HaptoFloater: Visuo-Haptic Augmented Reality by Embedding Imperceptible Color Vibration Signals for Tactile Display Control in a Mid-Air Image

Rina Nagano (iD), Student Member, IEEE, Takahiro Kinoshita (iD), Shingo Hattori (iD), Yuichi Hiroi (iD), Member, IEEE, Yuta Itoh (iD), Senior Member, IEEE, and Takefumi Hiraki (iD), Member, IEEE
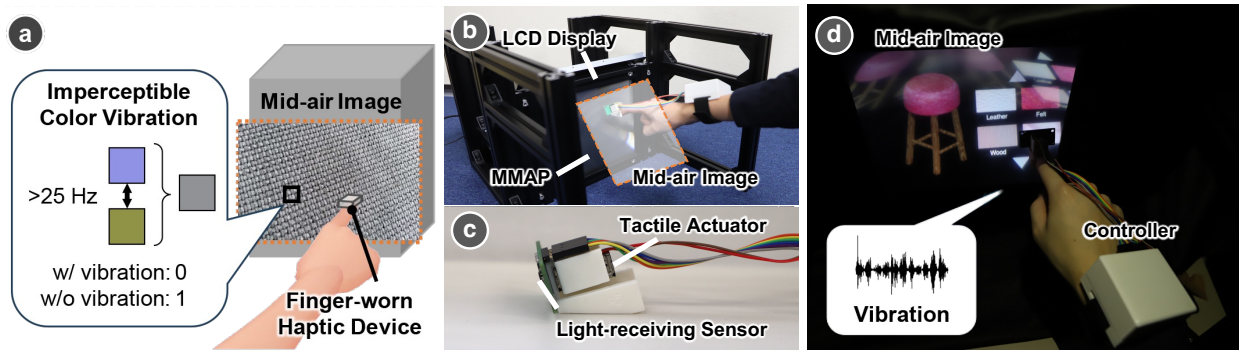

Fig. 1: (a) Conceptual image of HaptoFloater. Haptic vibration control information is embedded in each pixel of the mid-air image using imperceptible color vibration. This light information controls a finger-worn haptic device. (b) Overview of the assembly of a mid-air display. It consists of a high-brightness LCD display and a micro-mirror array plate (MMAP). (c) Overview of the finger-worn haptic device. It has a light sensor for receiving color vibration at the tip and a tactile actuator to present vibration on the side of the fingernail. (d) HaptoFloater displays a visuo-haptic texture image as an application scenario of texture design. The user can not only see the texture of a mid-air image but can also touch it.

Abstract— We propose HaptoFloater, a low-latency mid-air visuo-haptic augmented reality (VHAR) system that utilizes imperceptible color vibrations. When adding tactile stimuli to the visual information of a mid-air image, the user should not perceive the latency between the tactile and visual information. However, conventional tactile presentation methods for mid-air images, based on camera-detected fingertip positioning, introduce latency due to image processing and communication. To mitigate this latency, we use a color vibration technique; humans cannot perceive the vibration when the display alternates between two different color stimuli at a frequency of 25 Hz or higher. In our system, we embed this imperceptible color vibration into the mid-air image formed by a micromirror array plate, and a photodiode on the fingertip device directly detects this color vibration to provide tactile stimulation. Thus, our system allows for the tactile perception of multiple patterns on a mid-air image in 59.5 ms. In addition, we evaluate the visual-haptic delay tolerance on a mid-air display using our VHAR system and a tactile actuator with a single pattern and faster response time. The results of our user study indicate a visual-haptic delay tolerance of 110.6 ms, which is considerably larger than the latency associated with systems using multiple tactile patterns.

Index Terms—Visuo-haptic displays, mid-air images, LCD displays, imperceptible color vibration

---

◆

---

1 INTRODUCTION

Tactile sensation plays a crucial role in how we perceive and interact with the world around us, providing essential information about objects' texture and material properties. In augmented reality (AR), the integration of tactile feedback with visual stimuli to create a Visuo-Haptic AR (VHAR) experience becomes paramount to achieve a high level of realism and immersion [8]. This synthesis aims to allow users to see

and feel virtual objects as if they were physically present, enhancing the perception of object reality through tactile exploration.

To accomplish this, VHAR systems typically consist of two main components: an optical display system to present AR visuals and a physical haptic feedback system for tactile presentation. These systems must work seamlessly to merge haptic information with visual content, providing a coherent and immersive user experience. The challenge lies in ensuring that these tactile and visual cues are perfectly synchronized, enhancing the user's interaction with virtual environments.

However, current implementations of VHAR face significant limitations. The reliance on grounded haptic devices restricts users' movements, confining them to a small area where they can interact with the virtual environment. Moreover, these systems require meticulous calibration to ensure that tactile sensations align accurately with visual presentations [5, 15, 36, 37, 45]. This alignment is critical in maintaining the illusion of reality in VHAR, but it poses practical challenges in terms of the complexity of the setup and user experience.

A promising approach to address these challenges is to use haptic displays that are controlled by light from projection-based visual displays [30]. These VHAR systems expand the user's workspace

• Rina Nagano, Takahiro Kinoshita, Shingo Hattori, and Takefumi Hiraki are with University of Tsukuba. E-mail: nagano@pml.slis.tsukuba.ac.jp, kinoshita@hpcs.cs.tsukuba.ac.jp.
• Takahiro Kinoshita, Shingo Hattori, Yuichi Hiroi, and Takefumi Hiraki are with Cluster Metaverse Lab. E-mail: {s.hattori, y.hiroi, t.hiraki}@cluster.mu.
• Yuta Itoh is with the University of Tokyo. E-mail: yuta.itoh@iii.u-tokyo.ac.jp.

following projection volumes, enabling tactile interactions across a broader area. Notably, they eliminate the need for external tracking systems, enhancing the ease of use and achieving high spatio-temporal consistency between visual and haptic information.

Despite these advancements, projection-based VHAR systems still face their own set of challenges. One significant limitation is that tactile interactions are confined to the real surfaces on which the images are projected. This constraint means that while the interaction area is expanded, it remains bound to the projection surface, limiting the versatility and realism of the haptic feedback in more complex virtual environments.

However, mid-air imaging technology, which allows for the display of full-color images in real space without the limitations of physical surfaces, is garnering interest. This technology enables the overlay of computer graphics on the real-world view without necessitating devices like head-mounted displays (HMDs) to be worn around the eyes. The capability of mid-air imaging to enhance the visualization and entertainment of real-world information makes AR systems that utilize it a key player in these fields [4, 42].

An inherent challenge with mid-air image displays is the lack of tactile feedback, as the images are intangible. To address this, some researchers have explored methods to integrate tactile feedback with mid-air images through additional devices. However, visual and haptic representations in these systems are often managed through uncoordinated processes that rely on external tracking systems, such as infrared sensors, to function. This reliance frequently leads to delays and a mismatch in the positioning between visual and haptic information due to the processing and synchronization requirements of the sensing technologies. Consequently, there has yet to be a VHAR system capable of delivering diverse texture presentations that are spatio-temporally consistent and unrestricted by physical display surfaces.

This paper presents HaptoFloater, a Visuo-Haptic Augmented Reality (VHAR) system that integrates tactile feedback directly into the light of a mid-air image display (Fig. 1a). HaptoFloater can render visual and haptic information both independently and simultaneously while ensuring high spatio-temporal consistency between these two modalities. The system employs imperceptible color vibrations to control wearable haptic devices, embedding undetectable control signals to users within the projected image. This technique allows for embedding control information on a pixel-by-pixel basis without altering the image's visual quality, ensuring that the displayed images remain comfortable for human viewing.

The haptic displays are equipped with photo sensors that detect the control signals embedded in the image's light, enabling the presentation of vibrotactile stimuli corresponding to each color channel's vibration patterns. This direct transmission of control signals through the image light allows the haptic feedback to be synchronized with the visual display with virtually no perceptible delay. Through this approach, HaptoFloater solves the previously mentioned challenges, seamlessly integrating tactile and visual information in mid-air displays.

We implemented a proof-of-concept prototype of our HaptoFloater system consisting of a mid-air image display system and a finger-worn haptic device as shown in Fig. 1b. We use four vibration symbols that are suitable to use for visual applications to embed haptic control information, which is embedded into the mid-air image generated by a micro-mirror array plate (MMAP). Within the finger-worn device (Fig. 1c), a photodiode reads this color-vibration information and a tactile actuator presents the corresponding tactile information. With this configuration, our system provides a correspondence between the optical information embedded in the video and the position of the fingertip, and can present tactile information with low latency and positional consistency.

Through delay evaluation experiments in our prototype, we confirmed that our system can present visuo-haptic information with a delay of less than 100 ms, which is the reported tolerance of delay perception from vision to haptics [29, 30, 38]. This indicates that the HaptoFloater can also be used as an experimental platform to investigate human perceptual characteristics of visual-haptic perception of mid-air images. We evaluated the processing time (latency) of the

finger-worn device from the time it receives the light of a displayed mid-air image to the tactile presentation. In addition, we conducted user studies to determine whether the visual-haptic asynchrony (latency) of the system is acceptable to users. Finally, several applications in daily scenes were presented (Fig. 1d).

The contributions of this paper are that we

- developed a system that can superimpose tactile presentations on mid-air images by embedding imperceptible color vibration signals in the images and detecting these signals with a fingertip device.

- conducted the first user study of visual-haptic latency perception characteristics in mid-air displays using the developed system.

- developed three VHAR applications, an interactive touch panel, a texture design support system, and an interactive digital museum.

2 RELATED WORK

This section describes previous VHAR systems that have been specifically designed for visuo-haptic texture presentation. First, we provide an overview of existing haptic devices in VHAR and their presented limitations (Sec. 2.1). We then explain the principles of imperceptible color vibration, the method we use to embed signals for the control of the haptic device (Sec. 2.2). We also review the interface system that combines mid-air image displays and haptic devices, and discuss its potential application as a VHAR system (Sec. 2.3). Finally, we discuss guidelines for latency from haptic to visuo-haptic information (Sec. 2.4).

2.1 Haptic devices in VHAR

An early approach to haptic presentation was the use of grounded haptic displays such as PHANToM [40]. To use these for VHAR applications, they are usually combined with a visual display such as a half-mirror [15, 44] or a head-mounted display [10, 37]. These systems allow tactile information to be presented on a virtual image. The main limitations of these systems were that the use of grounded tactile devices limited the user's workspace to a small area that could be manipulated and required a cumbersome calibration process to align the tactile presentation with respect to the visual presentation.

One way to overcome these limitations is to integrate a haptic display into a flat panel visuo-haptic display [2, 6]. This system ensures the spatio-temporal consistency of the visuo-haptic display by integrating touch-sensing capabilities into the display. However, even in such visuo-haptic touch panels, the workspace is limited to a physical plane. In contrast, wearable haptic displays, such as finger-worn devices [33], styluses [21, 22], and armbands [13, 34], enable tactile presentation without restricting the user to a small workspace. These wearable displays can be used for both virtual object interaction [23, 27] and tangible interface interaction [3,7], but require external tracking systems and tend to introduce perceptible spatiotemporal errors.

Another approach to overcome the above limitations is to combine a projection-based visual display with a wearable haptic display controlled by projected light. HALUX [43] is a wearable haptic display that uses the luminance of projected light to control a vibrating actuator. In this display, the tactile actuators are controlled to move when the projected light from the projector hits them and to stop when it does not. This display uses the projected light to enable the presentation of tactile sensations, but does not present any visuo-haptic content to the human viewer, as the projected light is only used to control the tactile display. SenseableRays [35] use structured light to control the vibration of a vibrating actuator. The signal of visuo-haptic information is converted into a signal of tactile information, thereby presenting vibrotactile sensations. The amplified received light is used as a driving signal for the actuator. Since this system does not require an external tracking system, the tactile device can be made smaller and the latency of tactile presentation can be reduced. The disadvantage of this system is that the visuo-haptic and haptic information cannot be designed independently. In other words, the visuo-haptic information automatically determined the visual information because of the direct
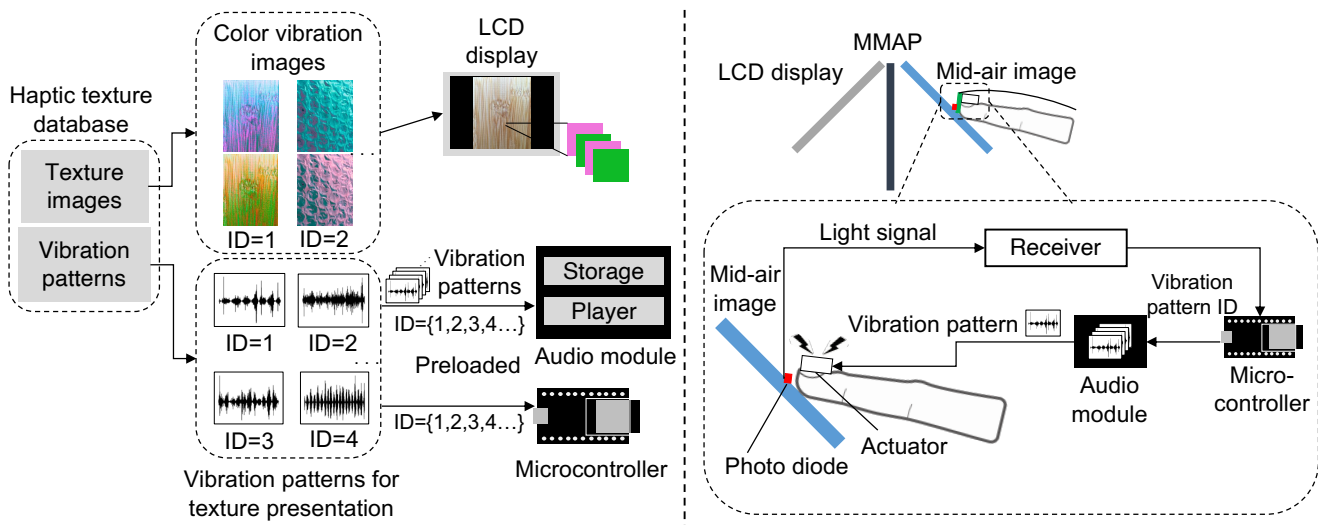
Fig. 2: System principle of HaptoFloater. (Left) Processing flow of visual and tactile information. (Right) Schematic illustration of the tactile feedback mechanism.

conversion between the projected light and the vibration pattern. HaptoMapping [30] is a VHAR system that controls haptic devices using pixel-level visible light communication (PVLC), which embeds imperceptible information in the projected image. PVLC is a technology that can embed luminance modulation information into images in a way that is invisible to humans by using a high-speed projector [19]. By using PVLC, this system enables independent visual and haptic design and consistent simultaneous visual and haptic presentation. However, this system had the disadvantage that it could only work when projected images designed to embed information were projected using a special high-speed projector.

Our VHAR system uses the imperceptible color vibration [1] to embed information into the displayed image. The imperceptible color vibration can be realized on LCD displays with a typical refresh rate of 60 Hz because the flicker frequency requirements for human perception are less stringent than those for luminance modulation. This allows independent design of visual and haptic sensations and consistent simultaneous visual and haptic presentation using a commercial LCD display.

2.2 Color vibration for embedding information

Efforts to embed information in images using color information have been studied mainly in the area of screen-camera interaction, while there are applications such as digital watermarking [50]. THAW [24] tracks a smartphone with a gradient-colored 2D pattern on the display that indicates the position of the smartphone when captured by the camera. CapCam [47] detects cameras with capacitive touchscreens and pairs the display with the camera by sending data as a flashing color pattern. VRCodes [46] is a technique that uses a rolling shutter smartphone camera that switches between complementary colors. While these methods are capable of unobtrusive data transmission, they can only output in grayscale, and the codes are visible, so they are embedded exclusively in the background of the content.

Imperceptible color vibration is a method that achieves the embedding of information by switching colors with constant luminance at a rate imperceptible to humans [1, 48]. This method takes advantage of the human visual property that observers perceive two intermediate colors when two different colors of the same luminance are rapidly alternated [16]. It is known that the critical color fusion frequency (CCFF), the frequency at which humans cannot perceive color vibration, is about 25 Hz, which is about half the critical fusion frequency (CFF), the frequency at which luminance flicker becomes imperceptible. Since the refresh rate of commercially available LCD monitors and projectors is 60 Hz or higher, the images displayed on these devices

can embed information through color vibration. The system can then send control signals to the device while displaying video content to humans. We exploit this imperceptible color vibration to implement a low-latency, low-displacement VHAR system as a mid-air display.

2.3 Tactile feedback on mid-air image displays

Mid-air imaging has attracted attention as a technique for displaying full-color images in real space without being constrained by physical surfaces. Several specialized optical systems have been proposed to realize such displays, including roof mirror arrays [25], retro-reflective mid-air imaging [41], and MMAP [26]. In particular, MMAP can display bright images and is easy to acquire and install, so MMAP-based AR systems play an important role in real-world information visualization and entertainment [18, 20].

In mid-air image display, there is usually no haptic feedback because the images are intangible. Therefore, research has proposed the use of ultrasonic phased arrays to add haptic feedback to mid-air images [28, 31, 49]. This phased array can be shaped so that the ultrasonic waves converge at this point. This would create pressure at a specific point in the air, and if a human finger were at that point, it would be possible to present a tactile stimulus without contact.

However, a drawback of ultrasonic phased arrays is that the frequency and intensity of the pressure can only be controlled to a very limited extent compared to tactile actuators. As a result, it is difficult to produce complex tactile sensations such as texture, although there have been efforts to expand the range of tactile sensations using ultrasound [9, 32]. In addition, because the visuo-haptic and haptic presentations in these systems operate in processes that are not coordinated with each other, they require external tracking systems such as infrared sensors. Therefore, the processing and synchronization caused by this sensing often results in latency and positional mismatch between visual and haptic information. The delay time resulting from communication between an external tracking system and a PC is generally said to be about 100 ms or more, and light control-based devices, which in principle do not incur this delay, are supposed to have an advantage for the human interface [14].

We propose a VHAR system that is spatio-temporally consistent and capable of presenting a wide variety of textures by controlling a wearable haptic display by embedding control signals in the light itself, which is composed of mid-air images.
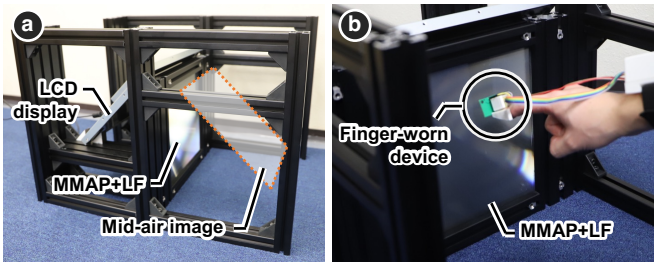
Fig. 3: Mid-air display hardware of HaptoFloater. (a) Images on the LCD display screen appear in the air, symmetrical to the MMAP and the vision control film, called Louver Film (LF). (b) Touching the mid-air image with a finger-worn haptic device.
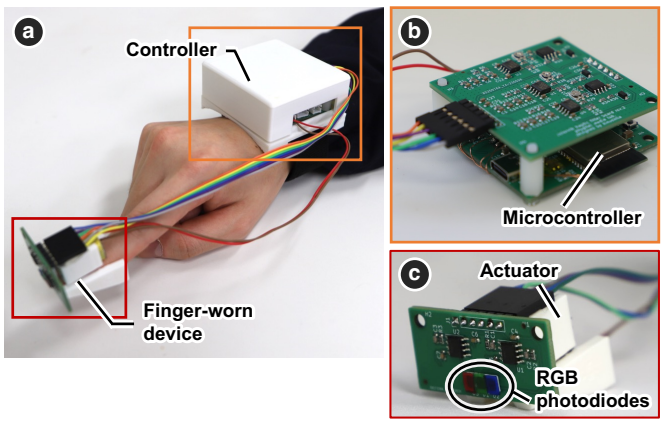

Fig. 4: (a) Hardware configuration of HaptoFloater consisting of a finger-worn haptic device and a controller unit worn on the arm. (b) The controller unit consists of an audio module that transmits tactile vibrations and a microcontroller. (c) The finger-worn haptic device consists of an actuator and light-receiving sensor board using RGB photodiodes.

2.4 Human perception characteristics for visuo-haptic latency tolerance

When designing a VHAR system, the presented visuo-haptic stimuli must be spatio-temporally consistent to enhance the reality of the content. Therefore, we need to know the threshold of delay perception in visuo-haptic presentation. With regard to delay perception, Miyazato et al. [29] investigated the acceptable latency range from visual to haptic sensation in a visuo-haptic teleconferencing system. A user study showed that the threshold for delay perception in their system was about 100 ms. Silva et al. [38] also investigated the acceptable range of perceivable latency of visuo-haptic sensation in video games. They found that the stimulus threshold for latency in their system was approximately 100 ms. Miyatake et al. [30] also investigated the acceptable range of visuo-haptic latency in a VHAR system with a tabletop projection display using finger-worn, stylus, and arm-worn tactile devices. The results showed that the latency stimulus thresholds were approximately 100, 159, and 500 ms for finger-worn, stylus, and arm-mounted devices, respectively.

Previous studies have focused on investigating the tolerance of visuo-haptic latency in visuo-haptic virtual reality systems or VHAR systems that involve contact with physical surfaces. Thus, there are still no design criteria for the consistency of visuo-haptic information in mid-air VHAR systems. In addition, the latency threshold in VHAR systems has only been studied when the haptic device is switched from off to on. Therefore, the stimulus threshold of latency in the VHAR system in both the on and off directions has not been investigated, and it would be interesting to understand the difference between them. In this paper, we report the results of an experiment on the temporal latency of visuo-haptic latency by superimposing visual and haptic sensations on a mid-air image display using the prototype system we developed.

3 HAPTOFLOATER

Figure 2 shows an overview of our HaptoFloater system. The system comprises three components: a mid-air image display (Sec. 3.1), color vibration images (Sec. 3.2), and a finger-worn haptic device (Sec. 3.3). The following sections describe each component in detail. We then describe how to estimate the latency generated by this system (Sec. 3.4).

3.1 Mid-air image display

Figure 3 shows the implementation of our mid-air image display. We implemented a mid-air image display using an LCD (ULF1505-IX, LITEMAX, 15" diagonal, 1000 cd/m^2 brightness), a micro-mirror array plate (MMAP) (ASKA3D-250NT, ASKANET, 250 mm × 250 mm), and a vision control film (WINCOS Vision Control Film W-0055, LINTEC).

An MMAP behaves as a transmissive mirror that can form an image at a plane-symmetric position to the optical element. The conceptual fabrication method for an MMAP is as follows: consider a series of mirrored glass (or acrylic) plates that are stacked on top of each other and then cut vertically into thin sheets. Given these two sheets, an MMAP can be obtained by rotating one sheet 90° along an axis perpendicular to its face and then gluing the two sheets together.

The vision control film, also called Louver Film (LF), is a unidirectional opaque film with opacity angles ranging from 0° to 55°. It is used to shield the light emitted by the LCD and transmitted through the MMAP from the user's eyes. In our mid-air display prototype, we placed the LCD at 45° to the MMAP and the LF on the opposite side of the LCD, overlapping the MMAP.

3.2 Generating color vibration image

Color vibrations alternately present two point-symmetric colors for the target color in the L*a*b* color space at constant luminance. Color pairs are quickly searched for using the method of Hattori et al. [11] to generate an array of multiple color pairs. Then, color pairs are extracted from the array that satisfies both the color vibration condition, which is imperceptible to humans, and the embedded signal condition, which allows the device to detect the vibration. These color pairs are embedded into pixels as time signals to obtain a color vibration image.

According to [11], color pairs can be generated with 9 representative colors (black, gray, white, red, green, blue, cyan, yellow, and magenta), which constitute color vibration. Furthermore, by generating color pairs using three RGB channels, for each representative color, at least 3 types of color pairs can be obtained, and with several of representative colors, 5 types of color pairs can be obtained. This means that, theoretically, at least 3 different vibro-tactile sensations can be presented in the representative colors, and 5 different types of vibro-tactile sensations with several of those colors. However, in terms of implementation, we found that vibration using the G channel sometimes reproduces a grayish color. For this reason, we generated color pairs using the R and B channels, i.e., we generated 3 color pairs in which the color vibration is (on, off), (off, on), or (on, on) in the R and B channels. These color pairs can be generated with 9 representative colors.

3.3 Finger-worn haptic device

Figure 4 shows our implementation of a fingertip device for tactile presentation. The sensor and actuator are controlled by a microcontroller worn on the arm (Fig. 4b). The fingertip device (Fig. 4c) consists of a light-receiving sensor for detecting color vibrations and a tactile actuator. The implementation details of each component are described below.

3.3.1 Light-receiving sensor

We implemented a light-receiving sensor circuit to receive and detect color vibrations using a photodiode and an operational amplifier (op amp).

Figure 5 shows a diagram of the light-receiving sensor. The signal from each photodiode (PD) first enters a transimpedance amplifier
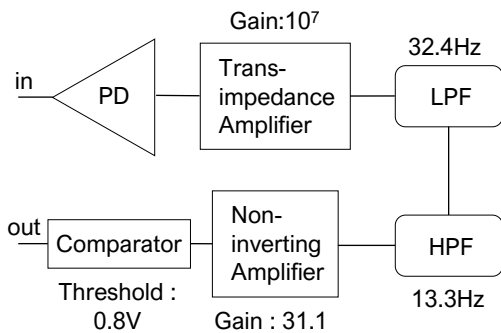
4

Fig. 5: Configuration diagram of light-receiving sensor.

circuit, where it is converted from current to voltage. Then, to extract only the 30 Hz color vibration signal with high sensitivity, the signal was passed through a low-pass filter (LPF) and a high-pass filter (HPF) in sequence. The LPF isolates the 50/60 Hz noise from the AC power supply and the ambient light from the signal. We implemented a 32.4 Hz Sallen-Key 4th order active LPF in the circuit. The signal then passes through a 13.3 Hz HPF and a non-inverting op-amp to match the phase of the input and output signals. Finally, the signal passes through the comparator to convert it to a square wave, making it easier for the microcontroller to process.

We used three photodiodes (NJL6407R, JRC) covered with an RGB color filter (NRS-1635, Nissho) as RGB photodiodes, an op-amp (AD822ARZ, Analog Devices, Inc.) for the trans-impedance amplifier, another op-amp (NJM4580E, JRC) as both the LPF and the non-inverting amplifier circuit, and a comparator (LM393DT, Texas Instruments). We also designed a CR-HPF circuit ourselves.

3.3.2 Tactile actuator

We use HAPTIC™ Reactor (Alps Alpine) for a tactile actuator embedded in the finger-worn haptic device. HAPTIC™ Reactor is a special type of linear resonant actuator (LRA), and it can convert sound transmitted through an audio amplifier into tactile vibrations, providing more complex tactile sensations over a wider frequency range than conventional LRAs.

The actuator is connected to an audio amplifier (MAX98357A, Analog Devices), and the amplifier is connected to the microcontroller (ESP32-WROOM-32E, Espressif Systems) via Inter-IC Sound (I^{2}S) interface (Fig. 4b).

The microcontroller detects texture by the presence or absence of vibration in each of the three RGB channels. To elaborate, each texture is represented by a unique identifier, which is associated with the pattern of presence or absence of vibration in each of the three RGB channels. The pattern of presence or absence of vibration indicates, for instance, that there is vibration in the case of R and no vibration in the case of G and B. Thus, by detecting the presence or absence of vibration in each of the three RGB channels, texture can be detected and identified. However, since two channels, R and B, are used in this implementation, the texture is detected and identified by the pattern of presence or absence of vibration in each of the two RB channels. In this way, the microcontroller inputs the corresponding ID information to the audio module based on the color vibration signals received from the light-receiving sensor circuit. The vibration signal is stored as an audio file, and the tactile vibration presented is controlled by determining the audio file to play by the microcontroller.

For this study, texture images and the corresponding tactile vibration information (audio files) were obtained from the LMT Haptic Texture Database [39], a tactile data set.

3.4 Latency estimation

The total latency of our system T_{total} is defined as the duration from the time when the haptic device is placed on a mid-air image with embedded control signals to the time when the device presents a vibration.

Table 1: Worst-case (Maximum) latency of the haptic from visual sensations using the finger-worn haptic device in the turn-on and turn-off conditions.

	T_{total}	=	T_{recv}	+	T_{vib}
turn-on	59.5 ms		16.7 ms		42.8 ms
turn-off	46.5 ms		16.7 ms		29.8 ms

T_{total} can be obtained by

$$T_{total} = T_{recv} + T_{vib}. \tag{1}$$

The details of each duration are as follows:

T_{recv} is the time from when the sensor enters the image area in which the color vibration is embedded to when the microcontroller detects the color vibration, which takes a value from 0 to half of the color vibration period, i.e., $1/(30\,\text{Hz} \times 2) = 16.7$ ms.

T_{vib} is time from when the microcontroller detects the color vibration to the time the actuator begins to vibrate. This is the sum of the processing time of the microcontroller and the audio amplifier and the mechanical response of the actuator.

4 EVALUATION

We conducted two experiments with our system. First, we measured the worst-case latency between visual and tactile sensations in our system (Sec. 4.1). Then, we evaluated the threshold at which users perceive latency between visual and tactile sensations (Sec. 4.2). Finally, we evaluated the reality of the tactile sensation reproduced by our device (Sec. 4.3).

The protocol was approved by the Ethical Review Committee of the University of Tsukuba Library, Information and Media Science (Registration number: 24-30) and Cluster, Inc. Research Ethics Committee (Registration number: 2023-004), and informed consent was obtained from all subjects.

4.1 Worst-case latency of haptic device

We estimate the worst case of T_{total} with our proposed finger-worn haptic device. The measurements were taken under conditions in which the actuator was switched from off to on (called "turn-on" condition) and from on to off (called "turn-off" condition). We measured T_{vib} using a μs-order precision timer inside the microcontroller. In the turn-on condition, the timer started when the microcontroller detected the color vibration and stopped when the actuator started to vibrate. In the turn-off condition, the timer started when the microcontroller stopped detecting color vibration and stopped when the actuator ceased vibration. The onset of vibration was detected by an acceleration sensor (KXR94-2050, Kionix) attached to the fingertip side of the device. We performed the measurements 100 times using a microcontroller for each condition in the turn-on condition and turn-off condition.

Table 1 summarizes the maximum T_{recv} and T_{vib} for each condition, and the maximum T_{total} as the sum of these latencies. In the turn-on condition, the mean value of T_{vib} was 42.8 ms with a standard deviation of 0.47 ms. In the turn-off condition, the mean value of T_{vib} was 29.8 ms with a standard deviation of 0.82 ms. Since maximum $T_{recv} = 16.7$ ms (Sec. 3.4), the worst case for the total system latency T_{total} is 59.5 ms and 46.5 ms for the turn-on and turn-off conditions, respectively.

The method using PVLC has a delay of 96.3 ms [30], thereby indicating that our method has a shorter delay time.

4.2 User study on latency perception

We conducted a user study to investigate whether users would notice a latency between visual and tactile sensations while using the system.

4.2.1 Experimental setup

In this experiment, we investigated the latency threshold between visual and haptic perception by introducing a delay time in the microcontroller.
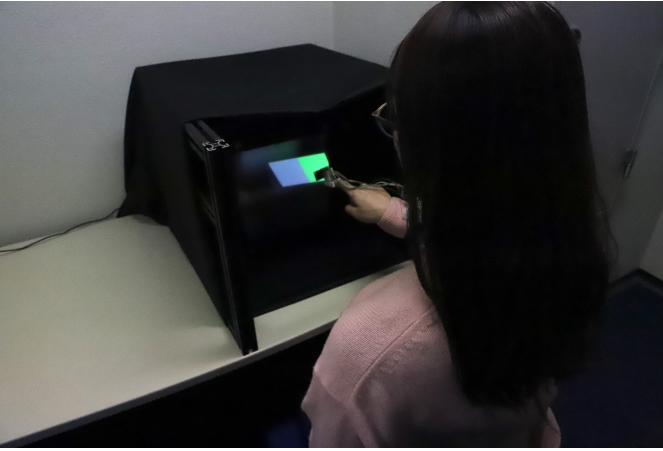
Fig. 6: Experimental situation during the user study on latency perception. Although the figure shows the room with the lighting on for clarity, the actual study was conducted with dimmed lighting.

Fig. 7: Percentages of positive answers in the experiment for a threshold time of a visual-haptic delay in the turn-on condition. The delay time perceived by users with 50% probability was found to be approximately 110.6 ms.

The image size of the mid-air image is 150 mm × 70 mm and is lined with gray and green regions side by side.

We conducted two types of perceptual threshold assessments: perceptual threshold for delay when changing from no tactile information to providing tactile information (turn-on condition) and vice versa (turn-off condition). By incorporating a sleep time on the microcontroller in addition to T_total, the delay time could be adjusted to ensure that the system correctly provided tactile feedback after the delay time specified by the control signal (≥ 59.5 ms in the turn-on condition and ≥ 46.5 ms in the turn-off condition).

4.2.2 Participants and experimental methods

14 people (7 males and 7 females, aged from 20 to 24, 12 right-handed and 2 left-handed) participated in this user study. Participants wore the haptic device on the index finger of their dominant hand. We prepared eight variations of latency (60, 80, 90, 100, 110, 120, 130, and 150 ms) for each of the turn-on and turn-off conditions, and these latencies also include device-dependent latencies. Figure 6 shows the experimental situation. During the turn-on condition, participants were instructed to move their hands from the gray area to the green area and report whether they perceived a delay corresponding to the tactile sensation from the moment they visually confirmed that their fingertips had entered the green area. In the turn-off condition, participants were instructed to move their hands from the green area to the gray area and report whether they perceived a delay from the moment they visually confirmed that their fingertip had entered the gray area until the tactile presentation ceased. To control the speed of the hand movements, the experiment participants were shown the experimenter moving her hand at a speed of 150 mm/s before the commencement of the experiment. Furthermore, a metronome was played at a rate of 60 beats per minute throughout the experiment, and participants were instructed to trace the image from one end of the area to the other. Participants rated 8 × 2 = 16 conditions per trial, with each trial repeated 8 times for a total of 128 ratings. The order of the conditions was randomized per trial and per participant.

4.2.3 Results

Figure 7 and Fig. 8 show the averaged percentage of positive answers obtained in the user study for each of the turn-on and turn-off conditions, respectively. The bars represent the standard error of the mean, and the curve is fitted using the sigmoid function defined as follows:

$$y = \frac{1}{1 + \exp(-k(x - x_0))} \times 100. \tag{2}$$

The fitting resulted in parameter values of $k = 0.0316$ and $x_0 = 110.64$ in the turn-on condition, and $k = 0.0346$ and $x_0 = 112.33$ in the turn-
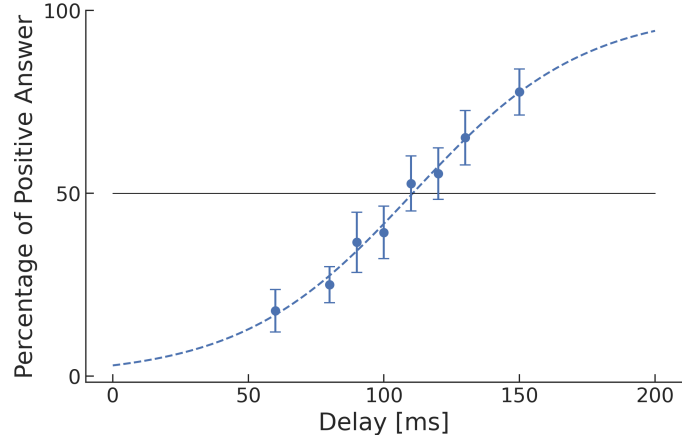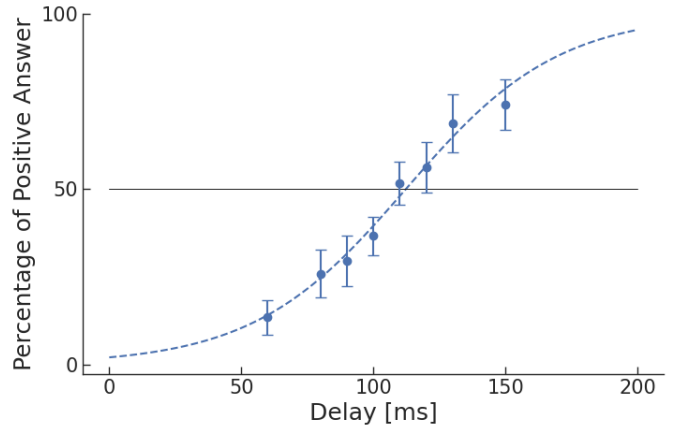

Fig. 8: Percentages of positive answers in the experiment for a threshold time of a visual-haptic delay in the turn-off condition. The delay time perceived by users with 50% probability was found to be approximately 112.3 ms.

off condition, which were used in the calculations. We identified the threshold values (T_th) for the acceptable delay time for visual and tactile perception. This was defined as the time at which the user perceived a delay with a 50% probability. The results indicated T_th in the turn-on and turn-off conditions are approximately 110.6 ms and 112.3 ms, respectively. In previous studies, a visual-haptic delay tolerance for users viewing a virtual finger in a display image using a finger-worn force feedback device [38] and the delay tolerance when viewing the user's own fingertip on a tabletop display using a finger-worn haptic device [30] have been reported, both of which are approximately 100 ms. These values are about 10 ms less than the latency tolerance T_th obtained for the mid-air display. This suggests that the latency tolerance is slightly looser than in the previous studies because VHAR with mid-air images involves moving the finger in the air, where there are no physical cues.

Since these latency thresholds are significantly larger than T_total, the proposed finger-mounted haptic device meets the requirement of having a latency that does not allow the user to perceive the visuo-haptic latency. Thus, our system maintains temporal consistency between visual and haptic modalities. This measured threshold can also be used as a design criterion for temporal consistency when considering other configurations of our system, such as the use of different actuators.

6

Fig. 9: Experimental situation during the user study on visuo-haptic texture perception. The participant touches the mid-air image and rate how realistic the material was.

4.3 User study on texture perception

We conducted a user study to investigate how the user's perception of the reality of the texture varies depending on the presence and appropriateness of the tactile presentation when the proposed VHAR system is used to display visuo-haptic textures.

4.3.1 Experimental setup

In this experiment, we investigated how the user's perception of texture reality changes when the type of tactile sensation is changed in a visuo-haptic texture presentation. The mid-air images were 120 mm × 70 mm in size, and five different textures (Brick, Cork, Sheep Skin, Paper, Jeans) were selected from the LMT Haptic Texture Database [39]. In selecting the textures, we referred to the "perceptual similarity between all our surface materials" information in the database and select five textures from each of the nine categorized groups. We used three types of tactile presentation: no tactile sensation (only a texture image), a single frequency (150 Hz) vibration, and a texture vibration. For texture vibration presentation, we used audio files provided by the LMT Haptic Texture Database.

The audio files from the database were normalized so that the maximum/minimum values were ± 8192, which is one quarter of the maximum/minimum recorded value in a 16-bit wav file. In addition, the amplitude of the single frequency audio file was also generated to be 8192. This process was done to match the volume levels of the database audio file and the single-frequency audio file. We did not include a sleep time in the microcontroller in this experiment, so the latency of the visuo-haptic latency is expected to be T_{total}.

4.3.2 Participants and experimental methods

Figure 10 shows the experimental result.

12 participants (7 males and 5 females, aged 20 to 24 years, 12 right-handed) participated in this user study. Participants wore the tactile device on the index finger of their dominant hand. We prepared five textures, as described in Sec 4.3.1. Figure 9 shows the experimental situation. Participants experienced the texture presentation to their satisfaction by moving their hands in the mid-air area where the texture image was displayed. After each experience in a single condition, participants answered the question "How realistic was the texture material?" on a 7-point Likert scale (1: not at all, 7: very much). Participants evaluated 3 conditions × 5 textures = 15 conditions. The order of texture presentation was randomized per participant and the order of presentation conditions was randomized per texture.

4.3.3 Results

Figure 10 shows user ratings of texture tactile reality on a 7-point Likert scale, where image only is when no tactile sensation is presented, 150

Hz sensation is when a sine wave is presented, and texture sensation is when texture tactile sensation is presented. For Image Only, 150 Hz Sensation, and Texture Sensation, in that order, the mean and standard deviation are as follows: For brick, the mean and the standard deviation were (1.83, 1.14), (2.17, 1.14), and (4.42, 1.19), respectively. Similarly, for cork, they were (1.75, 0.83), (2.42, 1.32), and (4.50, 1.38); for sheep skin, they were (2.00, 1.15), (2.08, 0.86), and (4.50, 1.32); for paper they were (1.75, 0.92), (2.00, 0.91) and (4.33, 1.60); for jeans they were (1.42, 0.49), (2.75, 1.42) and (4.17, 1.28).

We conducted a Shapiro-Wilk test on each of the data sets to ascertain the normality of the data. The results demonstrated that none of the experimental data with the five samples exhibited normality. Therefore, a Friedman's test, a non-parametric test, was performed. The results indicated that there were significant group differences in all experiments ($p < .001$). A Wilcoxon signed-rank test (Shaffer-corrected) was then performed on all combinations of the two groups as a subtest. The results showed that the texture vibration condition was significantly higher than the no tactile sensation condition ($p_{Brick} = .003$, $p_{Cork} = .002$, $p_{Sheep} = .009$, $p_{Paper} = .003$, $p_{Jeans} = .006$) and significantly higher in the texture vibration condition than in the single frequency vibration condition ($p_{Brick} = .006$, $p_{Cork} = .006$, $p_{Sheep} = .003$, $p_{Paper} = .009$, $p_{Jeans} = .045$).

Furthermore, for the Jeans texture only, the condition presenting single-frequency vibration was significantly higher than the condition not presenting tactile sensation ($p_{Jeans} = .011$), but no significant differences were observed for the other textures.

Thus, we found that presenting texture vibration is an effective method for presenting realistic visuo-haptic textures.

5 APPLICATIONS

We have developed three potential applications that use HaptoFloater.

5.1 Mid-air touch panel with multiple tactile feedback

Figure 11a shows a user interface system that allows users to experience a variety of tactile feedback while retaining the benefits of a mid-air touch panel, providing a richer experience. Specifically, the application allows users to feel tactile stimuli of different textures depending on where they touch when operating a button. Because this system does not limit the number of people using the system and is not a physical touch screen, it reduces the risk of infection from sharing a touch screen with multiple unidentified users, such as in public places. In addition, since no fingerprints are left on the screen, secure button operation can be achieved. While retaining the benefits of the mid-air touchscreens, this system provides a richer experience with a variety of tactile feedback.

5.2 Texture design support system

Figure 11b shows a texture design support system for physical surfaces. In this application, the signal is embedded in the texture image, and the corresponding tactile pattern is preloaded into the audio module. As the user touches and explores the surface, the tactile device selects the tactile pattern corresponding to that location and presents it to the user. It is hoped that users will benefit from this system when searching for desired textures from a group of samples for a new product. Since the texture is presented only as a mid-air image, the user can easily try different textures by changing the image and vibration pattern. This system will help the user to test textures and find the right texture without scattering physical texture samples all over the workspace.

5.3 Tactile presentation in museums

Figure 11c shows an example of an application for tactile presentation of exhibited objects in a museum. The mid-air image can present information associated with the real object while simultaneously viewing the real object beyond the image. By adding tactile information to the mid-air image in addition to visual information, the user can experience the tactile sensation of the texture of the exhibited object, which cannot actually be touched.
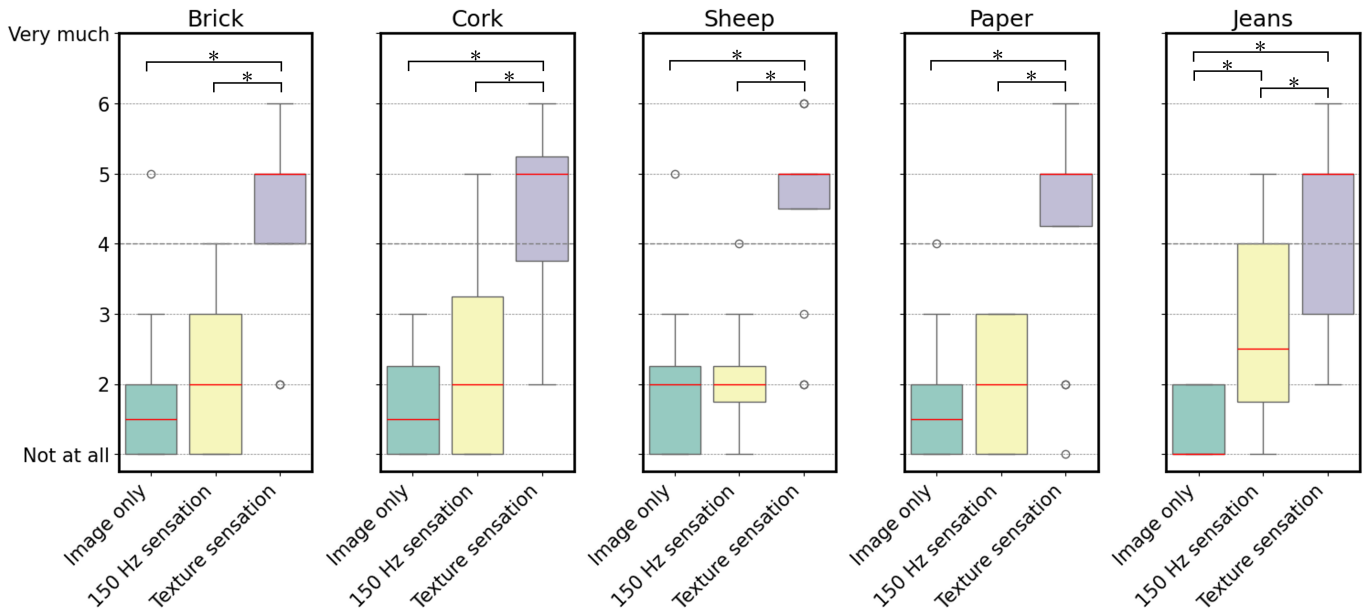
Fig. 10: User ratings of texture tactile reality on a 7-point Likert scale, where image only is when no tactile sensation is presented, 150 Hz sensation is when a sine wave is presented, and texture sensation is when texture tactile sensation is presented.
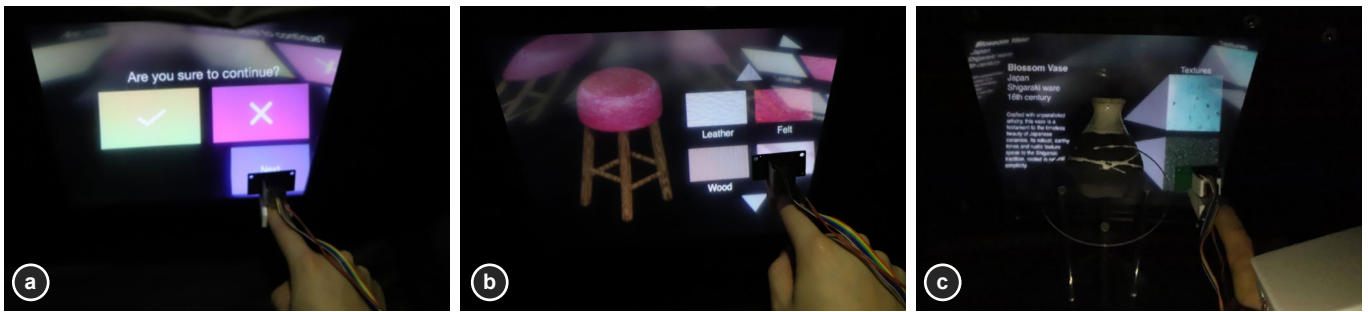

Fig. 11: (a) Mid-air touch panel with multiple tactile feedback. When the user touches a floating button with a finger close to it, they immediately perceive a different click feeling (tactile feedback) for each button. (b) Texture design support system. The user can see the visual texture as a mid-air image and experience the corresponding tactile feedback at precisely the appropriate timing and position. (c) Tactile presentation in museums. The vase the user is looking at is a real object, and the mid-air image is presenting VHAR from above it. The user can confirm the texture of the vase by touching a magnified mid-air image of the vase with a fingertip device.

6 DISCUSSION

6.1 Perception of image flicker

In the experiment, two-color images were displayed alternately at 60 Hz, resulting in color vibration at 30 Hz. Color fusion occurs at this 30 Hz color vibration because the critical fusion frequency of colors in the human eye is about 25 Hz [17]. Furthermore, in this paper, we used a La*b* color space designed to approximate human vision, based on the previous work [11], to produce color vibration images with stabler values of L, the luminance component. Therefore, it is expected that humans will not perceive the flicker.

However, in our experiments, flickering of the color vibration image was sometimes observed. The experimental results of the study [12] indicate that approximately 10% of users may experience flicker caused by the color pairs utilized in the user experiments on latency perception. Although the luminance of the two color vibration images was adjusted equally, the perception of flickering is believed to be caused by differences in the critical fusion frequency of the colors depending on the environment of the light source, the type of color, and the individual. Therefore, subject experiments should be conducted for each type of color.

On the other hand, many LCDs with refresh rates higher than 60 Hz, such as 120 Hz and 144 Hz, are now commercially available, and mLED displays that achieve higher refresh rates by representing gradation through high-speed LED blinking are also becoming more popular. Using these devices, it seems possible to solve this issue by embedding information in images by means of color vibration faster than 30 Hz.

6.2 Spatial extent of user location detection

The precise moment at which a specific mid-air image is deemed to have been touched is when the fingertip sensor receives light from the mid-air image. For example, in the case of the mid-air touch panel with multiple tactile feedback in Applications, the button is considered to have been touched when the fingertip sensor receives light from the mid-air image. Due to the nature of the mid-air image display, the light is received in the vicinity of the mid-air image plane. However, from an interaction perspective, it is necessary to determine the exact moment when the button is pressed. Therefore, it is necessary to ascertain the detection performance at different positions in space. It is currently understood that, due to the nature of mid-air image displays, light is more readily received at the rear of the mid-air image and less so at the front, as light tends to diverge. Nevertheless, future investigations should be conducted to ascertain the detection performance with respect

to light, depth, and other relevant factors.

7 CONCLUSION

In this paper, we propose HaptoFloater, a novel visual-haptic mid-air image display that can eliminate visual-haptic delay perceived by the users. We constructed an mid-air image display by combining an LCD and a micro-mirror array plate (MMAP), and embedded tactile control information in this image using invisible color vibration. We also implemented a finger-worn haptic device using a tactile actuator, a light-receiving sensor board, and a microcontroller, and used the color vibration information to control this device. We evaluated the performance of the finger-worn haptic device and showed that it could present tactile vibrations to the user with a delay of 59.5 ms after the fingertip entered the region of a mid-air image. We conducted the user study and found that the visual-haptic delay tolerance for a finger-worn haptic device in a mid-air display is 110.6 ms. From this result, we can conclude that our proposed finger-worn haptic device has a practical performance in visual-haptic mid-air displays. As future work, we will investigate the visual-haptic delay tolerance for other forms of haptic devices, such as stylus-type devices, as well as finger-worn device. In addition, we can explore more application scenarios using HaptoFloater.

REFERENCES

[1] S. Abe, A. Arami, T. Hiraki, S. Fukushima, and T. Naemura. Imperceptible Color Vibration for Embedding Pixel-by-Pixel Data into LCD Images. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17*, pp. 1464–1470, 2017. doi: 10.1145/3027063.3053074 3

[2] C. Basdogan, F. Giraud, V. Levesque, and S. Choi. A Review of Surface Haptics: Enabling Tactile Effects on Touch Surfaces. *IEEE Transactions on Haptics (ToH)*, 13(3):450–470, 2020. 2

[3] O. Bau and I. Poupyrev. REVEL: Tactile feedback technology for augmented reality. *ACM Transactions on Graphics*, 31(4):89:1–89:11, 2012. 2

[4] L.-W. Chan, T.-T. Hu, J.-Y. Lin, Y.-P. Hung, and J. Hsu. On top of tabletop: A virtual touch panel display. In *2008 3rd IEEE International Workshop on Horizontal Interactive Human Computer Systems*, pp. 169–176, Oct. 2008. 2

[5] F. Cosco, C. Garre, F. Bruno, M. Muzzupappa, and M. A. Otaduy. Visuo-haptic mixed reality with unobstructed tool-hand integration. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 19(1):159–172, 2013. 1

[6] A. Costes, F. Danieau, F. Argelaguet, P. Guillotel, and A. Lécuyer. Towards Haptic Images: A Survey on Touchscreen-Based Surface Haptics. *IEEE Transactions on Haptics (ToH)*, 13(3):530–541, 2020. 2

[7] X. De Tinguy, C. Pacchierotti, M. Marchal, and A. Lécuyer. Enhancing the Stiffness Perception of Tangible Objects in Mixed Reality Using Wearable Haptics. In *Proc. IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 81–90. IEEE, 2018. 2

[8] U. Eck, F. Pankratz, C. Sandor, G. Klinker, and H. Laga. Precise haptic device co-location for visuo-haptic augmented reality. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 21(12):1427–1441, 2015. 1

[9] E. Freeman, R. Anderson, J. Williamson, G. Wilson, and S. A. Brewster. Textured surfaces for ultrasound haptic displays. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, ICMI '17, 2 pages, p. 491–492. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3136755.3143020 3

[10] M. Harders, G. Bianchi, B. Knoerlein, and G. Székely. Calibration, registration, and synchronization for high precision augmented reality haptics. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 15(1):138–149, 2008. 2

[11] S. Hattori and T. Hiraki. Accelerated and optimized search of imperceptible color vibration for embedding information into lcd images. In

[12] S. Hattori, Y. Hiroi, and T. Hiraki. Measurement of the imperceptible threshold for color vibration pairs selected by using macadam ellipse, 2024. 8

[13] L. He, C. Xu, D. Xu, and R. Brill. PneuHaptic: Delivering Haptic Cues with a Pneumatic Armband. In *Proc. ACM International Symposium on Wearable Computers*, pp. 47–48, 2015. 2

[14] T. Hiraki, Y. Kawahara, and T. Naemura. Projection-based localization and control method of robot swarms for swarm user interfaces. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, UbiComp '18, 6 pages, p. 584–589. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3267305.3267307 3

[15] M. Inami, N. Kawakami, D. Sekiguchi, Y. Yanagida, T. Maeda, and S. Tachi. Visuo-haptic display using head-mounted projector. In *Proc. IEEE Virtual Reality (VR)*, pp. 233–240. IEEE, 2000. 1, 2

[16] Y. Jiang, K. Zhou, and S. He. Human visual cortex responds to invisible chromatic flicker. *Nat. Neurosci.*, 10(5):657–662, May 2007. 3

[17] Y. Jiang, K. Zhou, and S. He. Human visual cortex responds to invisible chromatic flicker. *Nature Neuroscience*, 10(5):657–662, May 2007. doi: 10.1038/nn1879 8

[18] T. Kikuchi, Y. Yahagi, S. Fukushima, S. Sakaguchi, and T. Naemura. AIR-range: Arranging optical systems to present mid-AIR images with continuous luminance on and above a tabletop. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 722–723, Mar. 2022. 3

[19] S. Kimura, R. Oguchi, H. Tanida, Y. Kakehi, K. Takahashi, and T. Naemura. PVLC projector: Image projection with imperceptible pixel-level metadata. In *Proc. ACM SIGGRAPH Posters*, p. 135:1, 2008. 3

[20] S. Kiuchi and N. Koizumi. Distortion-free mid-air image inside refractive surface and on reflective surface. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 606–614, Mar. 2022. 3

[21] K.-U. Kyung and J.-Y. Lee. Ubi-Pen: A Haptic Interface with Texture and Vibrotactile Display. *IEEE Computer Graphics and Applications*, 29(1):56–64, 2008. 2

[22] J. C. Lee, P. H. Dietz, D. Leigh, W. S. Yerazunis, and S. E. Hudson. Haptic Pen: A Tactile Feedback Stylus for Touch Screens. In *Proc. ACM symposium on User Interface Software and Technology (UIST)*, pp. 291–294, 2004. 2

[23] Y. Lee, S. Lee, and D. Lee. Wearable Haptic Device for Stiffness Rendering of Virtual Objects in Augmented Reality. *Applied Sciences*, 11(15):6932, 2021. 2

[24] S.-w. Leigh, P. Schoessler, F. Heibeck, P. Maes, and H. Ishii. THAW: tangible interaction with see-through augmentation for smartphones on computer screens. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '14*, pp. 89–96, 2015. doi: 10.1145/2677199.2680584 3

[25] Y. Maeda, D. Miyazaki, and S. Maekawa. Aerial imaging display based on a heterogeneous imaging system consisting of roof mirror arrays. In *2014 IEEE 3rd Global Conference on Consumer Electronics (GCCE)*, pp. 211–215, Oct. 2014. 3

[26] S. Maekawa, K. Nitta, and O. Matoba. Transmissive optical imaging device with micromirror array. In *Three-Dimensional TV, Video, and Display V*, vol. 6392, pp. 130–137. SPIE, Oct. 2006. 3

[27] M. Maisto, C. Pacchierotti, F. Chinello, G. Salvietti, A. De Luca, and D. Prattichizzo. Evaluation of Wearable Haptic Systems for the Fingers in Augmented Reality Applications. *IEEE Transactions on Haptics (ToH)*, 10(4):511–522, 2017. 2

[28] Y. Makino, Y. Furuyama, S. Inoue, and H. Shinoda. Haptoclone (haptic-optical clone) for mutual tele-environment by real-time 3d image transfer with midair force feedback. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, 11 pages, p. 1980–1990. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2858036.2858481 3

[29] T. Miyasato, H. Noma, and F. Kishino. Subjective evaluation of perception of delay time between visual information and tactile information. *The Journal of the Institute of Television Engineers of Japan*, 49(10):1353–1356, 1995. 2, 4

[30] Y. Miyatake, T. Hiraki, D. Iwai, and K. Sato. Haptomapping: Visuo-haptic augmented reality by embedding user-imperceptible tactile display control signals in a projected image. *IEEE Transactions on Visualization and*

Computer Graphics, pp. 1–1, 2021. doi: 10.1109/TVCG.2021.3136214 1, 2, 3, 4, 5, 6

[31] Y. Monnai, K. Hasegawa, M. Fujiwara, K. Yoshino, S. Inoue, and H. Shinoda. Haptomime: Mid-air haptic interaction with a floating virtual screen. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, 5 pages, p. 663–667. Association for Computing Machinery, New York, NY, USA, 2014. doi: 10.1145/2642918 .2647407 3

[32] T. Morisaki, M. Fujiwara, Y. Makino, and H. Shinoda. Midair haptic-optic display with multi-tactile texture based on presenting vibration and pressure sensation by ultrasound. In *SIGGRAPH Asia 2021 Emerging Technologies*, SA '21 Emerging Technologies, article no. 10, 2 pages. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3476122.3484849 3

[33] C. Pacchierotti, S. Sinclair, M. Solazzi, A. Frisoli, V. Hayward, and D. Prattichizzo. Wearable Haptic Systems for the Fingertip and the Hand: Taxonomy, Review, and Perspectives. *IEEE Transactions on Haptics (ToH)*, 10(4):580–600, 2017. 2

[34] L. Rahal, J. Cha, and A. El Saddik. Continuous tactile perception for vibrotactile displays. In *Proc. IEEE International Workshop on Robotic and Sensors Environments*, pp. 86–91. IEEE, 2009. 2

[35] J. Rekimoto. SenseableRays: opto-haptic substitution for touch-enhanced interactive spaces. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, pp. 2519–2528. Association for Computing Machinery, New York, NY, USA, Apr. 2009. 2

[36] C. Sandor, T. Kuroki, S. Uchiyama, and H. Yamamoto. Exploring visuo-haptic mixed reality. Technical Report 106, PRMU, Tokyo, Japan, 2007. 1

[37] C. Sandor, S. Uchiyama, and H. Yamamoto. Visuo-haptic systems: Half-mirrors considered harmful. In *Proc. WHC '07*, pp. 292–297, 2007. 1, 2

[38] J. M. Silva, M. Orozco, J. Cha, A. E. Saddik, and E. M. Petriu. Human perception of haptic-to-video and haptic-to-audio skew in multimedia applications. *ACM Trans. Multimedia Comput. Commun. Appl.*, 9(2):1–16, May 2013. 2, 4, 6

[39] M. Strese, J.-Y. Lee, C. Schuwerk, Q. Han, H.-G. Kim, and E. Steinbach. A haptic texture database for tool-mediated texture recognition and classification. In *2014 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE) Proceedings*, pp. 118–123, 2014. doi: 10.1109/HAVE.2014.6954342 5, 7

[40] D. Systems. *"Touch"[Sold as Phantom Omni at by SensAbleTechnologies]*. https://www.3dsystems.com/haptics-devices/touch (Accessed 08/11/2020). 2

[41] K. Tsuchiya and N. Koizumi. Aerial imaging by Retro-Reflection for Mid-Air image display and camera viewpoint transfer. *IEEE Access*, 9:25110–25117, 2021. 3

[42] Y. Ueda, H. Nii, K. Minamizawa, and S. Tachi. Haptomirage: An active-shuttered real image auto-stereoscopic display. In *22nd International Display Workshops, IDW 2015*, pp. 1137–1140. International Display Workshops, 2015. 2

[43] H. Uematsu, D. Ogawa, R. Okazaki, T. Hachisu, and H. Kajimoto. HALUX: projection-based interactive skin for digital sports. In *ACM SIGGRAPH 2016 Emerging Technologies*, number Article 10 in SIGGRAPH '16, pp. 1–2. Association for Computing Machinery, New York, NY, USA, July 2016. 2

[44] D. Wang, Y. Zhang, and Y. Wang. Analysis of registration accuracy for collocated haptic-visual display system. In *Proc. IEEE Haptics Symposium (HAPTICS)*, pp. 303–310, 2008. 2

[45] D. Wang, Y. Zhang, W. Zhou, H. Zhao, and Z. Chen. Collocation accuracy of visuo-haptic system: metrics and calibration. *IEEE Transactions on Haptics (ToH)*, 4(4):321–326, 2011. doi: 10.1109/TOH.2011.17 1

[46] G. Woo, A. Lippman, and R. Raskar. VRCodes: Unobtrusive and active visual codes for interaction by exploiting rolling shutter. In *Proceedings of the 2012 IEEE International Symposium on Mixed and Augmented Reality - ISMAR '12*, pp. 59–64, nov 2012. doi: 10.1109/ISMAR.2012.6402539 3

[47] R. Xiao, S. Hudson, and C. Harrison. CapCam: Enabling Rapid, Ad-Hoc, Position-Tracked Interactions Between Devices. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces - ISS '16*, pp. 169–178, 2016. doi: 10.1145/2992154.2992182 3

[48] G. Yamamoto, L. Sampaio, T. Taketomi, C. Sandor, H. Kato, and T. Kuroda. Imperceptible On-Screen Markers for Mobile Interaction on Public Large Displays. *IEICE Transactions on Information and Systems*, E100.D(9):2027–2036, 2017. doi: 10.1587/transinf.2016PCP0015 3

[49] K. Yoshida, Y. Horiuchi, S. Inoue, Y. Makino, and H. Shinoda. HaptoCloneAR: mutual haptic-optic interactive system with 2D image superimpose. In *ACM SIGGRAPH 2017 Emerging Technologies*, number Article 13 in SIGGRAPH '17, pp. 1–2. Association for Computing Machinery, New York, NY, USA, July 2017. 3

[50] L. Zhang, C. Bo, J. Hou, X.-Y. Li, Y. Wang, K. Liu, and Y. Liu. Kaleido: You Can Watch It But Cannot Record It. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking - MobiCom '15*, pp. 372–385, 2015. doi: 10.1145/2789168.2790106 3