instructions from the computer *via* serial communication to rotate servo motors and send photoreflexor values. The microcomputer instructed the pulse width to the PWM servo driver with I2C to control the servo rotation angle. The Servo motors were supplied the power from a 5V AC adapter. The microcomputer got photoreflexor values *via* an analog multiplexer (TC4052BP, TOSHIBA). The cables were covered by black tubes.

4.2. Robot Arm Control

We utilized inverse kinematics to move the robot arms to a localized 3D position. Inverse Kinematics calculated each arm joint angle corresponding to a provided position and arm tip posture. The calculated angles were sent to the microcomputer to rotate servo motors. The servo motors on the joints were rotated to the calculated angle to move the robot arms. In providing haptic stimulation, we kept the arm tip posture constant. Out of two DSM44 servo motors, one kept the azimuth angle at 60° , and the other kept the elevation angle horizontal to the ground. These angles were decided empirically. In measuring the cheek surface, we controlled the arm tip so that the photoreflexors could detect the skin appropriately as described in section 4.3. We defined the origin in robot arm coordinate as the position of the first arm joint.

4.3. Cheek Surface Estimation

We detected the cheek surface to calibrate the mapping physical and virtual space. We traced the cheek surface with the robotic arms and collected points on the trajectory. By fitting a quadratic surface to the collected 3D points, we estimated the cheek surfaces. The calibration was performed for each arm due to the asymmetrical facial geometry.

In collecting points on the cheek surface, the robot arm traced the cheek surface. At first, the robot arm moved straight from its initial position until it touched the cheek surface. While the arm was moving, the photoreflexors detected the distance from the arm tip to the cheek surface by converting the sensor values into an actual distance with a power-law model ($y = ax^b$). Then, the arm moved along the cheek surface in the following order: rightward, leftward, downward, rightward, and upward. At that time, the arm moved within a range $x_{smin} < x < x_{smax}$, $z_{smin} < z < z_{smax}$. When the arm moved to the right, the photoreflexor on the right side of the arm tip measured the distance to the cheek surface, and the next position to move was computed using Equation (1) (Figure 3A).

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} -\sin \alpha & \cos \alpha & 0 \\ \cos \alpha & \sin \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} dx \\ dy \\ 0 \end{pmatrix} \quad (1)$$

dx was 1.5 because the photoreflexor width was about 3.0 mm. Also, the arm detected the horizontal normal direction to the

cheek at the current position. The photoreflexors on the left and right sides detected the distance to the cheek surface, and then the gap between the normal and the current arm tip direction was calculated as follows: $\theta_{diff} = \arctan((dy_l + dy_r)/2 * dx)$. The tip of the arm was rotated by the difference from the horizontal normal direction of the face. Then, the arm moved to the next position. The arm iterated the above procedure until the arm's position x reached the right edge of the stimulation area (x_{smax}). Next, in case that the arm moved to the left direction, a similar procedure as described above was performed until the arm's position x reached the left edge of the stimulation area (x_{smin}). When the arm moved downward, the photoreflexor located on the bottom side of the arm tip measured the distance to the cheek surface, and the next position to move was calculated using formula 2 (Figure 3B).

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} \cos \beta & -\sin \beta & 0 \\ \sin \beta & \cos \beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ dr \\ dz \end{pmatrix} \quad (2)$$

dz was decided to 1.5 thanks to the photoreflexor width, and β was defined as follows: $\beta = \theta_0 + \theta_4 + 90degrees$ (θ_n was the arm's n th joint angle). The arm iterated to move the detected point until the arm's position z reached the bottom edge of the stimulation area (z_{smin}). A similar procedure as above was performed when the arm moved to the right and up direction. Thus, the arm moved along the cheek surface. By collecting the points on the arm trajectory, we acquired the points of the cheek surface. During the point collection, if the estimated distance dy_l, dy_r, dr was larger than a threshold, our system released the robotic arms from the user's cheek surface and automatically stopped. We empirically decided the threshold as 2.0.

Using the collected points, we estimated the cheek surface as a quadratic surface (Figure 3C). We fitted a quadratic surface to collected points to compute the parameters of the quadratic surface with the least-square method (Equation 3). This way, we obtained the cheek surface as the quadratic surface. The cheek surface estimation was performed for each arm.

$$ax^2 + by^2 + cx + dy + e = z \quad (3)$$

4.4. Haptic Stimulation

We encoded a target's spatial information into a stimulation point and touched it with the robotic arms to present directional cues (Figure 5). We obtained directional information from the target position (x_t, y_t, z_t). Based on the directional information, we determined which robot arm to use to stimulate. Then, we calculated the position on the cheek surface (x_s, y_s, z_s) corresponding to the directional information. We mapped the calculated position (x_s, y_s, z_s) within the stimulation area to determine a stimulation point (x_c, y_c, z_c).

As the first step, we converted the target position (x_t, y_t, z_t) into an azimuthal and an elevational angles (θ_t, φ_t) in the virtual space. By transforming the virtual world coordinate into the head coordinate, we translated the target position (x_t, y_t, z_t) into the position in the head coordinate (x_h, y_h, z_h). The position (x_h, y_h, z_h) was converted into a position in a spherical

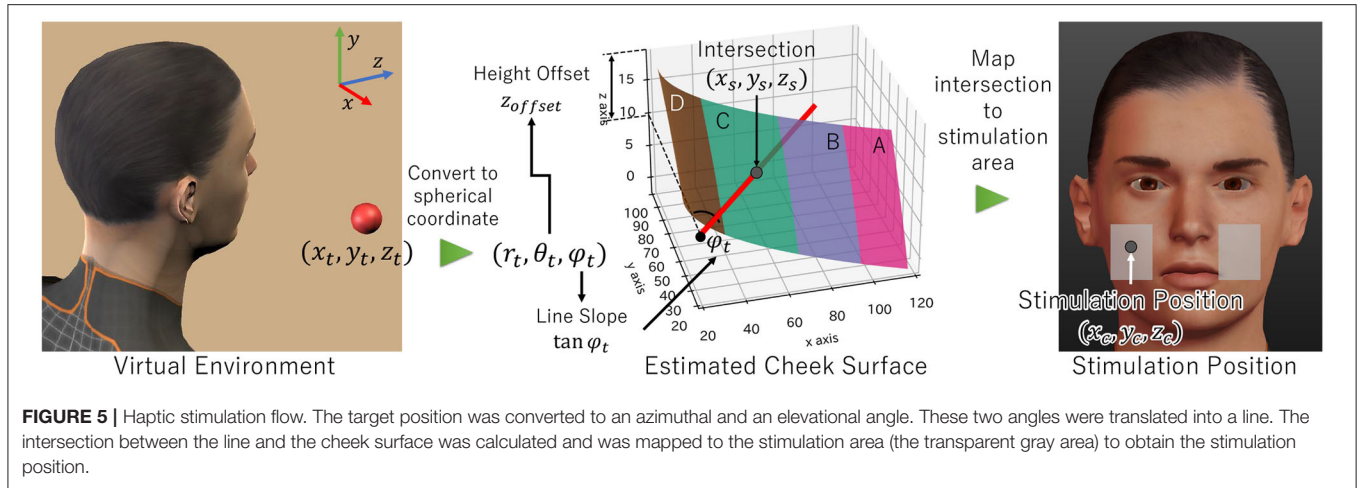


FIGURE 5 | Haptic stimulation flow. The target position was converted to an azimuthal and an elevational angle. These two angles were translated into a line. The intersection between the line and the cheek surface was calculated and was mapped to the stimulation area (the transparent gray area) to obtain the stimulation position.

coordinate (r_t, θ_t, ϕ_t) to obtain an azimuthal and elevational angles (ϕ_t, θ_t) .

In accordance with the azimuthal angle ϕ_t , we determined the robotic arms to use; if the azimuthal angle was within the front area ($abs(\phi_t) \leq \phi_f$), both arms were used; if the target was at the left side ($\phi_t < -\phi_f$), the left arm was used; if the target was at the right side ($\phi_f < \phi_t$), the right arm was used. In our implementation, we set ten degrees as the threshold of the front area ϕ_f .

Then, the azimuthal and elevation angles (ϕ_t, θ_t) were translated into a line to calculate a point on the cheek curved surface (x_s, y_s, z_s) . The elevation angle θ_t was converted into the height of the line z_{offset} as follows:

$$z_{offset} = \begin{cases} z_{min} & (\theta_t < 0) \\ \frac{\sin \theta_t}{\sin \theta_{max}} * (z_{max} - z_{min}) & (0 \leq \theta_t < \theta_{max}) \\ z_{max} & (\theta_{max} \leq \theta_t) \end{cases} \quad (4)$$

In our implementation, θ_{max} was 45 degrees, z_{min} was 0, and z_{max} was 20. The azimuthal angle ϕ_t was used as a slope of the line ($\tan \phi_t$). The line encoded the azimuthal and elevation angles (ϕ_t, θ_t) was as follows:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} \tan \phi_t \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ z_{offset} \end{pmatrix} \quad (t: \text{Constant}) \quad (5)$$

We calculated the intersection (x_s, y_s, z_s) between the estimated cheek surface (formula 3) and the line (formula 5) to obtain the point corresponding to the target's directional information.

According to the calculated position (x_s, y_s, z_s) , we determined stimulation points (x_c, y_c, z_c) . We classified the estimated cheek surface into four areas based on x_s as shown in **Figure 5** Center; area A was $x_f < x_s$; area B was $x_{c_{max}} \leq x_s < x_f$; area C was $x_{c_{min}} \leq x_s \leq x_{c_{max}}$; area D was $x_s < x_{c_{min}}$. x_f was the threshold of the front area, which corresponded to the threshold of the azimuthal angle ϕ_f . $x_{c_{max}}$ and $x_{c_{min}}$ were the maximum and minimum value of the stimulation area respectively. Depending

on which area the calculated point (x_s, y_s, z_s) was, we computed the stimulation position (x_c, y_c, z_c) as follows:

$$(x_c, y_c, z_c) = \begin{cases} (x_{c_{max}}, y_{x_{c_{max}}, z_s}, z_s) & ((x_s, y_s, z_s) \text{ in area A or B}) \\ (x_s, y_{x_s}, z_s) & ((x_s, y_s, z_s) \text{ in area C}) \\ (x_{c_{min}}, y_{x_{c_{min}}, z_s}, z_s) & ((x_s, y_s, z_s) \text{ in area D}) \end{cases} \quad (6)$$

$y_{x_{c_{max}}, z_s}$ was the y value in Equation (3) when $x = x_{c_{max}}$ and $z = z_s$. $y_{x_{c_{min}}, z_s}$ was the y value in Equation (3) when $x = x_{c_{min}}$ and $z = z_s$. If both robot arms were used, the stimulation position was calculated for each side of the cheek. The stimulus encoded azimuthal direction was discrete, while that of elevational direction was continuous. We expected that the continuous stimulation in the elevational plane would let users recognize the height information accurately. Finally, we decided the stimulation position (x_c, y_c, z_c) and touched it with the robot arms to present the directional cues on the cheek.

To protect the safety of the users, if the estimated distance dy_l, dy_r, d_r was larger than a threshold, our system moved the robotic arms to the opposite side of the users, and then stopped robotic arm actuation, as described in section 4.3.

5. EXPERIMENT 1: DIRECTIONAL GUIDANCE

We investigated how accurately haptic cues on the cheeks provided directional information in VR space. In this experiment, participants searched and pointed invisible targets in a virtual environment with haptic cues on cheeks. The targets appeared at 15° intervals of 180° in the azimuthal plane and of 90° in the elevation plane, for a total of 91 locations [13 (azimuth) * 7 (elevation)]. When the target appeared, the robot arms started to stimulate the participant's cheek. The robot arms provided the stimulation according to the participant's head posture and the target location. In the virtual environment, the head position was fixed, while the head rotation was reflected. Participants turned to the target direction and pressed a controller button to point the target direction. We

placed a reticle in front of participants to help them to select the target. We showed instruction texts to participants in the virtual environment. We evaluated the absolute azimuthal and the absolute elevational angular error and task completion time. There were 18 participants (16 males, 2 females; age was $M = 24.28$, $SD = 2.68$), and they had no disability related to tactile organs. The participants were recruited from the students of the Faculty of Science and Technology of our university. Out of 18 participants, 14 had experienced using VR (>10 times), and the others had experienced using VR (>once). Nine of 18 participants played VR less than an hour a week, and the rest of them played VR more than an hour a week. The participants provided written consent form before the experiment.

Our experiment had two sessions, namely, a rehearsal session and an actual performance session. In the rehearsal session, participants looked for the invisible target with the help of haptic stimulation. The rehearsal session consisted of 45 trials, and one target appeared in each trial. We set a five seconds interval between each trial. During the interval, participants were instructed to gaze at the front, and the robot arms stopped stimulating. Participants selected the direction until they selected the correct one. If the error between the selected direction and the actual target direction was within 10 degrees in azimuthal and elevational angles, we regarded it as correct. When participants selected the correct target, they heard the audio feedback for the correct answer and showed the actual target for three seconds. Then, the trail moved on the next. In the actual performance session, participants looked for the invisible target by relying on the haptic stimulation as in the rehearsal session. The actual performance session consisted of 91 trials. When participants selected the direction, they heard the audio feedback, and the trial moved on to the next. We instructed the participants to wear the haptic stimulation device, earplugs, and a noise canceling headphone. During the experiment, white noise was played through the headphone to reduce the noise from the robot arms so that participants hear only the sound from the VE. This study was conducted with a signed consent form collected from participants and an experimental protocol followed the guidelines provided by the research ethics committee at the Faculty of Science and Technology, Keio University.

5.1. Experimental Protocol

We first explained the task and the procedure of the experiment to the participants. Also, we asked the participants to tell us if the participants felt anomalies. We instructed the participants to put on the haptic stimulation device and earplugs and sit him on a chair. Then, we estimated the cheek surface of the participants. At that time, we instructed the participant not to move during the calibration. After the cheek surface estimation, we showed the VE and demonstrated visual and cheek haptic information of the target position to the participant. Here, we instructed the participants to move the head up, down, left, and right to present haptic stimulation when the target was located in front. After the demonstration, we put a noise canceling headphone on the participants. Then, the rehearsal session was performed. In the rehearsal session, the participant searched for the invisible target 45 times. Immediately after the rehearsal session, an actual

performance session was performed. In the actual performance session, the participant searched for the invisible target 91 times. After the actual performance session, the participant removed the earplugs, the headphone, and the device. The participants sit to the chair during the experiment. This experiment took about an hour.

To protect the safety of the participants, we limited the duration of the power supply to the robotic arms to only the duration of the experiment. Until the participants worn the device, we supplied no power to the robotic arms. While the participants performed the task, we carefully observed the robotic arms to abort our system if any incident occurred. At the end of the task, we immediately shut off the power of the robotic arms to allow the participants to take off the device.

5.2. Result

The result is shown in **Figure 6**. Some participants made unintended selections by touching the button. There were three such mistakes, and we removed them. The absolute azimuthal angular error was $M = 2.54$ degrees, $SD = 2.19$ degrees. On the other hand, the absolute elevation angular error was $M = 6.54$ degrees, $SD = 9.34$ degrees (**Figure 6**, left). The histograms of both angular errors respectively in **Figure 6** Center, Right. The task completion time was $M = 13.01$ s, $SD = 6.78$ s. We showed an example of temporal changes of the azimuthal and elevational angular gap between the target direction and head one (**Figure 7**). The histograms of both directions indicate that almost azimuthal and elevation angular errors were within 10° . It is considered because, in the rehearsal session, we defined the error within 10° of both angles as the correct answer.

However, the standard deviation of elevational angular error was larger than the azimuthal one. We expected that the elevational error was similar to the azimuthal one by providing continuous stimulation mapped the elevational direction. However, according to **Figure 7**, the participant, it seemed to be difficult to recognize the elevation angle intuitively. **Figure 7** Right indicates that the participants found the azimuthal direction soon. On the other hand, **Figure 7** Right shows that the participants repeatedly shook their heads in the elevational direction to find the accurate target direction. After the experiment, all participants reported that they had difficulty finding the elevational direction. Furthermore, one out of the participants reported that he could not recognize the base of the height of the stimulation, so he judged whether the stimulation stopped or not. Therefore, our haptic stimulation in elevational angle has room for improvement, as the simple continuous stimulation could result in unclearness.

6. EXPERIMENT 2: SPATIAL GUIDANCE

We investigated how our technique affected the task performance, usability, and workload of spatial guidance in VR space. Participants had to look for and touch the visible target in virtual space. We compared task completion time, System Usability Scale (SUS), and NASA-TLX score in visual+haptic condition with those in visual condition and visual+audio condition. In the visual condition, participants searched the

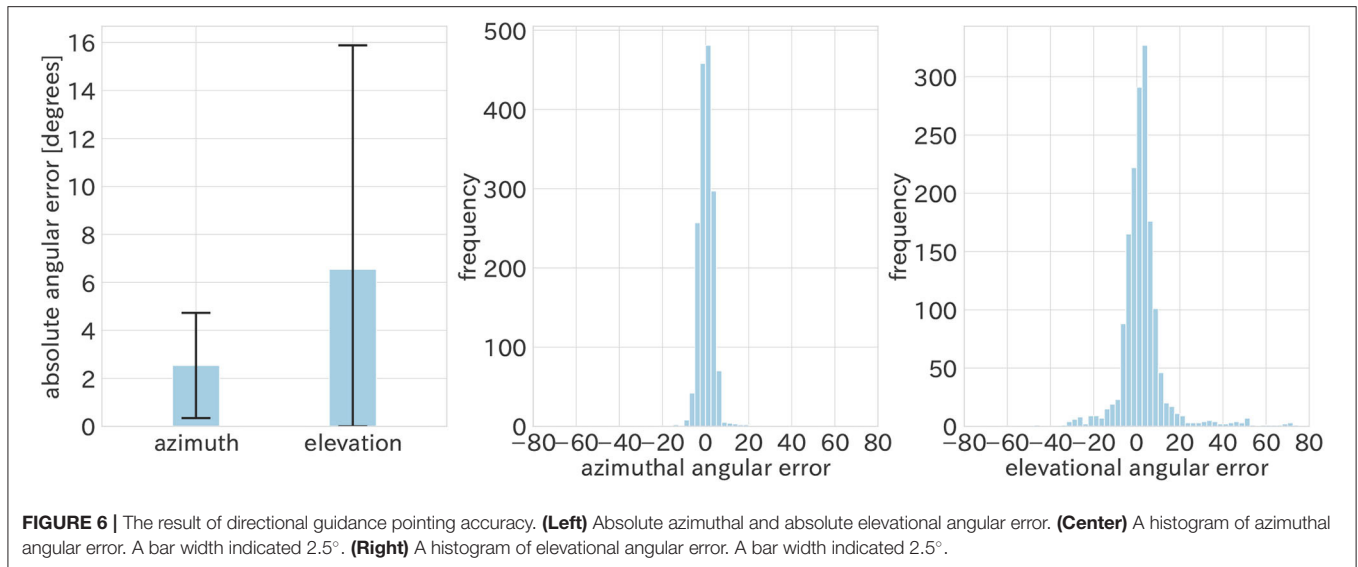


FIGURE 6 | The result of directional guidance pointing accuracy. **(Left)** Absolute azimuthal and absolute elevational angular error. **(Center)** A histogram of azimuthal angular error. A bar width indicated 2.5°. **(Right)** A histogram of elevational angular error. A bar width indicated 2.5°.

target with only visual information (color). In the visual+audio condition, participants looked for the target with visual and audio cues. We used a 440 Hz tone as the basis for the audio cues. We encoded the azimuthal and elevational angle into the difference between left and right sound and the sound frequency, respectively. The sound frequency was transformed based on the difference between the target direction and the participant's frontal direction as follows: $f_i = -440 * \text{abs}(\theta_{diff_e})/90 + 440$, where f_i was sound frequency, θ_{diff_e} was the elevational angular difference between the target direction and the participant's frontal direction. This audio modulation method was inspired by the previous study (Marquardt et al., 2019). However, this modulation method differs from the previous study (Marquardt et al., 2019) in the modulation of the elevational angle; in the previous study, the elevational angle was encoded using a function with a quadratic growth; and, in our study, the elevational angle was modulated using a linear function. In the visual+haptic condition, participants found the target with visual information and haptic stimulation to cheeks. We employed 18 participants (16 males, 2 females; age was $M = 24.28$, $SD = 2.68$) who had no disability related to tactile organs. We recruited the participants from students of the Faculty of Science and Technology of our university. The participants were the same at those who took part in Experiment 1. Of 18 participants, 14 had experienced VR more than ten times, and the rest had experienced VR more than once. Of the 18 participants, nine played VR less than an hour a week, and the others played VR more than an hour a week. In advance of the experiment, we got consent from the participants.

Participants looked for and detected targets of a specific color among many spheres (Figure 8). Participants performed 80 trials and took a break after every 20 trials. In each trial, 50 spheres, including a target and 49 fakes, appeared in a room ($3 \times 3 \times 2$ m). The targets were placed in randomly chosen positions. We generated 40 random positions, and the position components (x , y , z) were sorted randomly to obtain 80 counterbalanced

positions. The other 49 fakes randomly appeared at the beginning of the trial. At that time, we placed each sphere without overlap. The sphere colors were chosen randomly so that the distance between each color was >0.25 in HSV space when the trial started. When the spheres appeared, the robot arms started to stimulate the participant's cheeks. The robot arms provided the stimulation according to the participant's head posture and the target location. The head movements, including the position and posture, and the hand movements were reflected in the virtual environment. This study was conducted with signed consent forms collected from participants and an experimental protocol followed the guidelines provided by the research ethics committee at the Faculty of Science and Technology, Keio University.

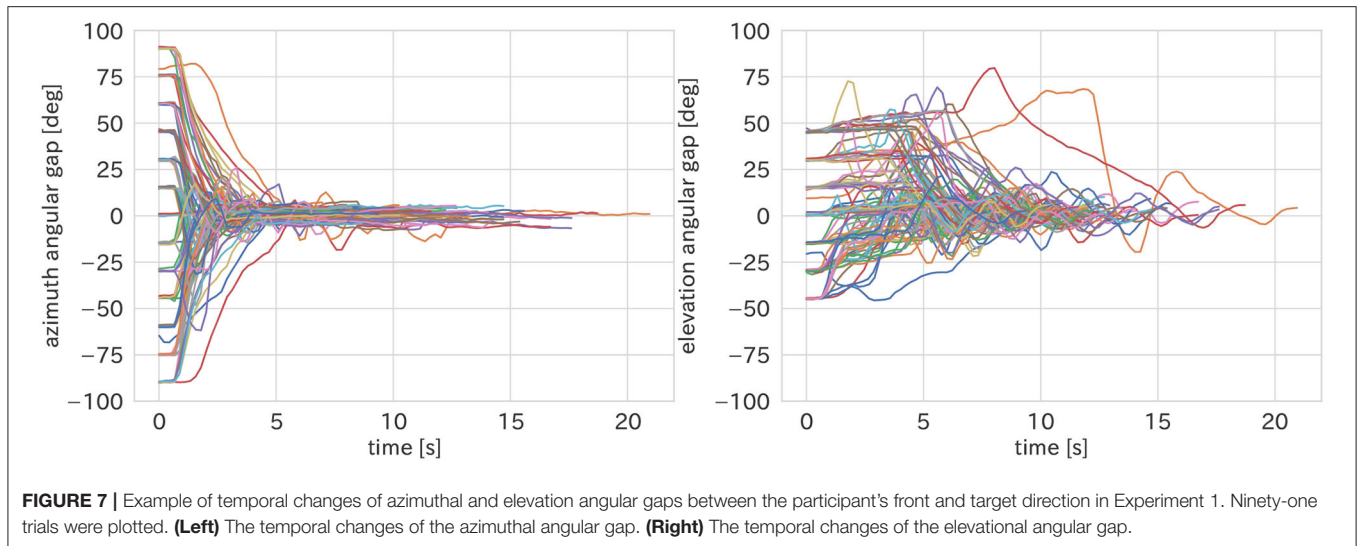
6.1. Experiment Protocol

First, we explained about the task and the conditions to the participants. At that time, we instructed the participants to let us know if the participants felt an issue. Then, we instructed to wear the haptic stimulation device and insert earplugs.

The experiment had two sessions, namely, a rehearsal session and an actual performance session. We first conducted the rehearsal session to let participants train to find the target. In the rehearsal session, participants searched a target with each condition's cues. The session consisted of 20 trials, and a target appeared in each trial. There was a five seconds interval between each trial. During the interval, participants were instructed to look at the front, and the robot arms stopped stimulating. When participants selected the correct target, they heard the audio feedback. In the actual performance session, participants looked for the target with each condition's cues. The session had 80 trials, and there was a break after every 20 trials.

We showed this experiment flow below:

1. The participant was instructed to go to the center of the room. The center was indicated as a circle on the floor.



2. The participant was shown the target to find on the instruction board and was heard a sound.
3. The participant looked for the target. The target to find was shown on the instruction board.
4. When the participant touched the target, the trial was finished.

After each condition, the participants answered the SUS and the NASA-TLX questionnaires. At first, the participants answered ten questions about the SUS by rating a 5-point Likert scale, from 1 “Strongly Disagree” to 5 “Strongly Agree.” Next, the participants rated six subjective subscales about the NASA-TLX scores, namely, “Physical Demand,” “Mental Demand,” “Temporal Demand,” “Performance,” “Effort,” “Frustration.” And then, for every pair of six subjective subscales (15 pairs), the participants answered which subscale was important for them in the task.

When we conducted this experiment, we randomized the order of the conditions for each participant with Latin square method to take a counterbalance of the condition order. The experimental procedure took about one and half hours.

For the safety of the participants, we managed the power supply of the robotic arms strictly. As described in section 5.1, we observed the participants and robotic arms carefully to keep the participant safe.

6.2. Result

The overall average result of task completion time is shown in **Figure 9** Left and the individual results are shown in **Figure 9** Right. The task completion time in the visual condition was $M = 6.39$ s, $SD = 3.34$ s. The task completion time in the visual+audio condition was $M = 5.62$ s, $SD = 3.12$ s. The task completion time in the visual+haptic condition was $M = 4.35$ s, $SD = 2.26$ s. Friedman test revealed that there was a significant difference ($p < 0.001$) in three conditions in terms of the task completion time. *Post-hoc* tests with Bonferroni correction found that there were significant differences between visual and visual+audio condition ($p < 0.05$), visual and visual+haptic condition ($p < 0.001$),

and visual+audio and visual+haptic condition ($p < 0.01$). The SUS score is shown in **Figure 10**. In the visual condition, the SUS score was $M = 55.83$, $SD = 20.40$. In the visual+audio condition, the SUS score was $M = 47.78$, $SD = 20.09$. In the visual+haptic condition, the SUS score was $M = 80.42$, $SD = 10.99$. Friedman test showed a significant difference ($p < 0.001$) in the three conditions in terms of the SUS score. *Post-hoc* test with Bonferroni correction found significant differences between visual and visual+haptic condition ($p < 0.001$), and visual+audio and visual+haptic condition ($p < 0.001$). The visual+haptic condition using our guidance technique achieved the highest SUS score among all conditions, and there were statistical significant differences between the other conditions. Also, the SUS score in visual+haptic condition indicates an “Excellent” system according to the guideline on SUS score interpretation (Bangor et al., 2009). The NASA-TLX overall score is shown in **Figure 11**, and each weighted NASA-TLX score are shown in **Figure 12**. In the visual condition, NASA-TLX total score was $M = 75.81$, $SD = 16.89$. In the visual+audio condition, NASA-TLX total score was $M = 67.57$, $SD = 14.96$. In the visual+haptic condition, NASA-TLX total score was $M = 38.83$, $SD = 18.52$. Friedman test indicated a significant difference ($p < 0.001$) in the three conditions regarding the workload. *Post-hoc* test with Bonferroni correction revealed significant differences between visual and visual+haptic condition ($p < 0.001$), and visual+audio and visual+haptic condition ($p < 0.001$). Mental Demand scores were $M = 16.07$, $SD = 10.44$ and $M = 19.33$, $SD = 7.85$ and $M = 9.31$, $SD = 7.14$ in the visual, the visual+audio, and the visual+haptic condition. Physical Demand scores were $M = 13.87$, $SD = 9.58$ and $M = 8.13$, $SD = 7.00$ and $M = 11.46$, $SD = 8.49$ in the visual, the visual+audio, and the visual+haptic condition. Temporal Demand scores were $M = 2.76$, $SD = 4.62$ and $M = 1.19$, $SD = 2.50$ and $M = 1.70$, $SD = 3.24$ in the visual, the visual+audio, and the visual+haptic condition. Performance scores were $M = 8.04$, $SD = 8.26$ and $M = 7.07$, $SD = 6.90$ and $M = 3.61$, $SD = 2.15$ in the visual, the visual+audio, and the visual+haptic condition. Effort scores were $M = 20.22$, $SD = 10.44$ and $M = 14.13$, $SD = 10.44$ and $M = 10.44$, $SD = 10.44$ in the visual, the visual+audio, and the visual+haptic condition.

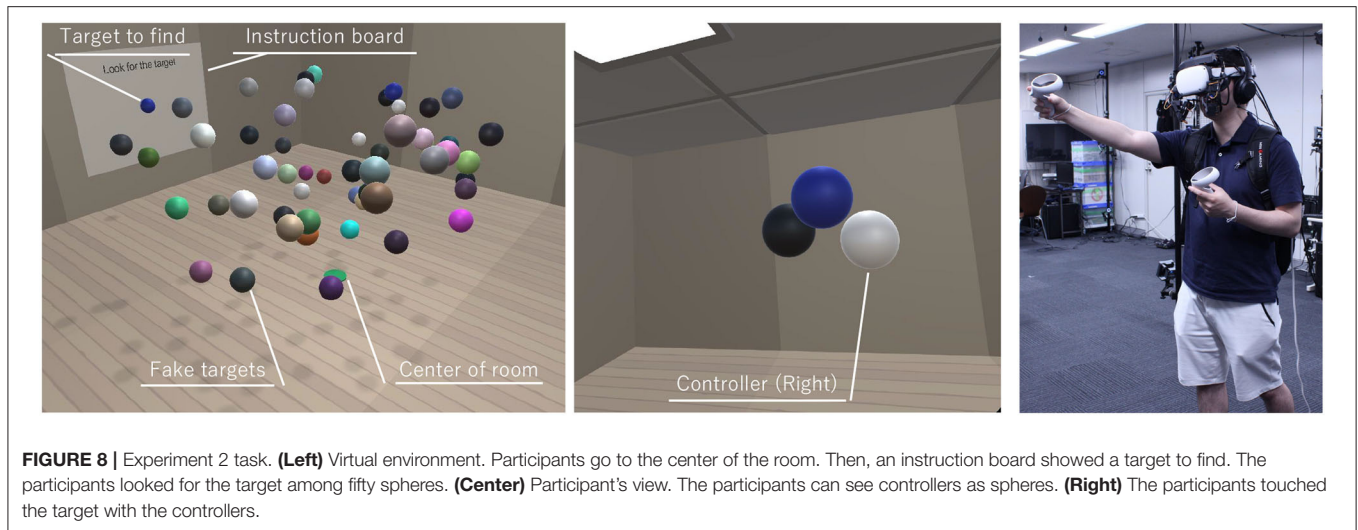


FIGURE 8 | Experiment 2 task. **(Left)** Virtual environment. Participants go to the center of the room. Then, an instruction board showed a target to find. The participants looked for the target among fifty spheres. **(Center)** Participant's view. The participants can see controllers as spheres. **(Right)** The participants touched the target with the controllers.

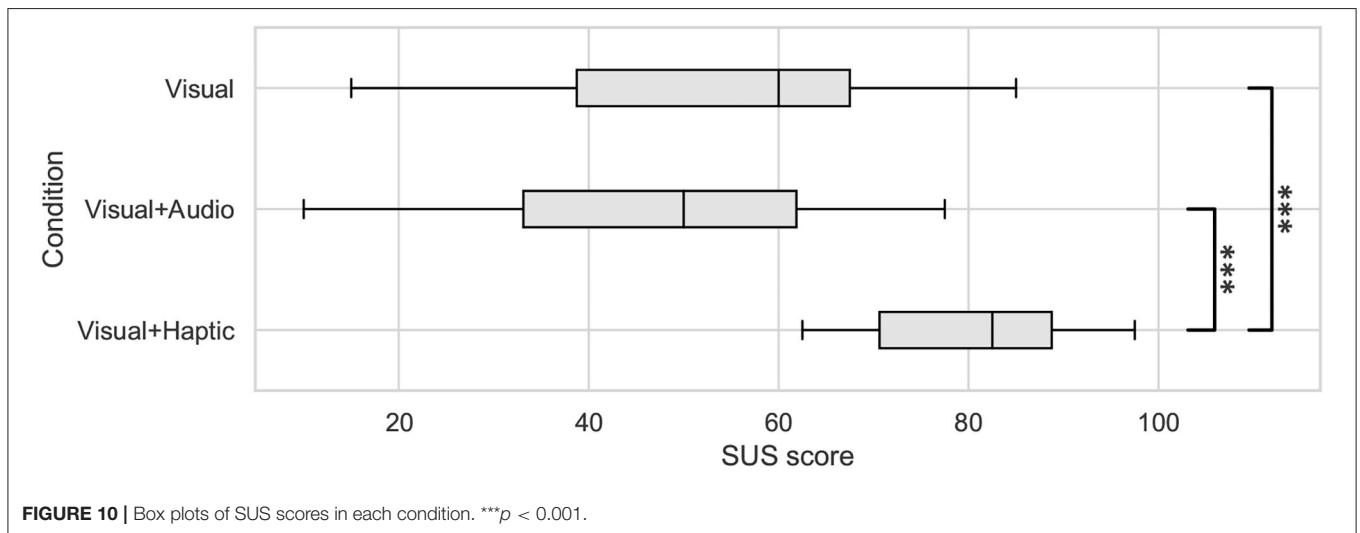
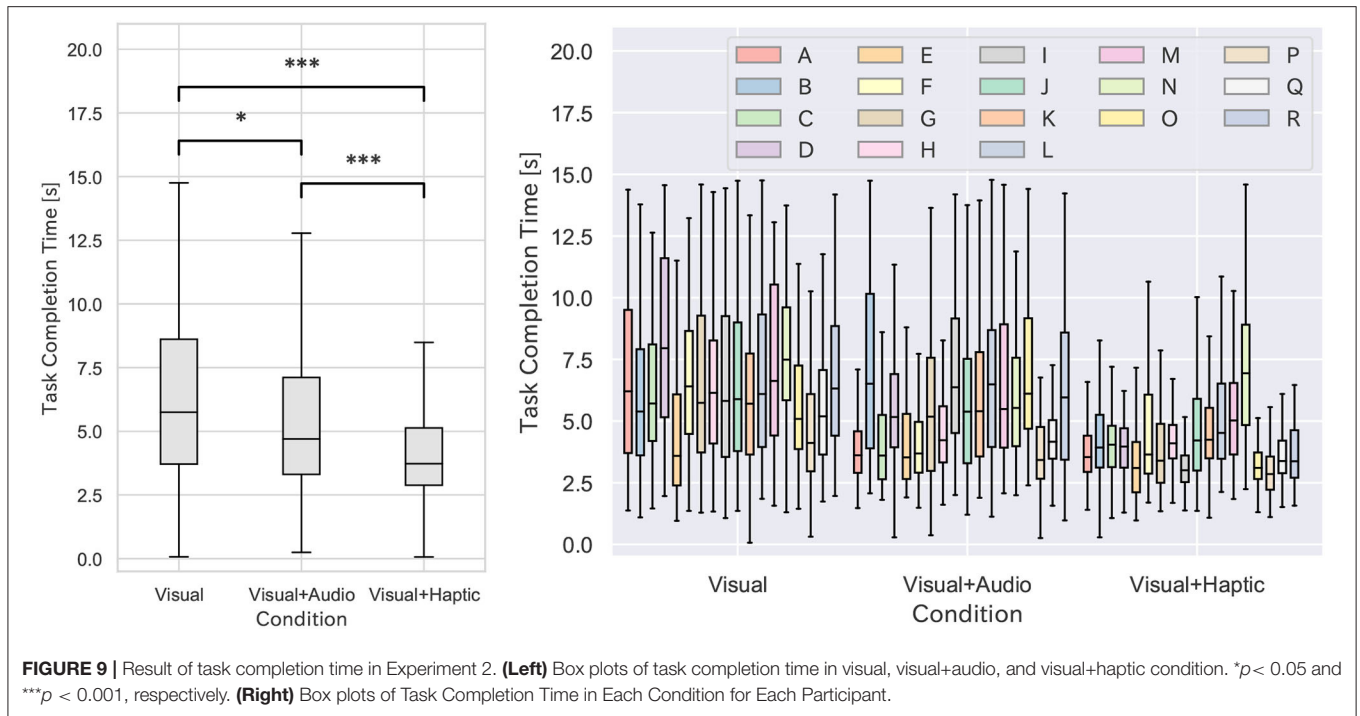
= 6.94 and $M = 17.50$, $SD = 5.22$ and $M = 7.61$, $SD = 6.82$ in the visual, the visual+audio, and the visual+haptic condition. Frustration scores were $M = 14.85$, $SD = 11.30$ and $M = 14.35$, $SD = 11.23$ and $M = 5.13$, $SD = 5.80$ in the visual, the visual+audio, and the visual+haptic condition. As for the Mental Demand score, Friedman test indicated that there is a significant difference between the three conditions ($p < 0.01$). *Post-hoc* tests with Bonferroni correction showed a significant difference between the visual+audio and the visual+haptic condition ($p < 0.001$) in the Mental Demand score. In the Physical Demand score, Friedman test indicated that there is a significant difference between the three conditions ($p < 0.01$). *Post-hoc* tests with Bonferroni correction indicated a significant difference between the visual and the visual+audio conditions ($p < 0.05$). As for the Effort score, Friedman test indicated a significant difference in the three conditions ($p < 0.001$). *Post-hoc* tests with Bonferroni correction revealed significant differences between the visual and the visual+haptic condition ($p < 0.001$) and the visual+audio and the visual+haptic condition ($p < 0.001$) in terms of Effort score. In the Frustration score, Friedman test showed a significant difference in the three conditions ($p < 0.01$). *Post-hoc* tests with Bonferroni correction indicated a significant difference between the visual and the visual+haptic conditions ($p < 0.05$), and visual+audio and visual+haptic conditions ($p < 0.01$) in terms of Frustration score. Therefore, our spatial guidance technique achieved shorter task completion time, higher usability, and less workload in the target searching task.

Looking at the participants' behavior during the experiment, in the visual+haptic condition, the participants tended to turn to the target direction more quickly when searching. On the other hand, the participants seemed to have difficulty detecting the target around the foot or above the head. In the visual+haptic condition, participants first were likely to search the target horizontally from their view. Therefore, it is considered that participants had difficulty finding the target that was located out of their view and at different heights. If the target was around the eye-level, the participants identified the target easily and quickly.

This implies that our technique presented the azimuthal direction appropriately. We also found that users were confused when searching for targets, even when the targets seemed to be hidden by fakes from the participant's view. Since our system did not provide depth information, such a lack of information might cause the participants to feel an inconsistency between the visual and haptic information when the targets were occluded by fakes.

After the experiment, several participants reported that they changed their behavior in visual+haptic conditions. According to the participant's report, in the visual condition, the participants made a great effort to learn the target's color without making mistakes. Although the visual+audio condition offered additional spatial information with audio cues, the participants mentioned that they had to concentrate on the audio to detect the directions of the target. Eventually, the visual+audio condition imposed the participants to pay attention to not only the target but also the audio cues, which caused a worse SUS score than the visual condition. On the other hand, in the visual+haptic condition, the participants reported that, thanks to the haptic cues, they paid less attention to the target's color than the visual condition. This less attention made the Effort and the Frustration scores better than the other two conditions (**Figure 12**). Especially, as the visual+haptic condition achieved the significantly better Mental Demand score than the visual+audio condition, additional haptic cues conveyed the spatial direction by paying a little attention. In addition, several participants mentioned that they were confident in touching the target because of the haptic cues. This increase of the confidence implies that our haptic stimulation augmented the spatial awareness in a successful manner. Therefore, the visual+haptic condition, which used our haptic stimulation system, allowed the participants to confidently find the target with less attention to the target and the haptic cues.

Also, some participants reported that they lost the stimulation when they changed their facial expressions after successfully detecting the target in the visual+haptic condition. This loss of stimulation was caused by our cheek surface estimation

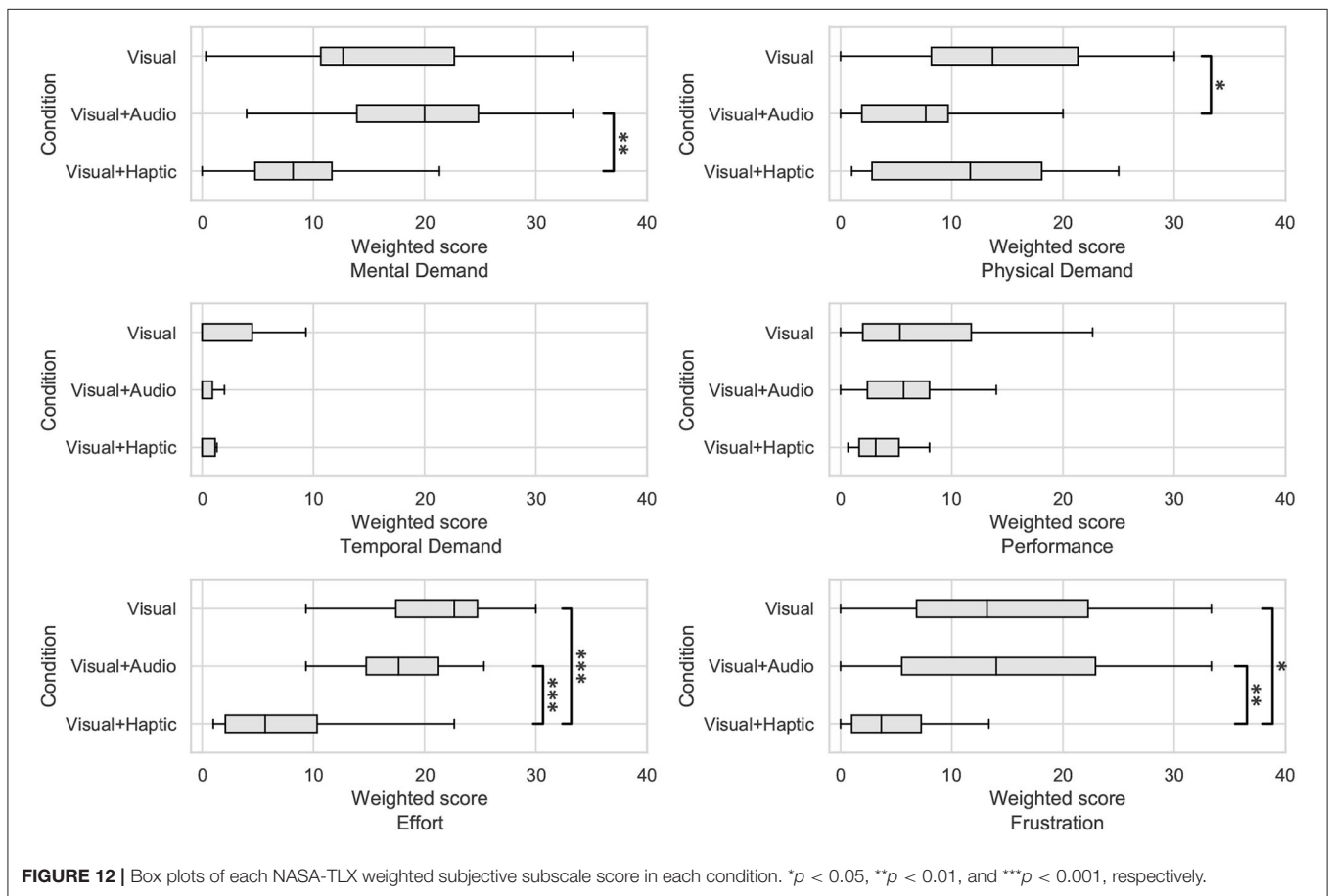
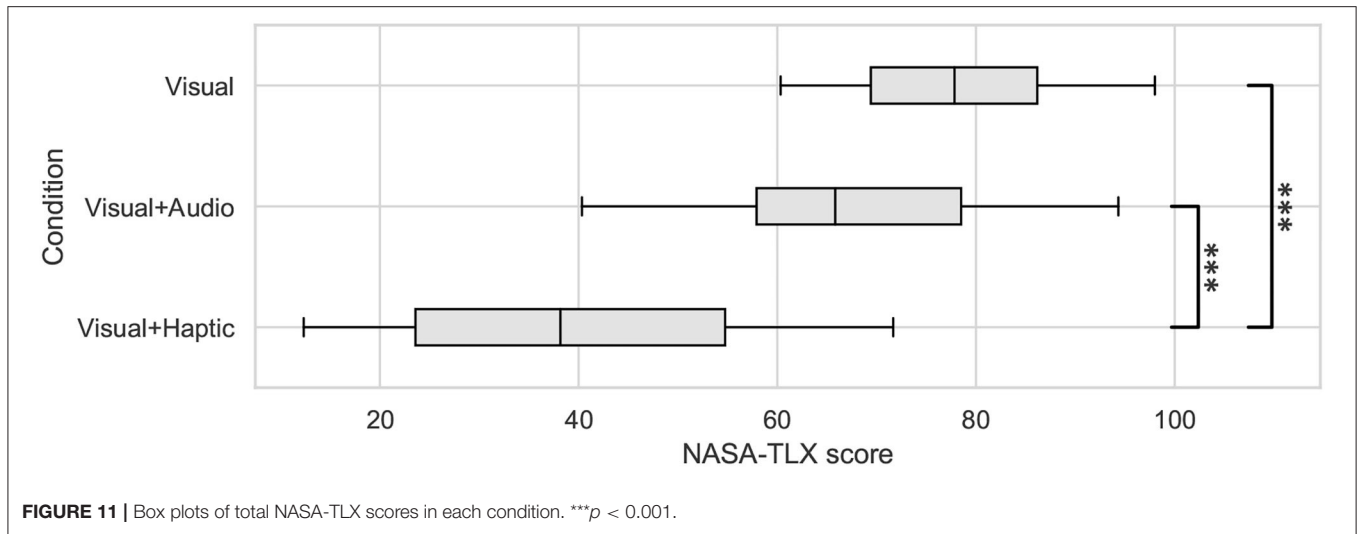


method. We detected the cheek surface only once before the stimulation. However, the cheek changes its shape, for example, when making facial expressions. Especially near the lips, the shape deforms significantly. Since our system stimulated the region around the lips when the target was in the front, it seems to lose the haptic stimulation that indicated the front direction.

7. LIMITATION

Our technique estimated the cheek surface as static to present the target direction. However, the cheek deforms when speaking,

eating, and forming facial expressions, which can cause a robot arm not to reach the cheek. In Experiment 2, some participants mentioned that they lost cheek cues by changing facial expressions. Especially, the geometry around the mouth deforms complexly thanks to several facial muscles, making it difficult for the robotic arms to reach. In addition, if an HMD position is shifted, the estimated surface also is shifted so that the robot arm may not touch the cheek. Similar issues caused by the misalignment between the HMD and the face have been mentioned in previous studies on photorelector-based sensing. Therefore, it is important to detect cheek surface in real-time to make the surface estimation



robust to the positional drift and capture the dynamic facial deformation.

However, in case of real-time sensing, servo delays significantly affect the system responsiveness. Our system calculated a point of cheek surface based on the distance predicted with photoreflectors and robotic arm position.

However, if our system get sensor data before arm actuation, our system mispredicts the positions, which causes wrong cheek surface estimation. Therefore, our current system waited for robot arm actuation for accurate cheek surface sensing. However, if sensing and stimulation are tried at the same time in our current system configuration, the stimulation is delayed by the

latency of the sensing at least. In our current implementation, the used servo motors are controlled by PWM control width of 20 ms. To actuate robotic arms more quickly, the high responsive servo motors are required.

In our spatial guidance experiment, we modulated elevational angular information into audio frequency to compare visual condition and visual+haptic condition. We translated the elevational angle into the pitch of audio cues. However, as the other method to modulate elevational information into frequency, head related transfer function-based methods were adopted (Sodnik et al., 2006; Rodemann et al., 2008). Therefore, by employing the head related transfer function-based approaches, the result may be changed.

We employed young participants in both experiments. The sensory systems of human, such as hearing and haptics, decline as we age. However, in the haptic sensation, tactile capability is attenuated and, especially, in the lower limbs. Therefore, the experimental results may be changed by elderly participants.

Our implementation provided slight stimulation on the cheek to present spatial directional information in a safe manner. However, it is still unclear whether the slight stimulation is appropriate in offering spatial guidance or not. Therefore, it is important to investigate comfortable force for using the cheek as a haptic display.

In our system, to convey the front direction, we stimulated both sides of the cheeks. This is because we avoided to stimulate the lips due to the susceptibility of the lips to injury. However, there is still room to consider how to present frontal direction by cheek stimulation in a safe manner. For example, when the frontal direction is presented, soft end effector and ambient stimulation are considered to be employed. As the future work, we will investigate the methods to present frontal direction.

Our system employed servo motors for the robot arm. However, the angular step of the servo motor was 1.8° . In the future, it will be tested whether using a high-resolution servo can make the stimulation more accurate.

We leveraged a robotic arm with some linkages as a stimulation device. However, the arm with linkages requires multiple joints with servo motors, which increases the weight of the arm. This weight increase makes the users tired. Therefore, it is necessary to construct a lightweight robot arm.

We measured facial surface using photoreflectors to provide slight touch on the cheek. According to several experimental participants, provided haptic stimulation was very light. Therefore, even our current configuration is sufficient to present safe stimulation. By introducing force probes on the tip of robotic arms, the force on the cheek can be detected the actual force more precisely, which makes the haptic presentation safer.

The result of experiment 1 indicated that the elevational angular error was larger than the azimuthal one. While participants seemed to have difficulty moving their heads slightly, there is a possibility that the limitation of our robot arm-based haptic stimulation led to this result. However, it is unclear how precisely we can control our heads in the elevational direction by ourselves. In the future, we will investigate the controllable angle of the head by ourselves.

Generally, continuous haptic stimulation can lose the sensation by sensory adaptation, while the participants did not report that in our experiments. We will investigate if such adaptation occurs when stimulating the cheek with our method for a longer duration.

8. CONCLUSION

We proposed Virtual Whiskers, a spatial directional guidance system that provided haptic cues on the cheeks with robot arms attached to an HMD. We placed robot arms to the left and right sides of the bottom of an HMD, allowing to offer haptic cues on both sides of the cheek. We detected points on the cheeks' surface using photo reflective sensors located on the robot arm tip and fitted a quadratic surface to the points to estimate the cheek surface. In stimulating the cheek, we calculated the point

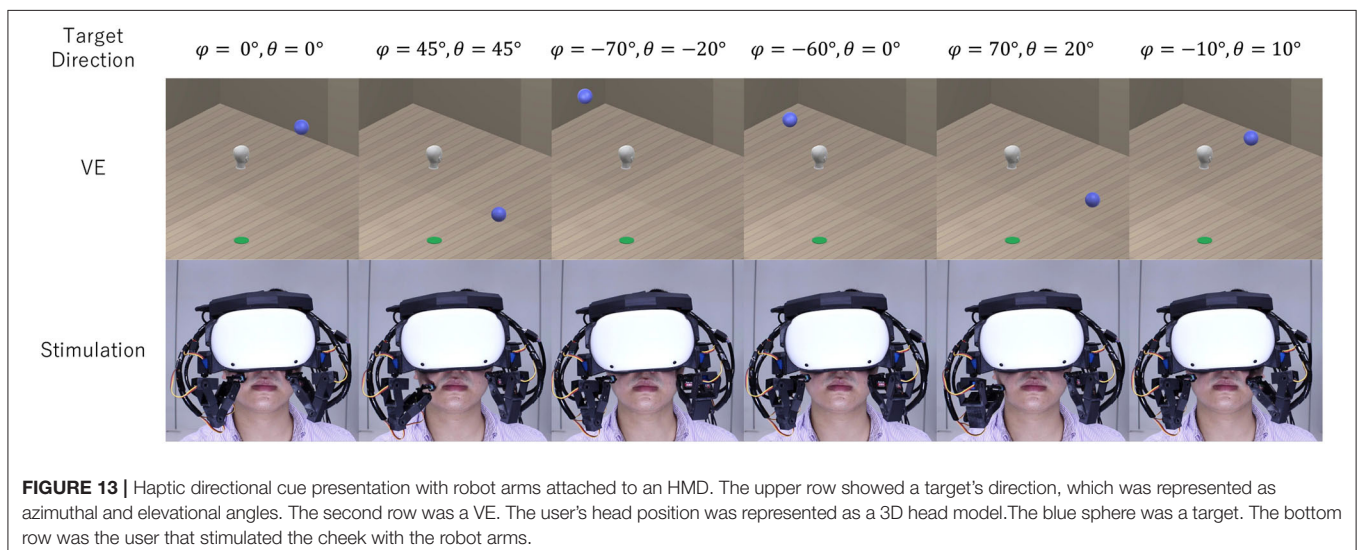


FIGURE 13 | Haptic directional cue presentation with robot arms attached to an HMD. The upper row showed a target's direction, which was represented as azimuthal and elevational angles. The second row was a VE. The user's head position was represented as a 3D head model. The blue sphere was a target. The bottom row was the user that stimulated the cheek with the robot arms.

on the estimated surface that was encoded the targets' azimuthal and elevational angles in the VR space and mapped the point to the stimulation position. According to spatial information of the target and user's head, we touched the cheeks with the arms to present directional cues (**Figure 13**).

We conducted an experiment on the directional guidance accuracy by haptic directional cues on the cheeks and evaluated the pointing accuracy. Our method achieved the absolute azimuthal pointing error of $M = 2.36$ degrees, $SD = 1.55$ degrees, and the absolute elevational absolute pointing error of $M = 4.51$ degrees, $SD = 4.74$ degrees. We also conducted an experiment on task performance in a spatial guidance task using our guidance technique. We compared the task completion time, SUS score, and NASA-TLX score in the target searching task in the three conditions; only visual information was presented; visual and audio cues were presented; visual and haptic information was presented. The result showed that the averages of task completion time were $M = 6.68$ s, $SD = 3.45$ s and $M = 6.20$ s, $SD = 3.30$ s and $M = 4.22$ s, $SD = 2.05$ s in visual and visual+audio and visual+haptic condition, respectively. The averages of the SUS score were $M = 44.17$, $SD = 21.83$, and $M = 44.44$, $SD = 23.24$, and $M = 86.11$, $SD = 8.76$ in visual and visual+audio and visual+haptic condition, respectively. The averages of NASA-TLX scores were $M = 75.22$, $SD = 22.93$ and $M = 71.03$, $SD = 12.13$ and $M = 35.00$, $SD = 17.16$ in visual and visual+audio and visual+haptic condition, respectively. Statistical tests showed significant differences in task completion time between each condition, In terms of SUS score and NASA-TLX score, statistical tests indicated significant differences between visual and visual+haptic condition and visual+audio and visual+haptic condition.

Through our experiments, we showed the effectiveness of cheek haptics on spatial guidance. Our technology could be applied to collaboration in a virtual environment. In collaboration with other people, it is important to share attention and interest. In such a case, our technique can indicate attention to users. Our system is integrated into an HMD and makes the user's hands free. Therefore, cheek haptic-based interaction allows users to collaborate effectively.

REFERENCES

- Al-Sada, M., Jiang, K., Ranade, S., Kalkattawi, M., and Nakajima, T. (2020). Hapticsnakes: multi-haptic feedback wearable robots for immersive virtual reality. *Virt. Real.* 24, 191–209. doi: 10.1007/s10055-019-00404-x
- Bangor, A., Kortum, P., and Miller, J. (2009). Determining what individual SUS scores mean: adding an adjective rating scale. *J. Usabil. Stud.* 4, 114–123.
- Berning, M., Braun, F., Riedel, T., and Beigl, M. (2015). "Proximityhat: a head-worn system for subtle sensory augmentation with tactile stimulation," in *Proceedings of the 2015 ACM International Symposium on Wearable Computers, ISWC '15* (New York, NY: Association for Computing Machinery), 31–38. doi: 10.1145/2802083.2802088
- Cassinelli, A., Reynolds, C., and Ishikawa, M. (2006). "Augmenting spatial awareness with haptic radar," in *2006 10th IEEE International Symposium on Wearable Computers* (Montreux: IEEE), 61–64. doi: 10.1109/ISWC.2006.286344

In our system, the robot arms were used to present a single direction by stimulating the cheek. However, the several robot arms attached to the HMD enable tangible interaction on a cheek for immersive HMD users. In the future, we will explore the potential application and the interaction technique with virtual objects.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

FN developed the system, implemented the hardware and software, designed and conducted experiments, and decided the direction of this study. AV implemented a virtual environment and designed the experiment. KS gave advice about the facial haptics and spatial guidance design and helped experiments. MF gave advice about the experimental design and analysis of experimental result. MS gave advice about the direction of this study as a supervisor. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

This project was supported by JST ERATO Grant Number JPMJER1701 and JSPS KAKENHI Grant Number 16H05870 and Grant-in-Aid for JSPS Research Fellow for Young Scientists (DC2) No.21J13664.

- Chen, D. K., Chossat, J.-B., and Shull, P. B. (2019). "Haptivec: presenting haptic feedback vectors in handheld controllers using embedded tactile pin arrays," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19* (New York, NY: Association for Computing Machinery), 1–11. doi: 10.1145/3290605.3300401
- de Jesus Oliveira, V. A., Brayda, L., Nedel, L., and Maciel, A. (2017). Designing a vibrotactile head-mounted display for spatial awareness in 3d spaces. *IEEE Trans. Visual. Comput. Graph.* 23, 1409–1417. doi: 10.1109/TVCG.2017.2657238
- Delazio, A., Nakagaki, K., Klatzky, R. L., Hudson, S. E., Lehman, J. F., and Sample, A. P. (2018). "Force jacket: Pneumatically-actuated jacket for embodied haptic experiences," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI '18* (New York, NY: Association for Computing Machinery), 1–12. doi: 10.1145/3173574.3173894
- Gil, H., Son, H., Kim, J. R., and Oakley, I. (2018). "Whiskers: exploring the use of ultrasonic haptic cues on the face," in *Proceedings of the 2018 CHI Conference*

- on *Human Factors in Computing Systems*, CHI '18 (New York, NY: Association for Computing Machinery), 1–13. doi: 10.1145/3173574.3174232
- Günther, S., Müller, F., Funk, M., Kirchner, J., Dezfali, N., and Mühlhäuser, M. (2018). “Tactileglove: assistive spatial guidance in 3d space through vibrotactile navigation,” in *Proceedings of the 11th Pervasive Technologies Related to Assistive Environments Conference, PETRA '18* (New York, NY: Association for Computing Machinery), 273–280. doi: 10.1145/3197768.3197785
- Ho, C., and Spence, C. (2007). Head orientation biases tactile localization. *Brain Res.* 1144, 136–141. doi: 10.1016/j.brainres.2007.01.091
- Hoppe, M., Knierim, P., Kosch, T., Funk, M., Futami, L., Schneegass, S., et al. (2018). “VRhapticdrones: providing haptics in virtual reality through quadcopters,” in *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia, MUM 2018* (New York, NY: Association for Computing Machinery), 7–18. doi: 10.1145/3282894.3282898
- Ion, A., Wang, E. J., and Baudisch, P. (2015). “Skin drag displays: dragging a physical tactor across the user’s skin produces a stronger tactile stimulus than vibrotactile,” in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15* (New York, NY: Association for Computing Machinery), 2501–2504. doi: 10.1145/2702123.2702459
- Jin, J., Chen, Z., Xu, R., Miao, Y., Wang, X., and Jung, T. P. (2020). Developing a novel tactile p300 brain-computer interface with a cheeks-STIM paradigm. *IEEE Trans. Biomed. Eng.* 67, 2585–2593. doi: 10.1109/TBME.2020.2965178
- Kaul, O. B., and Rohs, M. (2017). “Haptichead: a spherical vibrotactile grid around the head for 3d guidance in virtual and augmented reality,” in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17* (New York, NY: Association for Computing Machinery), 3729–3740. doi: 10.1145/3025453.3025684
- Liu, S., Yu, N., Chan, L., Peng, Y., Sun, W., and Chen, M. Y. (2019). “Phantomlegs: reducing virtual reality sickness using head-worn haptic devices,” in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 817–826. doi: 10.1109/VR.2019.8798158
- Maier, J. X., and Groh, J. M. (2009). Multisensory guidance of orienting behavior. *Hear. Res.* 258, 106–112. doi: 10.1016/j.heares.2009.05.008
- Marquardt, A., Trepkowski, C., Eibich, T. D., Maiero, J., and Kruijff, E. (2019). “Non-visual cues for view management in narrow field of view augmented reality displays,” in *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 190–201. doi: 10.1109/ISMAR.2019.000-3
- Matsuda, A., Nozawa, K., Takata, K., Izumihara, A., and Rekimoto, J. (2020). “Hapticpointer: a neck-worn device that presents direction by vibrotactile feedback for remote collaboration tasks,” in *Proceedings of the Augmented Humans International Conference, AHs '20* (New York, NY: Association for Computing Machinery), 1–10. doi: 10.1145/3384657.3384777
- Nakamura, F., Adrien, V., Sakurada, K., and Sugimoto, M. (2021). “Virtual whiskers: spatial directional guidance using cheek haptic stimulation in a virtual environment,” in *Proceedings of the 2021 Augmented Humans International Conference, AHs '21*, 1–10. doi: 10.1145/3458709.3458987
- Peiris, R. L., Peng, W., Chen, Z., Chan, L., and Minamizawa, K. (2017a). “Thermovr: exploring integrated thermal haptic feedback with head mounted displays,” in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17* (New York, NY: Association for Computing Machinery), 5452–5456. doi: 10.1145/3025453.3025824
- Peiris, R. L., Peng, W., Chen, Z., and Minamizawa, K. (2017b). “Exploration of cuing methods for localization of spatial cues using thermal haptic feedback on the forehead,” in *2017 IEEE World Haptics Conference (WHC)*, 400–405. doi: 10.1109/WHC.2017.7989935
- Pritchett, L. M., Carnevale, M. J., and Harris, L. R. (2012). Reference frames for coding touch location depend on the task. *Exp. Brain Res.* 222, 437–445. doi: 10.1007/s00221-012-3231-4
- Rodemann, T., Ince, G., Joublin, F., and Goerick, C. (2008). “Using binaural and spectral cues for azimuth and elevation localization,” in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2185–2190. doi: 10.1109/IROS.2008.4650667
- Shen, L., Saraiji, M. Y., Kunze, K., Minamizawa, K., and Peiris, R. L. (2020). “Visuomotor influence of attached robotic neck augmentation,” in *Symposium on Spatial User Interaction, SUI '20* (New York, NY: Association for Computing Machinery), 1–10. doi: 10.1145/3385959.3418460
- Siemionow, M. Z., (ed.). (2011). *The Know-How of Face Transplantation*. London: Springer London. doi: 10.1007/978-0-85729-253-7
- Sodnik, J., Tomazic, S., Grasset, R., Duenser, A., and Billingham, M. (2006). “Spatial sound localization in an augmented reality environment,” in *Proceedings of the 18th Australia Conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments, OZCHI '06* (New York, NY: Association for Computing Machinery), 111–118. doi: 10.1145/1228175.1228197
- Stokes, A. F., and Wickens, C. D. (1988). “Aviation displays,” in *Human Factors in Aviation* (Elsevier). doi: 10.1016/B978-0-08-057090-7.50018-7
- Suzuki, R., Hedayati, H., Zheng, C., Bohn, J. L., Szafir, D., Do, E. Y.-L., et al. (2020). “Roomshift: room-scale dynamic haptics for VR with furniture-moving swarm robots,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20* (New York, NY: Association for Computing Machinery), 1–11. doi: 10.1145/3313831.3376523
- Teo, T., Nakamura, F., Sugimoto, M., Verhulst, A., Lee, G. A., Billingham, M., et al. (2020). “WeightSync: proprioceptive and haptic stimulation for virtual physical perception,” in *ICAT-EGVE 2020 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, eds F. Argelaguet, R. McMahan, M. and Sugimoto (The Eurographics Association), 1–9.
- Tsai, H.-R., Chang, Y.-C., Wei, T.-Y., Tsao, C.-A., Koo, X. C.-Y., Wang, H.-C., et al. (2021). “Guideband: Intuitive 3d multilevel force guidance on a wristband in virtual reality,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY: Association for Computing Machinery), 1–13. doi: 10.1145/3411764.3445262
- Tsai, H.-R., and Chen, B.-Y. (2019). “Elastimpact: 2.5d multilevel instant impact using elasticity on head-mounted displays,” in *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, UIST '19* (New York, NY: Association for Computing Machinery), 429–437. doi: 10.1145/3332165.3347931
- Tseng, W.-J., Wang, L.-Y., and Chan, L. (2019). “Facewidgets: exploring tangible interaction on face with head-mounted displays,” in *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, UIST '19* (New York, NY: Association for Computing Machinery), 417–427. doi: 10.1145/3332165.3347946
- Van Erp, J. B. (2005). Presenting directions with a vibrotactile torso display. *Ergonomics* 48, 302–313. doi: 10.1080/0014013042000327670
- Wang, C., Huang, D.-Y., Hsu, S.-W., Hou, C.-E., Chiu, Y.-L., Chang, R.-C., et al. (2019). “Masque: exploring lateral skin stretch feedback on the face with head-mounted displays,” in *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, UIST '19* (New York, NY: Association for Computing Machinery), 439–451. doi: 10.1145/3332165.3347898
- Weber, B., Schätzle, S., Hulin, T., Preusche, C., and Deml, B. (2011). “Evaluation of a vibrotactile feedback device for spatial guidance,” in *2011 IEEE World Haptics Conference, WHC 2011* (Istanbul: IEEE), 349–354. doi: 10.1109/WHC.2011.5945511
- Wilberz, A., Leschtschow, D., Trepkowski, C., Maiero, J., Kruijff, E., and Riecke, B. (2020). “Facehaptics: robot arm based versatile facial haptics for immersive environments,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20* (New York, NY: Association for Computing Machinery), 1–14. doi: 10.1145/3313831.3376481

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Nakamura, Verhulst, Sakurada, Fukuoka and Sugimoto. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.