Received May 1, 2019, accepted May 27, 2019, date of publication June 7, 2019, date of current version July 1, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2921580

Imperceptible AR Markers for Near-Screen Mobile Interaction

AKIRA MATSUMOTO[1], SATOSHI ABE[1], TAKEFUMI HIRAKI[2], (Member, IEEE), SHOGO FUKUSHIMA[3], AND TAKESHI NAEMURA[1,3], (Member, IEEE)

[1]Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 1138656, Japan
[2]Graduate School of Engineering Science, Osaka University, Toyonaka 5608531, Japan
[3]Interfaculty Initiative in Information Studies, The University of Tokyo, Tokyo 1138656, Japan

Corresponding author: Akira Matsumoto (akiram@nae-lab.org)

This work was supported by the Japan Society for the Promotion Science (JSPS) KAKENHI under Grant Numbers JP16H01739 and JP17J04216.

ABSTRACT Owing to the pervasive use of displays and smartphones, mobile interactions with display screens have gained attention within the advertising and gaming industries as well as in human–computer interaction research. Communication through QR code-like markers and localization via AR markers are common examples of such interactions. However, these visible markers interfere with the display content; this problem is critical for localization over a wide range of interactions, and fewer markers result in less reliability and accuracy. Although some studies have addressed this issue, few have focused on near-screen interaction without additional hardware. To address this problem, we propose an easy-to-install localization method that uses an array of AR markers, which are made imperceptible to the human eye through chromaticity vibration at 30 Hz. We mainly focus on applications, such as digital signage, where users point their smartphones at the display content. Through four evaluations, we confirm that the pointing error is within 1 mm, and that the proposed system works, when the distance between the screen and smartphone is 4–24 times the size of the AR marker. In addition, we establish that our system is robust against rotation. Finally, we present two potential application scenarios, advertising and navigation.

INDEX TERMS AR markers, displays, human–computer interaction, imperceptible color vibration, mobile computing, ubiquitous computing.

I. INTRODUCTION

Owing to the widespread use of smartphones and the increase in the number of public and private displays, the interaction between display screens and smartphones has attracted attention in the fields of human–computer interaction (HCI) and advertising. In this mobile interaction with screens, the seamless connection of devices is crucial. The currently prevalent systems mainly involve a scenario in which smartphones are only used as gateways for the input and output operations of users, and hardly use the positional relationship between these devices.

Some researchers have focused on device interaction as a space-aware interaction (interaction that varies according to the positional relationship between the screen and smartphone). In this interaction, it is essential to localize the smartphone and measure the positional relationship between the devices. This measurement enables intuitive operation (e.g., a smartphone as a pointer or mouse on the display), enriching the interaction [1]–[3].

For measuring the relationship between the smartphone and screen, the capturing of AR markers using the smartphone camera is a convenient method [4]. However, the display of visible markers impairs the users' visual experience because they occlude the display contents [4], [5]. To solve this problem, various methods have been proposed. By using special hardware, a wide range of interactions can be realized; however, the installation is inconvenient [6], [7]. Feature tracking can be realized by off-the-shelf hardware [8], but the display content is limited to those containing rich features. Yamamoto *et al.* [9] embedded random dot markers [10] on screens utilizing imperceptible color vibration, which displays two different colors alternately. Color vibration can be generated in ordinary 60-Hz displays. However, approximately 20% of the random dot markers must be captured for reliable tracking [11], preventing closer interaction.

In view of the above, we utilize imperceptible color vibration to embed an array of AR markers for smartphone

The associate editor coordinating the review of this manuscript and approving it for publication was Charith Perara.

localization at a closer distance in space-aware mobile interaction with screens, in this study. We define this type of interaction as near-screen interaction. We mainly focus on applications, where a smartphone is pointed at the display content, in near-screen interaction such as digital signage applications. The main motivation for realizing near-screen interaction is because in some cases, users view signage (e.g. maps) from close proximity, and it is convenient for them to receive information in their smartphones because they can continue to access this information after moving away from the screen. In this case, we realize interaction using the positional relationship between the screen and smartphone by measuring the position pointed to by the smartphone in the display coordinate system. We measure the display-pointing accuracy, and clarify the relationship between the marker size and the distance over which marker detection works. We determine, whether our system works, even if the smartphone is tilted or moving. Moreover, we develop a sample application for demonstrating the proposed method, as shown in Figure 1.

In summary, the main contributions of this study are as follows:

- Mapping of the related systems in space-aware mobile interaction with screens, considering the ratio of the images to be captured and the advantages, and the presentation of the challenges to be addressed.
- Realization of smartphone localization by embedding an array of AR markers in the display content using imperceptible color vibration, and development of a method for designing the AR-marker size based on the imaging range and angle-of-view of the smartphone camera.
- Evaluation of the proposed system under laboratory conditions, for clarifying the pointing accuracy and working environment. The results show that our marker-size design method is reasonable, and our system is robust against camera shake and tilt.
- Demonstration of the user experience with a prototype application, using a public display and a smartphone.

II. RELATED WORK

A. SPACE-AWARE MOBILE INTERACTION WITH SCREENS

Some studies have explored the space-aware interaction between screens and smartphones. There are various methods for tracking smartphones, which determine the distance between the smartphone and screen. These methods can be roughly divided into two types: near-screen and far-range. Figure 2 compares the related work on space-aware mobile interaction with screens in terms of 1) the ratio of the image to be captured, and 2) three advantages: the unobtrusive feature (UF), off-the-shelf hardware (OH), and content independency (CI).

1) NEAR-SCREEN INTERACTION

The iPvlc realizes interaction between a screen and a smartphone placed on the screen [7]. This system transmits position information utilizing pixel-level visible light communication (PVLC) [12], and the smartphone receives information through photodetectors. In the PVLC method, independent data are embedded in each pixel by a

(a) Advertising

(b) Navigation

FIGURE 1. Application scenarios: By pointing the smartphone at a restaurant on the map (first frame), the user can (a) download the menu or (b) be guided to the restaurant (second frame).

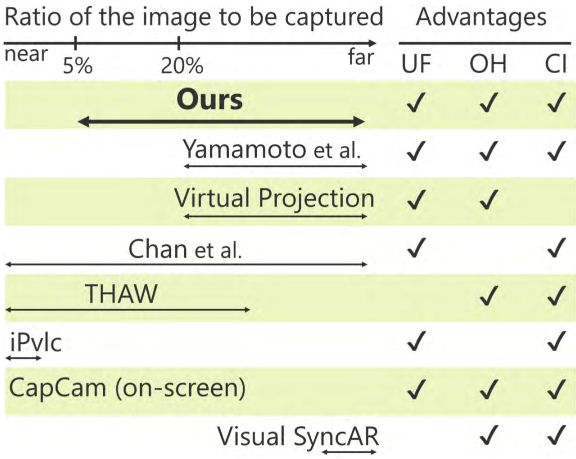


FIGURE 2. **Comparison of the related work on space-aware mobile interaction with screens. The advantages of the respective systems are mapped as per three categories: unobtrusive feature (UF), off-the-shelf hardware (OH), and content independency (CI).**

DLP projector. The PVLC embeds bit patterns as high-speed flicker (≈ 8 kHz), which are imperceptible to the human eye. Although this method realizes precise tracking, it requires special hardware.

The THAW [5] tracks a smartphone that is placed on or hovered over a screen by displaying a two-dimensional (2D) color pattern in the camera's field-of-view alone. This system requires only off-the-shelf hardware. Although its color pattern is meant to be occluded by the smartphone, it is impossible to hide the pattern completely.

The CapCam [13] uses a touchscreen for tracking. In addition, it automatically establishes a wireless link between the smartphone and screen by displaying a sequence of color to the smartphone rear camera. Although this system is completely independent of the displayed content, the smartphone needs to touch the screen.

2) FAR-RANGE INTERACTION

Chan *et al.* [6] combined an infrared and ordinary color projector to project visible content and invisible markers, simultaneously. This system projects AR markers through an infrared projector and detects the markers using an infrared camera connected to a smartphone. It dynamically changes the marker size such that the camera can detect the markers from various distances. Although this approach realizes a wide range of interaction by changing the marker size dynamically, special hardware, such as the infrared projector and infrared camera, are undesirable because they are difficult to install.

Virtual Projection [8] tracks a smartphone by detecting the feature points [14] in the displayed image. While this system does not require obtrusive markers or special hardware, its display content is limited to that containing rich features, i.e., tracking depends upon the display content. Moreover, the smartphone needs to capture 20% of the image for reliable tracking.

Visual SyncAR [15] surrounds the display content with a white frame, and tracks a smartphone by detecting the frame. This system sends a timestamp of the content through a digital watermark. It vibrates the pixel values of the content slightly such that the human eye cannot distinguish the modulated images. Although this system does not depend upon the display content, the smartphone has to capture the entire display for tracking.

Yamamoto *et al.* [9] embedded random dot markers [10] on screens utilizing imperceptible color vibration (color vibration will be explained in detail in the next section). Although this system uses an ordinary 60-Hz display and does not need to capture the entire display, approximately 20% of the random dot markers must be captured for tracking [11].

B. IMPERCEPTIBLE COLOR VIBRATION

Imperceptible color vibration is a promising method for embedding invisible code into images, in the temporal domain. The maximum chromatic flicker frequency perceptible to the human eye is approximately 25 Hz, which can be generated by an ordinary 60-Hz display. Several studies have utilized imperceptible color vibration [9], [16], [17]. The system proposed by Yamamoto *et al.* [9] is unsuitable for our purpose because it uses a webcam and desktop PC for decoding. When digital signage is used, it is preferable, if the receiver is a device usually carried around by people. Abe *et al.* proposed a method for embedding QR codes into images utilizing imperceptible color vibration, and extraction using a smartphone camera [16]. This system satisfies our requirements because the receiver is a smartphone.

To transmit information without impairing the display content, Abe *et al.* modulated the original color of each pixel with two colors that have the same luminance as the original [16]. Imperceptible color vibration can be generated by displaying these two colors alternately, using ordinary 60-Hz displays. The code pattern can be embedded either by vibrating or not vibrating each pixel, thereby representing black or white, respectively. For decoding, a video is captured at 120 fps using a smartphone. The two modulated colors alternately appear in the captured frames, as shown in Figure 3. The vibrating pattern can be extracted by considering the difference between the two frames, and thresholding the output.

III. METHOD

We propose a smartphone localization system by embedding an array of AR markers into images using imperceptible color vibration, and a method for designing the marker size based on the imaging range and angle-of-view of the smartphone camera. As shown in Figure 4, we embedded an array of AR markers into the display content using imperceptible color vibration, and detected them using a smartphone. We can calculate the position of the smartphone from the detected markers. The size of the AR markers and their arrangement determines the maximum and minimum distances of the smartphone from the screen. Let M be the length of the marker-side and G be the interval between markers.

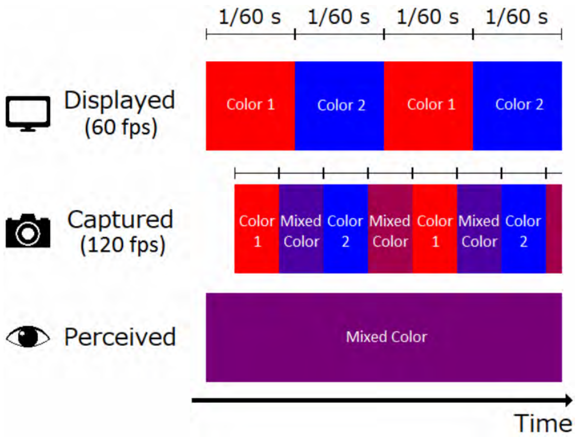


FIGURE 3. Time sequences of the frames for the system components and human eye [16].

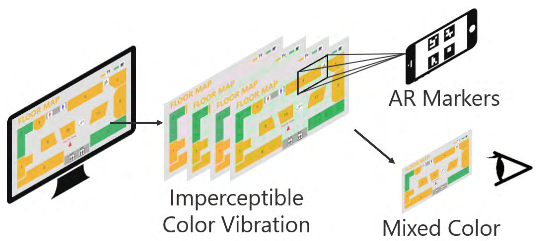


FIGURE 4. System operation with the proposed method. Two color-vibrated images are displayed alternately. The smartphone camera extracts the embedded AR markers, while the human eye perceives the normal image alone.

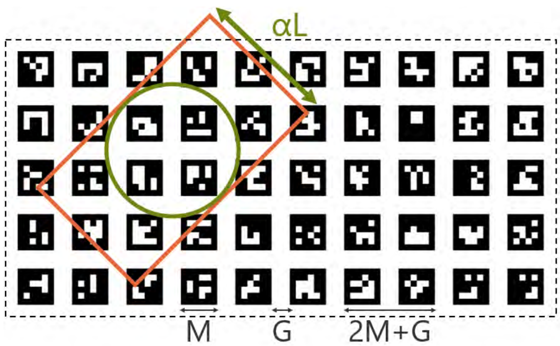


FIGURE 5. Marker-alignment example. The orange rectangle indicates the region captured by the smartphone. The green circle indicates the region always captured by the smartphone, regardless of its rotation.

We assume that the center of the smartphone camera points to a location within the dashed rectangle, as shown in Figure 5. The distance between the rectangular border and the outer-most markers is $G/2$. Let θ be the diagonal angle-of-view of the smartphone camera, $m : n$ $(m \geq n)$ be the aspect ratio of the video, and L be the distance between the smartphone and screen. Assuming that the smartphone is parallel to the screen, the area of the screen captured by the smartphone will be a rectangle with a diagonal length of $2L \tan \frac{\theta}{2}$ (the orange rectangle in Figure 5). The length of the shorter-side of the rectangle is αL, where $\alpha = \frac{2n \tan \frac{\theta}{2}}{\sqrt{m^2+n^2}}$.

To determine the minimum L needed to capture an entire marker, regardless of how the smartphone is translated and rotated, it is sufficient, if a marker is always present within the circle, centered on the center of the rectangle and inscribed within the rectangle (green circle in Figure 5). This condition is satisfied, if the diameter of the circle is greater than the diagonal length of the square surrounding 2 × 2 markers. Therefore,

$$\alpha L \geq \sqrt{2}(2M + G) \tag{1}$$

must be satisfied to capture an entire marker.

Next, we consider the maximum L at which camera localization is possible. Let R be the number of pixels of the captured image corresponding to αL, P be the number of pixels of the captured marker corresponding to M, and P_{min} be the minimum P required for detecting the marker. The value of P can be calculated from the product of R and the ratio of M to the length of the shorter-side of the rectangle. Hence,

$$P = R\frac{M}{\alpha L} \geq P_{min}. \tag{2}$$

In summary, the range of L is

$$M\frac{R}{\alpha P_{min}} \geq L \geq \frac{\sqrt{2}(2M + G)}{\alpha}. \tag{3}$$

The value of P_{min} is calculated experimentally and discussed in the next section.

IV. EXPERIMENTS

We conducted four experiments for measuring the display pointing accuracy and the maximum length between the smartphone and screen, and for evaluating the robustness against motion and rotation, respectively. We used a 13.3-inch 1080p laptop (MB-S250XN1-EX3, Mouse Computer) as the screen and the Galaxy S8 (Samsung) as the smartphone ($\theta \approx 70^\circ, m = 16, n = 9, R = 720$ px). The experiments were conducted in an assembled darkroom (ADR-F2, ASONE).

A. EXPERIMENT-1: POINTING ACCURACY

To avoid error between the ground truth and the actual position of the smartphone, we estimated the relative position from a certain position, where the smartphone was pointed at.

We embedded 10 × 5 AR markers ($M = 120$ px, $G = 60$ px) into a gray single-color image ((R, G, B) = (128, 128, 128)). We used the ArUco [18], [19] to generate 4 × 4 bit markers. The modulated colors were (R, G, B) = (158, 117, 131) and (R, G, B) = (85, 138, 125). Figure 6 displays the actual images. We used an XY plotter to move the smartphone in increments of 50 mm, parallel to the edge of the display. The smartphone was positioned 140-mm above the display. We measured at 5 × 3 points, starting from a certain point on the upper-left of the display. We performed 100 measurements at each point, and calculated the relative position by subtracting the measured value from the starting point. In addition, to determine the accuracy of pointing using the

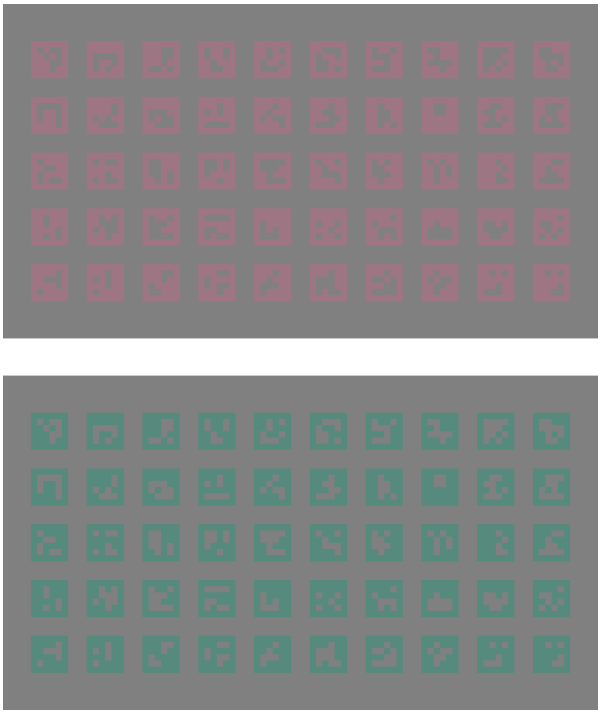


FIGURE 6. Images used in Experiment-1; these two images are displayed alternately at 60 Hz.

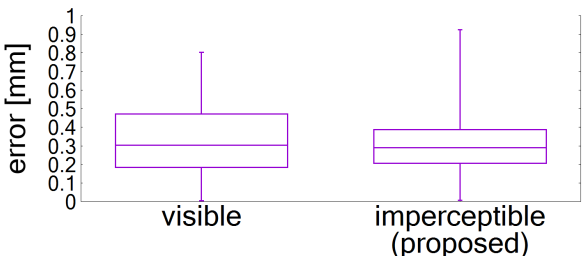


FIGURE 7. Error results; there were no significant differences between the visible and imperceptible markers.

AR markers themselves, we conducted the same experiment with visible markers, i.e., black ((R, G, B) = (0, 0, 0)) markers on a white ((R, G, B) = (255, 255, 255)) background using the images shown in Figure 6.

We defined the error as the distance between the ground truth and the measured value. The results are depicted in Figure 7. The median of the error with visible markers was 0.30 mm and the maximum error was 0.80 mm. In contrast, the median of the error with imperceptible markers was 0.29 mm and the maximum error was 0.92 mm. There were no significant differences between the visible and imperceptible markers. It was confirmed that our system has sufficient accuracy for practical applications because the error was within 1 mm.

B. EXPERIMENT-2: MAXIMUM DISTANCE BETWEEN THE SMARTPHONE AND SCREEN

We measured the detection rate at various values of L (distance between the smartphone and screen) and

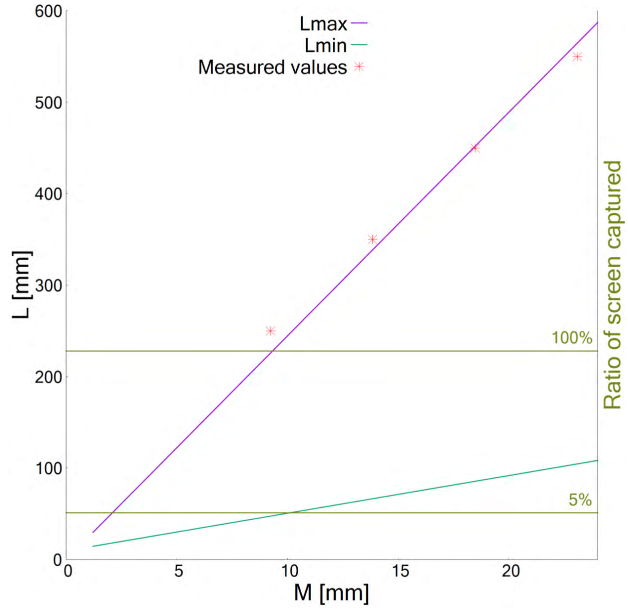


FIGURE 8. Range of L according to M. The points indicate the largest L at which the measured detection rate is higher than 50% for each M, L_{max} is their least squares approximation, and L_{min} is a line derived from the right-side of inequality (3). Our system can be used in the range between L_{max} and L_{min}. The green horizontal line indicates the ratio of the screen captured by the smartphone from a distance, L.

M (marker length), in this experiment. We positioned the smartphone to capture the center-part of the display, and moved it perpendicular to the display in increments of 50 mm. We performed 100 measurements at each point and calculated the detection rate. We used the same gray image and modulated colors as in Experiment-1. We used multiple sizes and varied the numbers of markers (M = 9.24, 13.85, 18.47, 23.09 mm).

We can estimate the maximum L with respect to M from the result. We plotted the largest L for which the detection rate was higher than 50% for each M (Figure 8); line, L_{max}, is their least-squares approximation ($L_{max} \approx 24M$). The value of P_{min} can be calculated from the left-side of inequality (3) and the slope of L_{max}, i.e., $P_{min} \approx 43$. Line, L_{min}, is derived from the right-side of inequality (3), when G = 30 px ($L_{min} \approx 4M + 9.5$). Hence, the values of L between L_{max} and L_{min} are the distances over which our system works. For example, the smartphone captured approximately 5% of the screen at $L = 50$ mm and 100% at $L = 230$ mm, under the conditions of the conducted experiment.

C. EXPERIMENT-3: DETECTION RATE WITH A MOVING SMARTPHONE

We measured the detection rate, when the smartphone was moving. The smartphone was set on an electrical linear slide (EZS6-D085-AZMAD-1, Oriental Motor) and moved at constant speed, 140 mm above the display . The distance moved from the starting point was measured. We performed 100 measurements at each speed and calculated the detection rate. The same gray image, modulated colors, and embedded markers, as in Experiment-1, were used.

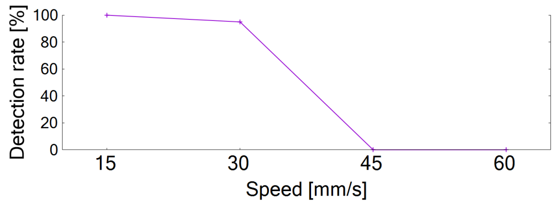


FIGURE 9. Detection rates for different smartphone movement speeds.

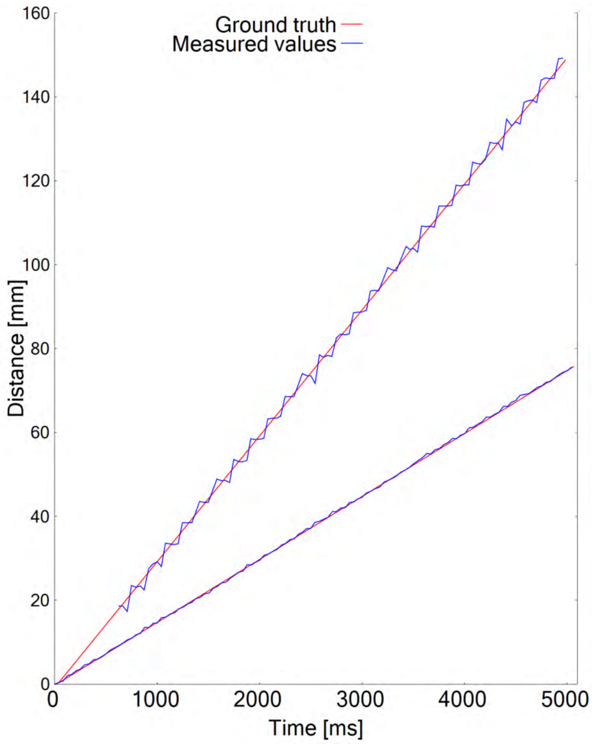


FIGURE 10. Measured values for smartphone speeds of 15 and 30 mm/s. The ground truth includes raw data from the linear slide. The markers could not be detected immediately, after the movement commenced at 30 mm/s. This result implies that our system can withstand camera shake.

The result is depicted in Figure 9. The markers could not be detected at more than 45 mm/s. Figure 10 shows the measured values and ground truth at 15 mm/s and 30 mm/s. The ground truth includes the raw data from the linear slide. The measured values were smooth at 15 mm/s and jagged at 30 mm/s. Although our system performance was unsatisfactory when the smartphone was moving, this result implies that our system can withstand camera shake.

D. EXPERIMENT-4: DETECTION RATE WITH A TILTED SMARTPHONE

We measured the detection rate, when the smartphone was tilted. The smartphone was positioned 90-mm above and parallel to the display, and rotated in increments of 15° along the three axes defined in Figure 11, respectively. We measured the distance from the pointing position at 0° to the pointing position at each angle. Errors within 8 mm were defined as correct detection. We performed 100 measurements at each angle and calculated the detection rate. We used the same

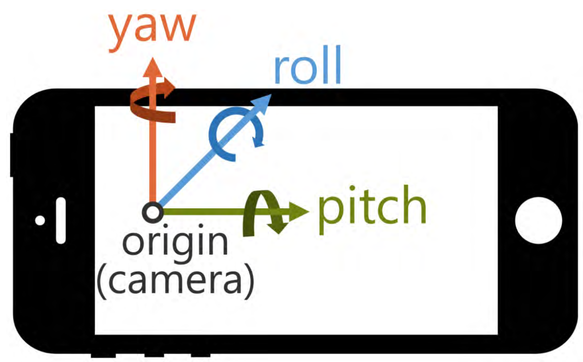


FIGURE 11. Definition of the three rotation axes.

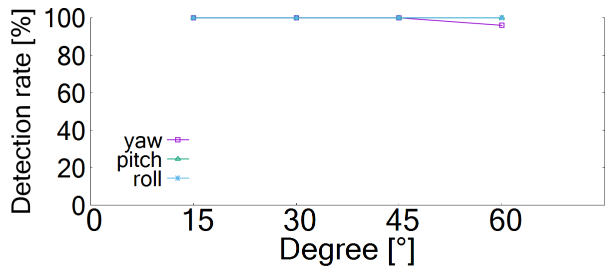


FIGURE 12. Detection rates for the three rotation axes. It is confirmed that our system is robust against rotation.

gray image, modulated colors, and embedded markers as in Experiment-1. The result is depicted in Figure 12, confirming that our system is robust against rotation.

V. APPLICATION

We developed an application in which the display screen depicts a floor map and a user acquires information on a region of the map by pointing a smartphone at it (Figure 1). The user can either jump to a related webpage or obtain directions to the spot indicated on the map. We confirmed that the AR markers worked for various positions of the hand-held smartphone. This application demonstrates the potential of this method for application in fields such as advertising and indoor navigation.

VI. CONCLUSION

In this study, we proposed an easy-to-install method for localizing a smartphone near a screen, with embedded AR markers through imperceptible color vibration. We demonstrated that our system has sufficient accuracy and robustness against rotation for practical applications, and clarified the relationship between the marker-size and distance-range. Specifically, our system works, even when the smartphone captures only 5% of the screen. In addition, we developed an application to demonstrate our method, and established the potential of the proposed method for application in fields such as advertising and indoor navigation.

A limitation of our method is that our system does not work, when the smartphone is placed on the screen. In future, we aim to develop a method for tracking smartphones that are placed on the display screen. Our system will be more practical, if we can create an embedded marker that can handle a wider distance range.

REFERENCES

[1] M. Baldauf, M. Salo, S. Suette, and P. Fröhlich, "Display pointing: A qualitative study on a recent screen pairing technique for smartphones," in *Proc. 27th Int. BCS Hum. Comput. Interact. Conf. (BCS-HCI)*, 2013, pp. 45:1–45:6.

[2] S. Siddhpuria, S. Malacria, M. Nancel, and E. Lank, "Pointing at a distance with everyday smart devices," in *Proc. CHI Conf. Hum. Factors Comput. Syst. (CHI)*, 2018, pp. 173:1–173:11.

[3] J. Y. Lee, M. S. Kim, D. W. Seo, S. M. Lee, and J. S. Kim, "Smart and space-aware interactions using smartphones in a shared space," in *Proc. 14th Int. Conf. Hum.-Comput. Interact. Mobile Devices Services Companion (MobileHCI)*, 2012, pp. 53–58.

[4] N. Pears, D. G. Jackson, and P. Olivier, "Smart phone interaction with registered displays," *IEEE Pervasive Comput.*, vol. 8, no. 2, pp. 14–21, Apr. 2009.

[5] S. Leigh, P. Schoessler, F. Heibeck, P. Maes, and H. Ishii, "THAW: Tangible interaction with see-through augmentation for smartphones on computer screens," in *Proc. 9th Int. Conf. Tangible, Embedded, Embodied Interact. (TEI)*, 2015, pp. 89–96.

[6] L.-W. Chan, H.-T. Wu, H.-S. Kao, J.-C. Ko, H.-R. Lin, M. Y. Chen, J. Hsu, and Y.-P. Hung, "Enabling beyond-surface interactions for interactive surface with an invisible projection," in *Proc. 23nd Annu. ACM Symp. User Interface Softw. Technol. (UIST)*, 2010, pp. 263–272.

[7] Y. Kato, N. Fukasawa, and T. Naemura, "iPvlc: Pixel-level visible light communication for smart mobile devices," in *Proc. ACM SIGGRAPH Posters*, 2011, p. 45.

[8] D. Baur, S. Boring, and S. Feiner, "Virtual projection: Exploring optical projection as a metaphor for multi-device interaction," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. (CHI)*, 2012, pp. 1693–1702.

[9] 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 Trans. Inf. Syst.*, vol. E100.D, no. 9, pp. 2027–2036, 2017.

[10] H. Uchiyama and H. Saito, "Random dot markers," in *Proc. IEEE Virtual Reality Conf. (VR)*, Mar. 2011, pp. 35–38.

[11] L. Yang, J.-M. Normand, and G. Moreau, "Robust random dot markers: Towards augmented unprepared maps with pure geographic features," in *Proc. 20th ACM Symp. Virtual Reality Softw. Technol. (VRST)*, 2014, pp. 45–54.

[12] S. Kimura, M. Kitamura, and T. Naemura, "EmiTable: A tabletop surface pervaded with imperceptible metadata," in *Proc. 2nd Annu. IEEE Int. Workshop Horizontal Interact. Hum.-Comput. Syst. (Tabletop)*, Oct. 2007, pp. 189–192.

[13] R. Xiao, S. Hudson, and C. Harrison, "CapCam: Enabling rapid, ad-hoc, position-tracked interactions between devices," in *Proc. ACM Int. Conf. Interact. Surfaces Spaces (ISS)*, 2016, pp. 169–178.

[14] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, 2008.

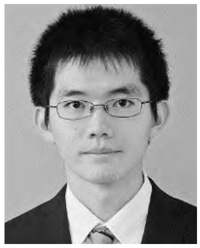
[15] S. Yamamoto, H. Tanaka, S. Ando, A. Katayama, and K. Tsutsuguchi, "Visual SyncAR: Augmented reality which synchronizes video and overlaid information," *J. Inst. Image Electron. Eng. Jpn.*, vol. 43, no. 3, pp. 397–403, 2014.

[16] S. Abe, T. Hiraki, S. Fukushima, and T. Naemura, "Screen-camera communication via matrix barcode utilizing imperceptible color vibration," in *Proc. 31st Annu. ACM Symp. User Interface Softw. Technol. Adjunct (UIST)*, 2018, pp. 166–168.

[17] G. Woo, A. Lippman, and R. Raskar, "VRCodes: Unobtrusive and active visual codes for interaction by exploiting rolling shutter," in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR)*, Nov. 2012, pp. 59–64.

[18] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognit.*, vol. 47, no. 6, pp. 2280–2292, 2014.

[19] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and R. Medina-Carnicer, "Generation of fiducial marker dictionaries using mixed integer linear programming," *Pattern Recognit.*, vol. 51, pp. 481–491, Mar. 2016.



AKIRA MATSUMOTO received a B.S. degree from the University of Tokyo, Japan, in 2018, where he is currently pursuing a master's degree in the Graduate School of Information Science and Technology. His research interests include the areas of augmented reality, human–computer interaction, and ubiquitous computing.

SATOSHI ABE received the B.S. and M.S. degrees from the University of Tokyo, Japan, in 2017 and 2019, respectively. His research interests include the areas of augmented reality, human–computer interaction, and ubiquitous computing.

TAKEFUMI HIRAKI (S'16–M'19) received the Ph.D. degree in information and communication engineering from the University of Tokyo, Japan, in 2019. He is currently a Japan Society for Promotion of Science (JSPS) Research Fellow with the Graduate School of Engineering Science, Osaka University. His research interests include the areas of augmented reality, human–computer interaction, and ubiquitous computing. He is a member of the ACM.

SHOGO FUKUSHIMA received the Ph.D. degree in engineering from the University of Electro-Communications, in 2013. He was a Visiting Student with the Camera Culture Group, MIT Media Lab supported by the "Japan Society for Promotion of Science (JSPS) Research Fellowships for Research Abroad," and a Project Researcher with the Graduate School of Information Science and Technology, the University of Tokyo. He is currently an Assistant Professor with the Interfaculty Initiative in Information Studies, The University of Tokyo, and a Researcher at Japan Science and Technology Agency (JST) PRESTO. His research interests include intelligence amplification (IA), virtual reality, entertainment computing, and human emotions. He is a member of the ACM.

TAKESHI NAEMURA (M'01) received the Ph.D. degree in electrical engineering from the University of Tokyo, Japan, in 1997. He is currently a Professor with the Interfaculty Initiative in Information Studies, the University of Tokyo. He was a Visiting Assistant Professor of computer science with Stanford University, USA, supported by the "Japan Society for Promotion of Science (JSPS) Postdoctoral Fellowships for Research Abroad" from 2000 to 2002. His research interests include virtual reality and human interface. He is a member of the ITE, IEICE, and ACM.