Natural-Language Scene Descriptions for Accessible Non-Visual Museum Exhibit Exploration and Engagement

Stacy A. Doore
Department of Computer Science, Bowdoin College, Brunswick, ME, USA
sdoore@bowdoin.edu

Anais C. Sarrazin
Department of Computer Science, Bowdoin College, Brunswick, ME, USA
asarraz@bowdoin.edu

Nicholas A. Giudice
Virtual Environment and Multimodal Interaction Lab, Department of Spatial Informatics, University of Maine, Orono, ME, USA
nicholas.giudice@maine.edu

Abstract

Indoor navigation and exploration of museum environments present unique challenges for visitors who are blind or have significant vision impairments (BVI). Like other indoor spaces, museums represent dynamic indoor environments that requires the need for both guided and self-tour experiences to allow for BVI visitor independence. In order to fully engage with a museum and its exhibits, BVI visitors need assistive technologies that support natural-language (NL) spatial descriptions that provide flexibility in the way system users receive descriptive information about gallery scenes and exhibit objects. In addition, the user interface must be connected to a robust database of spatial information to interact with mobile device tracking data and user queries. This paper describes the results of an early-stage demonstration project that utilizes an existing graph database model to support a NL information access and art gallery exploration system. Specifically, we investigated using a commercially available voice assistant interface to support NL descriptions of a gallery space and the art objects within it. Future work involves refining the language structures for scene and object descriptions, the integration of the voice assistant interface with tracking and navigation technologies, and additional user testing with sighted and BVI museum visitors.

2012 ACM Subject Classification Human-centered computing; Interface design prototyping

Keywords and phrases indoor environments, scene descriptions, voice assistant interface

Funding *Stacy A. Doore*: https://orcid.org/0000-0002-8322-6798
Nicholas A. Giudice: https://orcid.org/0000-0002-7640-0428

Acknowledgements We want to thank the Bowdoin College Museum of Art and the Bowdoin College InSite Design Team (Spring 2019).

1 Introduction

Think about your last visit to a museum. It might have been an art museum, a natural history museum, or a science museum. While you may have other sensory memories of your exploration of the museum exhibit space, much of the information about the space you were able to access was likely visual in nature. For example, there was the navigation of the museum's public spaces, the objects within the exhibit spaces, and the spatial configuration of the objects within the space. The objects displayed in the exhibit were probably arranged in a way that was visually appealing or perhaps provocative, however, they were likely completely inaccessible to touch. The exhibit placards, or an audio guide, might have provided a limited set of additional details about the object's origins, artist, or cultural significance but it

In Proceedings Speaking of Location 2019: Communicating about Space, Regensburg, Germany, September-2019.
Editors: K. Stock, C.B. Jones and T. Tenbrink (eds.);
Published at http://ceur-ws.org

probably was unlikely to provide enough information about the exhibit item to form an adequate conceptual map or model without the use of vision.

A critical aspect of a visit to a museum is the ability to freely explore its exhibition spaces according to one's own interests, making connections between personal lived experiences, and historical, and/or aesthetic context of the museum's collections. Although there have been many attempts to make museums of all types more accessible for BVI visitors, there are still relatively few solutions that focus on the problem of standardizing spatial language for multiple levels of scene descriptions as are necessary for a museum environment. Indoor navigation and exploration of museum environments present unique challenges for BVI users for a variety of reasons:

- Museums are dynamic indoor environments: Museum spaces are designed to be reconfigured regularly and the objects within gallery spaces are constantly changing.
- Museums require visitors to explore spaces to construct meaning: Independence of movement through a museum environment is a primary way in which visitors connect to museum exhibits and develop personal connections between exhibits and objects.
- Multiple granularities of spatial descriptions would be necessary for non-visual exploration of a museum environment: BVI visitors would require different levels of detail to provide natural-language information about route, gallery, exhibit objects.
- Museum collection databases are not designed to directly connect gallery space information to exhibit object information: Most museums rely on standardized relational database systems that are designed to store information for web publishing, object conservation, and digital asset management. However, this information is disconnected from the location and spatial configuration of objects in the physical museum space. These existing museum information systems are not designed to integrate tracking data, navigation, and natural-language voice assistant interface queries and commands.

Additional challenges for creating natural-language assistive technology applications revolve around: 1) collecting, storing, and representing the spatial granularity of the indoor space and objects within the space, 2) the need for tracking the BVI visitor within the indoor space to provide localized descriptions, and 3) a standardized set of non-visual based spatial descriptions representing different spatial scales and dimensions. System complexity arises from the nature of museums themselves. As buildings with both public and private spaces designed with various functions, traffic flows, and rules about visitor interactions with the dynamic galleries and changing objects within them. Often it is precisely the unique characteristics of museums that make them uninviting or inaccessible spaces for members of the BVI community who might want to explore the exhibits, yet find they are unable to do so without significant additional supports in place. In some cases, the costs of providing these types of assistive supports are beyond the scope of modest museum budgets despite the desire of their staff to create accessible spaces for all of the visiting public. This paper describes a proof of concept demonstration project that is designed to support BVI visitor exploration of a traditional art museum with automated spatial descriptions for the space (museum galleries), the collections in the space (exhibit objects and spatial relationships to one another), and the exhibit object scene descriptions (artwork spatial scenes).

2 Related Work

The creation of new information access technologies for BVI people is allowing for the exploration of indoor environments such as airports, shopping centers, and hospitals [15][22]. Advances in indoor location positioning technology has improved significantly in the past five years due to development of hybrid systems that use smart devices to receive and send signals from a combination of sources such as wifi, beacons, and near field communication tags [16]. These systems provide a mechanism for an individual to be tracked within the indoor environment at the building and room levels, while usually also receiving non-visual route directions to navigate and explore the indoor space.

However, within a museum environment, this type of information is rarely available or if present, likely requires additional support from a real-time person in the loop system to help navigate and describe the space and objects. These types of language-based support systems, while helpful, decrease independence and come with subscription costs for both the BVI user and the museum, therefore, they are relatively rare and not realistic for implementation in most medium and small size museums. For example, researchers [20] conducted a large-scale study of BVI accessibility features across 28 museums located within several major European cities. The results of the study were visualized in a matrix to demonstrate what types of accessibility strategies were used in the museums based on 35 potential accessibility strategies identified in a literature review. The researchers found less than half of the museums in the study had any non-visual accessibility supports in place at all. Three quarters of the museums that did have these types of assistive supports were more likely to implement non-visual accessibility features for only navigating the museum building, as opposed to accessing information about the actual exhibit spaces or exhibit objects. More specifically, museums reported avoiding implementing tactile representations due to costs and generally used a minimal level of descriptive information in their audio guides that replicated the same language for BVI visitors as was provided for sighted visitors. This replicated information was created for a sighted audience who could visually reference the artwork as they read the plaque or listened to the audio guide.

There are renewed efforts to better engage BVI visitors who require augmented natural-language location and object information. One approach has been to partner with BVI visitors as exhibit co-designers to yield new insights for museum accessibility solutions. In a recent study [4] the authors asked BVI focus group participants about what types of supports they believed were needed to independently explore a museum environment. In some cases, participants indicated they had never visited a museum independently, due to the inability to know where to navigate within the museum without having to rely on someone else for assistance. For those who had previously visited museums, participants indicated there were two primary accessibility barriers in museums they had experienced: 1) indoor navigation issues and 2) inaccessible information about the exhibit artwork and objects. The participants also reported that although many museums provide Braille exhibit books or audio guide information about exhibit objects, without detailed spatial descriptions about the gallery spaces and detailed object descriptions, they were unable to form an adequate conceptual model from the minimal information in the plaque descriptions. BVI participants also reported that they were most interested in having more control over their museum visit experience as well as the different types of information they could access about the space. Specifically, participants indicated they wanted an accessible application which they could use at their own leisure, providing indoor navigation information and detailed NL descriptions of the objects and their locations within the gallery spaces. Interviews conducted for the project described in this paper support these BVI user insights and system requirement suggestions. Several other studies [8][17] explored co-creation partnerships between museum curators and BVI community members to create more inclusive and engaging intellectual environments. As a result, a set of guidelines were developed to help museums in the implementation of co-designed exhibits to increase exhibit accessibility. Additional benefits of the co-design strategy were found to be the empathetic links that developed between the museum curators and the BVI collaborators.

Additional guidance for the design of NL museum supports is provided by the large body of route navigation and scene description literature, which demonstrates that humans construct spatial representations that reflect individualized needs, sensory capacities, and goals [21][28][24]. Cognitive processing of space is used for a variety of everyday tasks including wayfinding and navigation, interpreting visual symbolic representations (e.g., maps), using non-iconic spatial language (e.g., verbal route directions), and the forming of cognitive mental models [19][25]. Indoor scene descriptions are specifically designed to communicate spatial information about entities within indoor environments [10][23][13][11]. Automated annotation of information about indoor environments has been driven by innovations such as NL scene descriptions for use in robotic automation, indoor navigation applications, and location-based services for indoor public spaces. Understanding how people conceptualize and communicate about object relations in specific kinds of indoor spaces (e.g., airports, hospitals, shopping malls, and museums) is still a difficult problem for these automated systems because of the importance of context for spatial cues and determining this context is hard for current automated systems and algorithms [1]. Complexities of spatial language in indoor spaces can be seen in studies of the formation of structured vocabularies for describing indoor scenes [18], the roles that physical structure objects (e.g., walls, windows, doors, etc.) play in indoor scene descriptions [11], and how spatial relations are perceived and communicated in space [25][26][14][27].

Recently there has been a movement to improve online information access to museum collections for BVI individuals through text descriptions about the exhibit objects [5][2][12][6]. In some cases, the enhanced online verbal descriptions of individual objects within the museum collection provide information found on most exhibit plaques (e.g., artist, provenance, title, artistic medium, spatial dimensions, owner of the object, etc.). However, in addition to the regular plaque information. There is often a longer description that helps to orient the observer to object features using vivid and specific language, generally describing the visual concepts with other sensory analogies and providing contextual information about the historical and social significance of the work. In the case of [2], a Talking Museum application was developed using wifi and Bluetooth technology to detect an application on a visitor's smart device, thereby enabling an automated descriptive story about objects within the gallery space as the user freely explored the museum. This system was more focused on the automated task of detecting the visitor in proximity to the art object to trigger the description and less so on the quality or type of NL descriptions and accessibility features for BVI application users, however, the system does provide some real-time solutions to support independent exploration of museum spaces by BVI visitors. The prototype described in the remainder of this paper combined: 1) the guidelines for online enhanced object descriptions, 2) a flexible graph database model, and 3) location of BVI visitors through a real-time indoor tracking system within a traditional art gallery space.

3 InSite Design and NL Scene Language Descriptions

The InSite prototype was designed to use low cost, commercially available technologies to support an information system of scene representation for the space (gallery) and for the place (objects and their spatial and semantic relationships to one another), which provided more spatial information than the variety of alternative approaches (e.g., tactile maps, Braille plaques: recorded audio guides). This spatial language-based system was accessed by the visitor through a commercially available voice assistant and a Bluetooth headset allowing a BVI museum visitor to:

- explore the museum space and gallery objects through an online pre-journey tour,
- navigate within the real museum indoor environment of the museum galleries with descriptions of the different galleries and exhibits,
- follow route directions to explore the galleries based on a pre-journey list or impromptu exploration,
- receive a scene description of the gallery configuration, request a localized description of an area within the gallery with a list of the objects contained within that gallery area, and,
- select from a layered set of descriptions for individual objects within that area that provide general, spatial, and thematic descriptions based on the user's preference and interests.

The system uses an extended place graph model to support the layered description system of gallery and artwork based on existing BVI scene and artwork description guidelines. The InSite Prototype is designed to interact with low-energy Bluetooth beacons to assist in user tracking and navigation of gallery zones represented in the graph database. Artwork is organized by zones, and localization is achieved by interaction with NFC signals that connect through TCP to smartphone. The zone and painting NL descriptions are accessed through a smartphone screen reader (iOS) or through a voice assistant that is supported through two-way Bluetooth headset.

4 Extended place graph model

The museum project described in this paper employed an adapted spatial property graph model [7][27] that represented the gallery's collection and spatial information in the form of spatial triples (e.g., trajector, landmark, relation). The gallery was modeled as a series of three types of nodes: grid location nodes, painting/object nodes, and zone nodes. The grid location nodes were configured to represent a coordinate referenced space with x and y values given to each of the nodes. Each grid location node was given a property (occupied) consisting of a Boolean variable that represented whether or not it was occupied (i.e. someone could walk through the space). Grid location nodes were then connected to gallery zones (Zones 1-7). These zones represented sets of paintings that had a similar theme and were physically located within the same gallery zone. The zone nodes connected sets of paintings to a subset of grid location nodes, along with data on the zone location, name, and description of the theme of the zone. The painting nodes contained data about the painting, including the artist, the artistic medium, and several layers of descriptions designed to represent the painting using natural language descriptions. Tracking data from the gallery was provided to the InSite database in x and y coordinates to determine the specific grid location where the individual might be located (Figure 1). The database was designed to answer a set of queries to help the BVI user independently explore the gallery and its objects using both a smart phone mobile application. For example, if the user asked the system "What zone am I in?", they might receive the following information.

"You are in Zone 4, the ITALIAN Zone on the left wall of gallery. It contains seventeenth century Italian paintings, which include religious and mythological scenes, scenes of everyday life, and one portrait. The centerpiece is a large unattributed painting titled 'Fish Shambles', which portrays a scene of a fishmonger and a female customer. On the counter between them rests an impressive display of Mediterranean fish. Subjects in the other paintings include an enthroned Virgin Mary surrounded by adorers, Christ clearing a temple full of animals, and

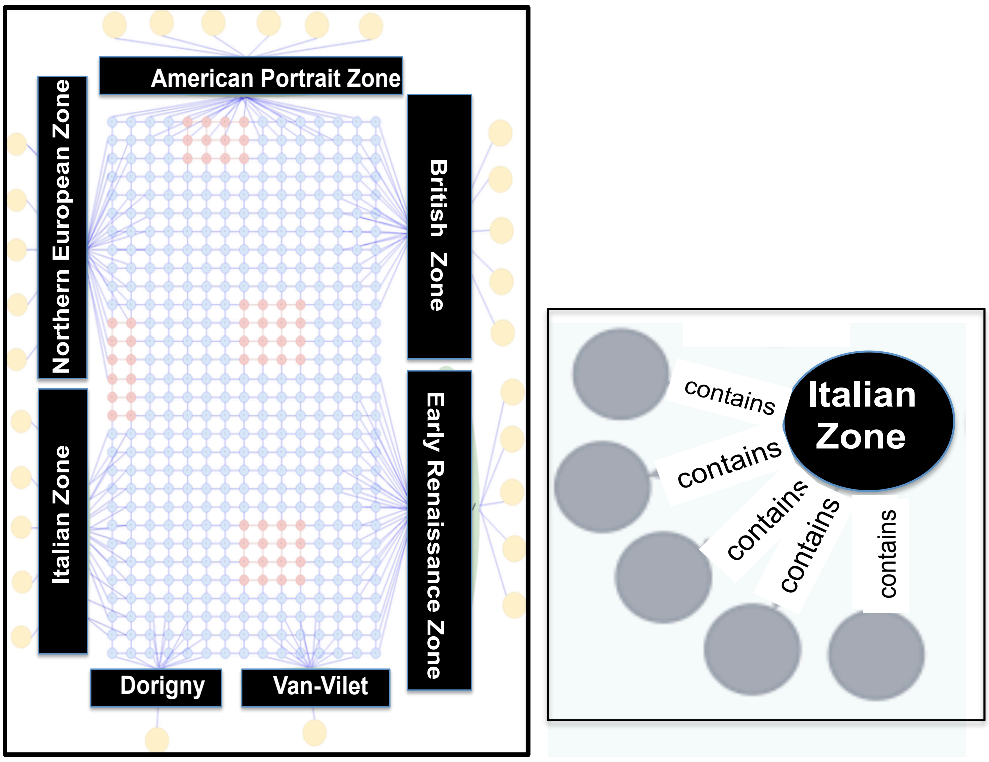

Figure 1 Figures 1 and 2: Visualization of InSite gallery model (Figure 1-left) and zone model with empty painting nodes (Figure 2-right).

Apollo chasing Daphne."

The graph structure of the database also allowed BVI visitors to use the navigation features of the mobile app. One could know, given their current location, which way they need to go in order to get to the gallery exit(s), or how to get to specific zones, without running into known obstacles (e.g., display cases, sculptures, etc.) in the room. Although the natural-language navigation interface was not fully implemented in this first version of the mobile application, the database structure would support future navigation features and queries. The flexible nature of the graph database model permits the addition of new nodes and relationships between nodes, as well as the reconfiguration of gallery objects based on the creation of new exhibits. This provides a spatial and temporal history of museum objects for the curators that moves far beyond what is traditionally available in a relational database spreadsheet or list of artwork and their respective locations. It also allows for the customization of different types of information about gallery contents that is responsive to diverse groups of visitors and their sensory information preferences.

5 Voice Assistant Interface

The primary user interface was implemented through an Amazon's Alexa Voice Service (AVS)[3], allowing for a rich set of verbal and auditory user interactions. Using Amazon's voice assistant interface skill developer toolkit, developers can create custom applications for use on Amazon Alexa-enabled devices. Developing and deploying a basic skill can take as little as 30 minutes, however, many skills can be tremendously complex, involve other

Bluetooth-enabled devices (e.g., headsets, smart-phones, etc.), APIs, and multiple layers of user input/output. Using the Alexa voice assistant interface in the demo project allowed the user to query information about the gallery such as a scene description of the gallery, description of sections of the gallery, pieces of artwork contained within the gallery, including spatial, curatorial and stylistic descriptions. As with most voice assistants, Alexa listens for key 'wake' words or command phrases and if the voice assistant does not understand the voice command that is given, it asks for a repeat of the command, additional input, or eventually quits the specified invocation or program. The demo system had several types of information the user could query using the voice interface such as:

- "Alexa, what gallery am I in?"
- "Alexa,what gallery zone am I in now?"
- "Alexa, what artwork is in this zone?"
- "Alexa, what zone contains the painting by Pontormo?"

For example, if the user wanted more information about a specific painting in the zone they would say "Alexa, give me a description of the Apollo and Daphne painting." The voice assistant would then ask the user if they would prefer a basic, spatial, or visual description of the painting that would provide various levels of detail about the artwork. The language structures in different types of descriptions are based on previous work and guidelines on non-visual scene descriptions such as the length of the individual sentence, the number of nouns and prepositions used in each sentence, the ordering of the scene description, and references to object scale (absolute and relative) [5][6][11]as well as information collected from interviews conducted with recent BVI museum visitors. These interviews made it clear that BVI visitors wanted flexibility in the types of information they could access from the prototype system. In some cases, they might only want the basic information available on the artwork's physical plaque or audio guide. In other cases, they wanted to know more about the spatial configuration of the objects within the artwork or extended thematic information that included the spatial information in the scene. [5][6]For example, if the user was interested in Pontormo's painting of Apollo and Daphne[9], they could request three different descriptions from the voice interface:

Basic: "This painting is titled Apollo and Daphne and was completed by Jacopo da Pontormo in 1513. It is a black and white oil on canvas. It is rectangular in shape. 24 3/8 inches tall by 19 1/4 inches wide. It is displayed in a bright metallic gold wooden frame."

Spatial: "The figures of Apollo and Daphne are located in the center of the portrait. The two figures occupy about 1/2 of the picture. Apollo is standing in the right of the portrait. He is reaching out to Daphne with his right hand. Daphne is located in the left of the painting. She is running with her back to Apollo. Daphne's fingers and arms are growing up into tree branches as she runs away from Apollo."

Visual: "Pontormo's haunting rendition of the tragic legend of Apollo and Daphne takes place in a shadowy black and white land. This portrait portrays the futile romance between Apollo and Daphne; Apollo is deeply in love with Daphne, a woman who abhors Apollo's advances. Located in the foreground center, Apollo is dressed in armor reaching out with his right hand towards Daphne's shoulder. As Apollo reaches out to lay his hand on Daphne, her fingers and arms in the left top background begin transforming into tree branches. In the portrait, we witness the initial instant of Daphne's transformation branches springing upward from her arms while the rest of her body remains in human form. The look of longing on Apollo's face is apparent, both of them appear to still be moving leftward out of the image, as if in mid-chase. The transformative depiction of Daphne signals the end of the chase as she turns into a tree, beginning with her hands and the top of her head. Apollo's touch reveals the futility of his desire, as well as the violence and persistence with which he pursued Daphne."

Each of these descriptions are stored as properties within the graph database and are triggered by key words and phrases built into the the voice interface skills. Although the NL descriptions may be improved by curators or by visitor feedback, the voice interface infrastructure and query labels themselves do not change, thus preserving the system commands while permitting new nodes and descriptions to be added or deleted in the gallery, without disruption or redesign of the system.

6 Prototype Limitations

The InSite prototype has distinct advantages over many current assistive technologies created for BVI museum visitors, such as interactive navigation, multiple levels of gallery/object spatial configuration information, and user choice in objects scene description information retrieval. However, as with any prototype there are some limitations associated with using a proprietary platform (e.g., AWS) which limits the system's ability to integrate other non-AWS technologies at the no/low cost level of this semester long project. Currently, the Alexa platform does not allow for two-way Bluetooth interactions, making it difficult for user NL queries to be heard clearly in a crowded space. User reaction latency is also often too long, causing problems with Alexa's default sleep setting that is activated after 16 seconds of inactivity. For novice users, it often takes more time than the Alexa will allow to learn to say the 'wake' words and command phrases in the correct order. In addition, the InSite graph database was built using an open-source Neo4j platform. Currently the AWS model does not allow for no-cost interoperability with other database development applications, therefore, a parallel AWS supported database was created to mirror the information contained in the orginal Neo4j database. This was a temporary solution for the prototype until another open-source voice assistant model could be developed that would be more suitable for accessibility applications and would meet the project requirements of no or low-cost design components. Due to the time limitations of the demo project and the temporary installation of the system, this version was only tested with sighted visitors. Subsequent versions will improve the integration of the navigation and museum route directions with both the mobile application and the voice interface.

7 Future Work

Future versions of the prototype will allow users to speak directly and discreetly into a Bluetooth headset to communicate with the voice assistant to interact with the gallery graph database, give a system query, request navigation information to move around the zones of the gallery, or exit the gallery to explore the rest of the museum. The visitor descriptions will continue to be refined and tested in order to train future versions of the system to generate suggested automated scene descriptions for curator and BVI visitor review.

References

1 Imad Afyouni, Cyril Ray, and Christophe Claramunt. Spatial models for context-aware indoor navigation systems: A survey. *Journal of Spatial Information Science*, (4):85–123, 2012.

2 Flora Amato, Angelo Chianese, Antonino Mazzeo, Vincenzo Moscato, Antonio Picariello, and Francesco Piccialli. The talking museum project. *Procedia Computer Science*, 21:114–121, 2013.

3 Amazon. Alexa voice service overview (v20160207). URL: `https://developer.amazon.com/docs/alexa-voice-service/api-overview.html`.

4 Saki Asakawa, João Guerreiro, Dragan Ahmetovic, Kris M Kitani, and Chieko Asakawa. The present and future of museum accessibility for people with visual impairments. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 382–384. ACM, 2018.

5 Elisabeth Salzhauer Axel and Nina Sobol Levent. *Art beyond sight: a resource guide to art, creativity, and visual impairment*. American Foundation for the Blind, 2003.

6 Sina Bahram and Anna Chiaretta Lavatelli. Using coyote to describe the world. In *Museums and the Web 2018*. Museums and the Web, 2018.

7 Hao Chen, Maria Vasardani, Stephan Winter, and Martin Tomko. A graph database model for knowledge extracted from place descriptions. *ISPRS International Journal of Geo-Information*, 7(6):221, 2018.

8 Anne Chick. Improving intellectual access in temporary exhibitions for sight loss visitors through co-creation and co-assessment. *The Design Journal*, 21(4):561–582, 2018.

9 Jacopo da Carrucci (called Pontormo). Apollo and daphne. URL: `http://artmuseum.bowdoin.edu/Obj4997?sid=6950&x=167853&port=4922`.

10 Mauro DiManzo, Giovanni Adorni, and Fausto Giunchiglia. Reasoning about scene descriptions. *Proceedings of the IEEE*, 74(7):1013–1025, 1986.

11 Stacy Anne Doore. *Spatial Relations and Natural-Language Semantics for Indoor Scenes*. PhD thesis, University of Maine, 2017.

12 Alison F Eardley, Louise Fryer, Rachel Hutchinson, Matthew Cock, Peter Ride, and Joselia Neves. Enriched audio description: Working towards an inclusive museum experience. In *Inclusion, disability and culture*, pages 195–207. Springer, 2017.

13 Zoe Falomir. Qualitative distances and qualitative description of images for indoor scene description and recognition in robotics. *AI Communications*, 25(4):387–389, 2012.

14 Nicholas A Giudice, Maryann R Betty, and Jack M Loomis. Functional equivalence of spatial images from touch and vision: Evidence from spatial updating in blind and sighted individuals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(3):621, 2011.

15 Nicholas A Giudice, William E Whalen, Timothy H Riehle, Shane M Anderson, and Stacy A Doore. Evaluation of an accessible, real-time, and infrastructure-free indoor navigation system by users who are blind in the mall of america. *Journal of Visual Impairment & Blindness*, 113(2):140–155, 2019.

16 João Guerreiro, Dragan Ahmetovic, Daisuke Sato, Kris Kitani, and Chieko Asakawa. Airport accessibility and navigation assistance for people with visual impairments. In *I n Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI'19)*, 2019.

17 Katrine Hesseldahl, Chris McGinley, and Georgia Monk. Using design thinking to develop new methods of inclusive exhibition making. *Studies in health technology and informatics*, 256:151–160, 2018.

18 Dahua Lin, Chen Kong, Sanja Fidler, and Raquel Urtasun. Generating multi-sentence lingual descriptions of indoor scenes. *arXiv preprint arXiv:1503.00064*, 2015.

19 Alan M MacEachren. Application of environmental learning theory to spatial knowledge acquisition from maps. *Annals of the Association of American Geographers*, 82(2):245–274, 1992.

20 Susana Mesquita and Maria João Carneiro. Accessibility of european museums to visitors with visual impairments. *Disability & Society*, 31(3):373–388, 2016.

21 Daniel R Montello. Scale and multiple psychologies of space. In *European conference on spatial information theory*, pages 312–321. Springer, 1993.

22 Paymon Rafian and Gordon E Legge. Remote sighted assistants for indoor location sensing of visually impaired pedestrians. *ACM Transactions on Applied Perception (TAP)*, 14(3):19, 2017.

23 Timothy H Riehle, P Lichter, and Nicholas A Giudice. An indoor navigation system to support the visually impaired. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4435–4438. IEEE, 2008.

24 Thora Tenbrink. Relevance in spatial navigation and communication. In *International Conference on Spatial Cognition*, pages 358–377. Springer, 2012.

25 Barbara Tversky and Herbert H Clark. Prepositions aren't places. *Behavioral and Brain Sciences*, 16(2):252–253, 1993.

26 Andrea Tyler and Vyvyan Evans. *The semantics of English prepositions: Spatial scenes, embodied meaning, and cognition.* Cambridge University Press, 2003.

27 Maria Vasardani, Sabine Timpf, Stephan Winter, and Martin Tomko. From descriptions to depictions: A conceptual framework. In *International Conference on Spatial Information Theory*, pages 299–319. Springer, 2013.

28 Steffen Werner and Kristine Schmidt. Environmental reference systems for large-scale spaces. *Spatial cognition and computation*, 1(4):447–473, 1999.