CARDS: Comprehensive AR Docent System

Seung Un Lee*† Jiyoung Yun*† Dain Kim*† Dooyoung Kim ‡ Seo Young Oh § Sang Ho Yoon¶

KAIST

2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) | 978-1-6654-5365-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISMAR-ADJUNCT57072.2022.00156

ABSTRACT

We present CARDS, a comprehensive Augmented Reality (AR) docent system to promote informative and engaging art experiences for exhibition visitors. While utilizing AR to add aesthetic visual effects has been actively explored, facilitating visitor engagement in educational exhibitions through AR has been relatively unexplored. Thus, the proposed system facilitates visitor engagement in informative exhibitions by providing AR-specific interaction features which include context-based sequential AR pins, position-based AR pin configuration, and orientation-based AR visual aids. With CARDS, we aim to enrich the onsite art experience of visitors by presenting information on exhibits in a systematic way, as well as to enhance visitor engagement by taking the bodily movements of visitors as interaction inputs. Furthermore, we suggest a design guideline for future educational AR guide applications using proxemic interaction methods.

Index Terms: Human-centered computing—Interaction design—Interaction design process and methods—User centered design

1 INTRODUCTION

Over the last decades, exhibitions have attempted to apply Human-robot technology, Augmented Reality (AR), and Virtual Reality (VR) to enhance the art experience of onsite visitors [7, 8, 15, 17, 21]. Among these technologies, Augmented Reality (AR) is gaining attention as it offers a virtual experience grounded in the real world. Especially, its application in exhibition guide systems is notable in terms of a more effective and interactive information delivery [9–12, 19, 20]. The information delivery functions offered by the AR-based guide systems help overcome the limitations of traditional appreciation measures. In physical exhibitions, visitors typically rely on traditional appreciation measures such as audio narrations, descriptions written on walls, and brochures [13]. These measures make it difficult for visitors to match where the information details are referring to. Furthermore, displaying visual aids such as reference works is rather limited, requiring visitors to search for them on their mobile phones.

With AR-based systems, localized AR cues and visual aid augmentation functions solve the discontinuous art experience brought about by the traditional approaches. Localized AR cues visually map the descriptions to where they are referring to the exhibits. Visitors can touch an AR cue to access information on that specific point on an exhibit [17, 19]. Furthermore, visual aids such as reference works are augmented adjacent to the original exhibit [18]. Visitors can actually view the visual aids near the original exhibit without the need to look them up on their mobile phones. Thus, an AR-based guide system provides a continuous and immersive art experience in terms of information delivery.

*Co-First Authors: S. Lee, J. Yun, D. Kim
†e-mail: [seungun.lee, clatya96, dainkim]@kaist.ac.kr
‡dooyoung.kim@kaist.ac.kr
§seoyoung.oh@kaist.ac.kr
¶Corresponding Author: sangho@kaist.ac.kr

The two main features of the existing AR guide systems, localized AR cues and visual aid display, still have rooms for improvement. At the moment, localized AR cues and AR visual aids remain stationary regardless of visitors' position and orientation. However, in actual exhibitions, visitors appreciate exhibits in various ways - as a whole, in part, from different angles - by moving back and forth and moving sideways. Thus, if localized AR cues and AR visual aids could be dynamically adjusted based on visitors' movements, the semantic relations between visitors' movements (interaction input) and the augmented information contents (interaction output) would be stronger, resulting in a more engaging information delivery process. In this regard, we considered the concept of 'proxemic interaction,' which focuses on the spatial relationships in designing the user-device interaction process. Proxemic interaction adopts continuous user movements such as position and orientation as inputs to interact with surrounding digital devices [2].

In this respect, we present Comprehensive AR Docent System (CARDS), an AR guide system that integrates refined versions of localized AR cues and visual aid display based on a coherent proxemic interaction method to strengthen the semantic relations between visitors' movements and the augmented information contents in the interaction process. The key features of CARDS are the following:

1. Context-based Sequential AR Pins
2. Position-based AR Pin Configuration
3. Orientation-based AR Visual Aids

Context-based sequential AR pins refer to AR cues connected by arrows, which allow visitors to access information on exhibits in a certain order. The pins are configured into three layers according to the scope of information, which are accessible by changing the visitor-exhibit distance. During the process, visitors can view relevant visual aids augmented over the exhibit by moving sideways. Based on the key features, the main contributions of CARDS are the following:

1. Provides an informative onsite art experience for visitors through systematic presentation of exhibit information
2. Presents a novel AR interaction method that facilitates visitor engagement by taking continuous body movements of visitors - position and orientation change - as interaction inputs
3. Suggests a design guideline for future educational AR guide applications willing to incorporate proxemic interaction into their designs

2 RELATED WORKS

Localized AR Cues Previous studies explored the usage of AR technology in terms of displaying localized descriptions on specific locations of exhibits. Sugiura et al. [19] presented a tablet AR system that augmented virtual labels referring to certain parts of medical specimens. An AR system presented by Chiu et al. [5] provided a more interactive AR annotation system by utilizing augmented red cues to present localized information once the cues are touched. The study revealed that those who experienced the interactive AR guide

2771-1110/22/$31.00 ©2022 IEEE
DOI 10.1109/ISMAR-Adjunct57072.2022.00156

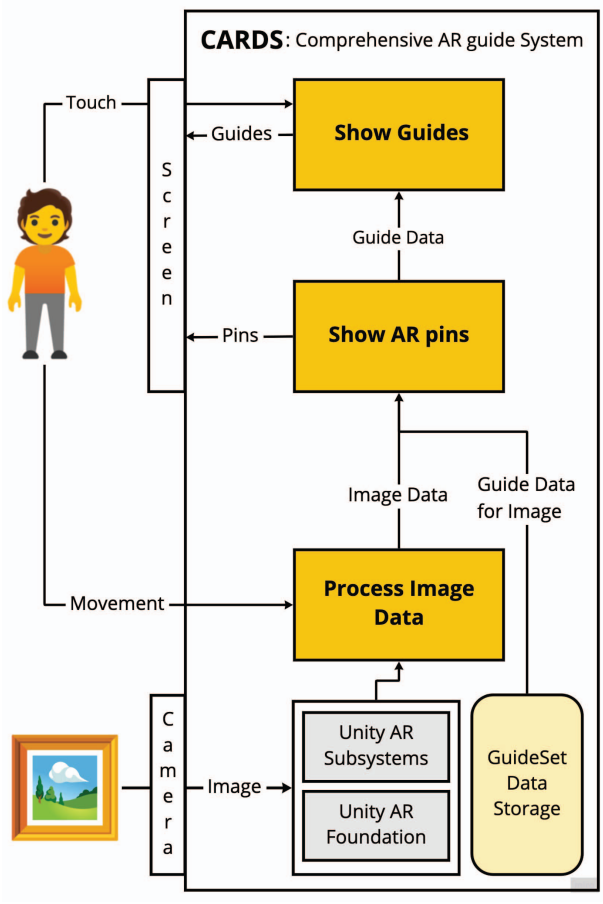

Figure 1: System Flow Diagram - CARDS recognizes exhibit images based on marker detection. Once an image is recognized, the system displays the corresponding guide data. Different AR pins and guides are displayed based on touch input and visitor movements

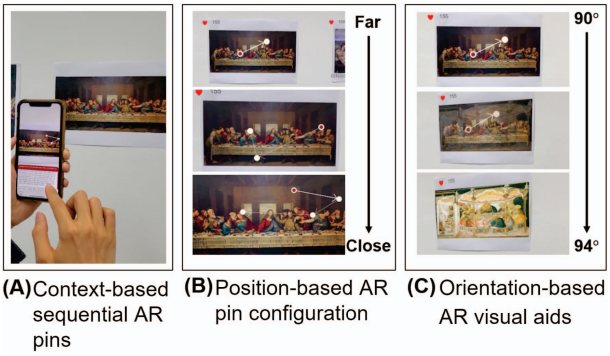

Figure 2: Implementation of (A) Context-based sequential AR pins, (B) Position-based AR pin configuration, and (C) Orientation-based AR visual aids

showed higher learning outcomes than those who did not. Similarly, a study by Chang et al. [3] showed consistent results that localized AR cues led to higher learning effectiveness and a positive user experience. However, the localized AR cues are not presented in a specific order, which makes it difficult for visitors to figure out the connections among the cues. Thus, a structuralized presentation of localized information based on specific guiding orders would improve the AR guide system.

Position-based Interaction The importance of considering the spatial relations between a user and a device in the interaction design process has been stressed by the previous literature [1, 2, 14]. In the case of utilizing mobile AR, the same logic can be applied between a user - with a mobile device - and a physical object [16]. The importance of considering the spatial relations is especially relevant in the exhibition setting, as the spatial positions of visitors is crucial due to their direct relationship to the art experience [4, 6]. A recent study by Chen et al. applied the proxemic interaction method to their AR guide system, augmenting different scopes of exhibit information according to the visitor-exhibit distance [4]. The study opens the possibility for future studies to apply proxemics to information delivery in AR.

Visual Aid Display AR-based digital guide systems enrich the art experience in exhibitions by augmenting visual aids. An AR app offered by PACE Gallery allowed visitors to view virtual reference artworks adjacent to the original exhibit [18]. The AR guide systems introduced by Sugiura et al. [19] and Chiu et al. [5] also assisted visitors with additional reference images and videos augmented on the side of exhibits. Despite the convenience of displaying visual aids with a simple touch, such interaction does not reflect visitors' art-appreciating behaviors. As in the case of position-based interaction, the proxemic approach seems to be applicable to the interaction method for displaying visual aids.

3 System Design and Implementation

Based on the previous studies, we discovered that the AR guide systems help enhance the art experience of exhibition visitors by augmenting localized exhibit descriptions and visual aids, and also by applying the proxemic interaction method to the information delivery process. However, each feature has rooms for improvement in terms of user experience design. Therefore, this study presents CARDS, a comprehensive AR guide system that integrates all three features - localized AR annotation, position-based interaction, and visual aid display - with each of them refined.

3.1 System Design

Figure 1 shows the system overview of CARDS. CARDS prototype is implemented with Unity 3D engine (2020.3.30f1). Visitors are equipped with android or iOS mobile devices. Interactions with virtual pins are developed in Unity with the AR Foundation. CARDS image recognition is marker-based, with image targets (exhibits) loaded to the reference image library beforehand. The descriptions and visual aids that correspond to the image targets (exhibits) are stored in the unity app local storage in advance.

3.2 CARDS Implementation

3.2.1 Image Recognition and Pin Interaction

Figure 3 shows the image recognition and pin interaction process. When the visitor positions the mobile device towards the target image (2D exhibit), the image is recognized, triggering a heart icon and a touch text guide (Figure 3.A). The number on the heart icon reflects the sum of 'like' reactions to the AR pins. We refer to AR visual cues as AR 'pins', due to their visual similarity.

Once the user touches the exhibit on the screen, AR pins appear on top of the image (Figure 3.B). The pins are connected with white lines, indicating that the pins are accessible in a specific order. The connected pins, referred to as 'sequential pins', allow the visitor to

appreciate exhibits in a certain order, which is carefully designed by art professionals. The first pin in order is marked with red, and the rest is marked with yellow. Different AR pins are shown in different layers, which are accessible through changing the visitor-exhibit distance. Details on the layer change function will be explained in the next section.

When the visitor touches the first pin, the localized description is displayed along with exhibit caption, pin number, reaction button, and navigation buttons (Figure 3.C). Once triggered, the aforementioned UI elements of the pin are fixed to the mobile screen, which does not require the visitor to keep the mobile device pointed towards the exhibit throughout the art experience. The visitor can put the mobile device down while reading the descriptions, and hold it back up afterwards for further interaction. Visitors can move to the next pin by touching the navigation button on the bottom of the screen. The pin that the visitor is currently accessing is marked with red. During the process of navigating through the pins, the visitor can press the reaction button, which is reflected in the heart icon number on the top. The UI elements can be turned off by touching anywhere on the exhibit.

3.2.2 Position-based AR Pin Configuration

A notable art appreciating behavior is moving back and forth to view an exhibit as a whole and in part. The insight derived from this behavior is that visitors acquire different visual information from the exhibit depending on their positions. In this respect, CARDS grouped the AR pins into three layers according to visitor-exhibit distance: close, middle, and far (Figure 4). The layers map the scope of the exhibit information to distance; the farthest layer contains general information and the closest layer contains specific details on the exhibit. The sequential AR pins form a single storyline per layer.

The layers are accessible by changing the visitor-exhibit distance. The initial layer is set according to the position measured at the moment of exhibit image recognition. Raycast-based distance calculation happens in real-time to immediately adjust the corresponding layers. The layer change occurs when the distance goes over or below a pre-set threshold values. The threshold values for CARDS were set to 0.2m and 0.4m, each marking the close-middle layer change point and the middle-far layer change point. The threshold values as well as the number of layers can be adjusted manually based on the size of the target exhibit and the exhibition environment.

3.2.3 Orientation-based AR Visual Aids

Visitors are able to access additional visual aids by moving sideways with their mobile devices pointed towards the exhibit as if they are seeing 'other aspects' of it. The rotation angle between the visitor and the exhibit is measured in real-time, which is reflected in the alpha value of the visual aid. Figure 5 illustrates the calculation process of the rotation angle. Based on the calculation, the visitor views the original exhibit within the range of 88° to 92°. When the visitor goes beyond the range, the visual aid is viewable (Figure 6.A and B). The change of angle is further notified to the visitors in the form of a curved bidirectional arrow combined with a circle, as shown in Figure 6.C. The circle, which represents the visitor's current rotation angle, smoothly moves along the curved arrow in real-time as the visitor moves sideways.

4 User Scenario

We envisioned a typical user scenario of CARDS to be as follows. First, the visitor points to the exhibits with a mobile device. A heart icon with a number is augmented along with a touch text guide for each exhibit. The visitor scans the exhibits to get a rough idea of the preference of the general public. After the scanning, the visitor selects an exhibit of interest and touches it. In this step, the initial visitor-exhibit distance is measured, and the corresponding layer

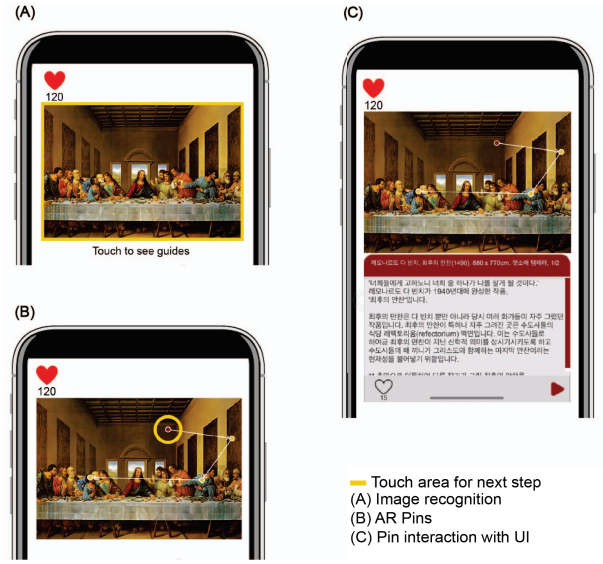

Figure 3: The process of Image Recognition and Pin Interaction - (A) Image recognition (B) Context-based sequential AR Pins (C) Pin interaction with UI

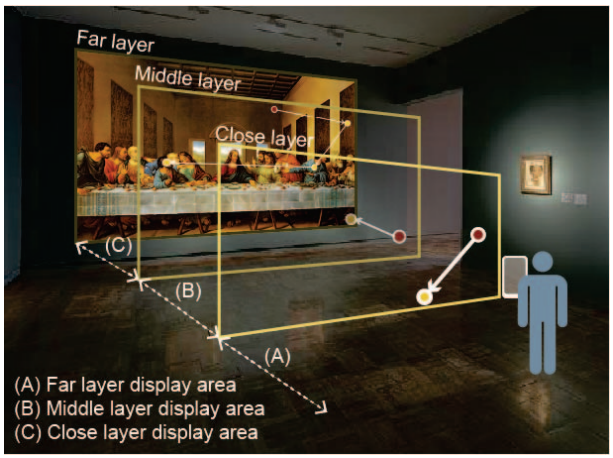

Figure 4: Position-based AR Pin Configuration - Context-based sequential AR pins are configured into three layers based on visitor-exhibit distance. When the visitor's position is within range (A), AR pins of close layer is displayed. Same applies to range (B) and (C), each corresponding to middle and far layer.

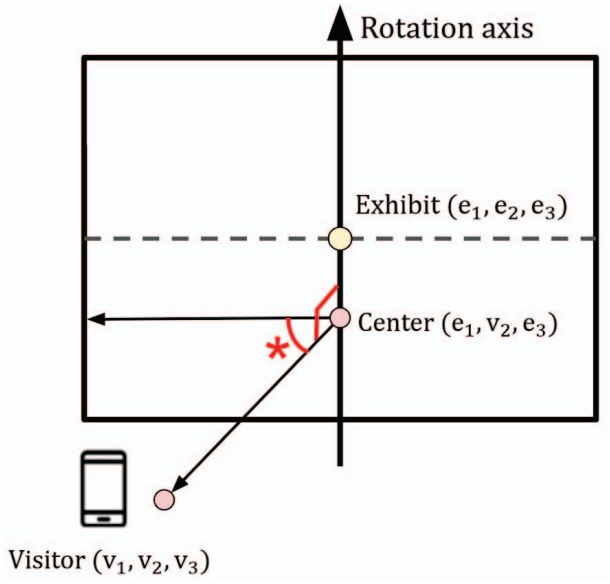

Figure 5: Rotation Angle Calculation Model - The center of rotation is the position on the rotation axis that goes by (e_1, v_2, e_3). The alphabet 'e' refers to the exhibit, and the alphabet 'v' refers to the visitor. Thus, the x and z position value of the center is identical to the x and z position value of the exhibit. Likewise, the y position value of the center is identical to the y position value of the visitor. The rotation axis is the vertical line (y-axis) that crosses the exhibit position that goes by exhibit(e_1, e_2, e_3). The visitor-center vector, the vector that connects the center and the visitor(v_1, v_2, v_3), is perpendicular to the rotation axis. The fixed vector points to the left edge of the exhibit from the center. The rotation angle (marked with $*$), which determines the orientation of the visitor, refers to the angle formed between the visitor-center vector and the fixed vector.

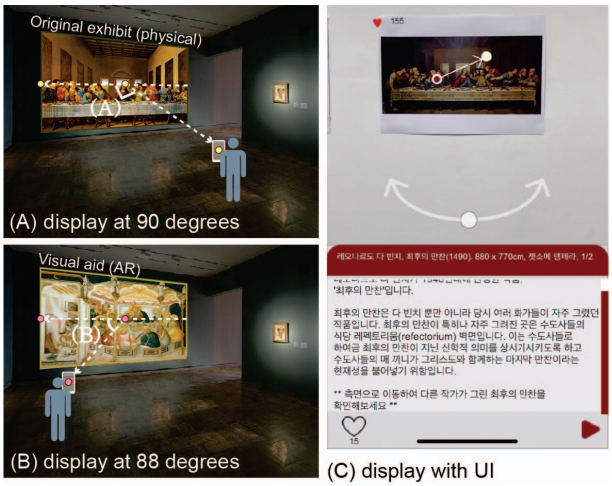

Figure 6: Orientation-based AR Visual Aids - When the rotation angle between the visitor and the exhibit of interest is near 90 degrees, the visitor views the original exhibit (A). When the visitor moves to the left, and the rotation angle goes below 88 degrees, visual aid is augmented on top of the original exhibit (B).

pops up. The layer contains AR pins that the visitor can interact with.

Second, the visitor selects the AR pins to read the corresponding descriptions. The context-based sequential AR pins allow the visitor to appreciate exhibits in a certain order, which is carefully designed by art professionals. The visitor navigates back and forth to access the connected descriptions. If the visitor wants to view the related image mentioned in the descriptions, the visitor moves sideways with the mobile device pointed towards the exhibit. The visitor checks the orientation angle based on the arrow and circle UI. After returning back to the description panel, the visitor presses the reaction button if the description pin is satisfying. Once the button is pressed, the button turns into a red heart, and the total number of buttons pressed is updated.

Third, the visitor moves back and forth to gain access to other layers. If the visitor wants to know the general information, such as historical backgrounds, the visitor steps away from the exhibit. If the visitor wants to learn about the specific details, such as painting techniques, the visitor walks toward the exhibit.

5 Conclusion

In this poster, we presented CARDS, a comprehensive AR guide system that assists visitors during their art experience through context-based sequential AR pins, position-based AR pin configuration, and orientation-based AR visual aids. The integration of the three functions allows visitors to acquire structuralized information through AR-specific interactions in a systematic way. In this respect, CARDS is expected to bring educational benefits as well as an enhanced sense of engagement through body movements. In this respect, CARDS suggests a novel AR interaction method that promotes an immersive and continuous art experience in informative exhibitions. Hence, we expect CARDS to provide design guidelines for future educational AR guide applications in terms of incorporating proxemic interaction inputs, such as user position and orientation, into their designs.

The limitation of this study is the lack of evaluation, as the study is yet in its early stages of research. We plan to conduct a user study with a minimum of 20 participants to evaluate CARDS in terms of usability, user experience, and learning effectiveness. Furthermore, CARDS only supports 2D exhibits and does not support a multi-user mode at the moment. Considering that exhibits vary across 2D and 3D forms, we are planning to extend CARDS to operate on 3D exhibits based on object recognition. Moreover, CARDS is currently a self-paced single-user application. However, taking into account that the traditional docent guide is conducted at a group level, CARDS can be extended to include a group-paced mode where a group leader holds control of the information delivery flow. An additional audio supplementary function could be added for the group-level option.

Acknowledgments

This work was supported by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government(MOTIE) (P0012746, The Competency Development Program for Industry Specialist). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agency.

References

[1] C. Ardito, P. Buono, M. F. Costabile, and G. Desolda. Interaction with large displays: A survey. *ACM Computing Surveys (CSUR)*, 47(3):1–38, 2015.

[2] T. Ballendat, N. Marquardt, and S. Greenberg. Proxemic interaction: designing for a proximity and orientation-aware environment. In *ACM International Conference on Interactive Tabletops and Surfaces*, pp. 121–130, 2010.

[3] K.-E. Chang, C.-T. Chang, H.-T. Hou, Y.-T. Sung, H.-L. Chao, and C.-M. Lee. Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum. *Computers & education*, 71:185–197, 2014.

[4] W. Chen, Y. Shan, Y. Wu, Z. Yan, and X. Li. Design and evaluation of a distance-driven user interface for asynchronous collaborative exhibit browsing in an augmented reality museum. *IEEE Access*, 9:73948–73962, 2021.

[5] C.-C. Chiu, W.-J. Wei, L.-C. Lee, and J.-C. Lu. Augmented reality system for tourism using image-based recognition. *Microsystem Technologies*, 27(4):1811–1826, 2021.

[6] N. Chung, H. Lee, J.-Y. Kim, and C. Koo. The role of augmented reality for experience-influenced environments: The case of cultural heritage tourism in korea. *Journal of Travel Research*, 57(5):627–643, 2018.

[7] L. Errichiello, R. Micera, M. Atzeni, and G. Del Chiappa. Exploring the implications of wearable virtual reality technology for museum visitors' experience: A cluster analysis. *International Journal of Tourism Research*, 21(5):590–605, 2019.

[8] J. Geigel, K. S. Shitut, J. Decker, A. Doherty, and G. Jacobs. The digital docent: Xr storytelling for a living history museum. In *26th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–3, 2020.

[9] R. Hammady and M. Ma. Designing spatial ui as a solution of the narrow fov of microsoft hololens: Prototype of virtual museum guide. In *Augmented reality and virtual reality*, pp. 217–231. Springer, 2019.

[10] R. Hammady, M. Ma, and A. Powell. User experience of markerless augmented reality applications in cultural heritage museums:'museumeye'as a case study. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pp. 349–369. Springer, 2018.

[11] D. Karadimas, L. Somakos, D. Bakalbasis, A. Prassas, K. Adamopoulou, and G. Karadimas. Current and potential applications of ar/vr technologies in cultural heritage."inception virtual museum hamh: A use case on bim and ar/vr modelling for the historical archive museum of hydra greece". In *International conference on transdisciplinary multispectral modeling and cooperation for the preservation of cultural heritage*, pp. 372–381. Springer, 2018.

[12] J.-K. Ko, D. W. Koo, and M. S. Kim. A novel affinity enhancing method for human robot interaction-preliminary study with proactive docent avatar. In *2021 21st International Conference on Control, Automation and Systems (ICCAS)*, pp. 1007–1011. IEEE, 2021.

[13] P. J. Locher. A measure of the information content of visual art stimuli for studies in experimental aesthetics. *Empirical studies of the arts*, 13(2):183–191, 1995.

[14] N. Marquardt and S. Greenberg. Proxemic interactions: From theory to practice. *Synthesis Lectures on Human-Centered Informatics*, 8(1):1–199, 2015.

[15] B. Neuhofer, D. Buhalis, and A. Ladkin. A typology of technology-enhanced tourism experiences. *International journal of tourism research*, 16(4):340–350, 2014.

[16] M. O'grady, G. O'hare, and S. Keegan. Interaction modalities in mobile contexts. In *Intelligent Interactive Systems in Knowledge-Based Environments*, pp. 89–106. Springer, 2008.

[17] J. Ping, Y. Liu, and D. Weng. Study on mobile ar guide system to enhance user experience in cultural heritage sites. In *2020 International Conference on Virtual Reality and Visualization (ICVRV)*, pp. 79–86. IEEE, 2020.

[18] L. Projects. Seeing picasso augmented reality app pace gallery palo alto.

[19] A. Sugiura, T. Kitama, M. Toyoura, and X. Mao. The use of augmented reality technology in medical museums. In *Teaching Anatomy*, pp. 337–347. Springer, 2020.

[20] L. Wein. Visual recognition in museum guide apps: do visitors want it? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 635–638, 2014.

[21] W. Xu, T.-T. Dai, Z.-Y. Shen, and Y.-J. Yao. Effects of technology application on museum learning: a meta-analysis of 42 studies published between 2011 and 2021. *Interactive Learning Environments*, pp. 1–16, 2021.