3DMovieMap: an Interactive Route Viewer
for Multi-Level Buildings

Seita Kayukawa
Waseda University
Tokyo, Japan
Miraikan - The National Museum of Emerging Science and
Innovation
Tokyo, Japan
k940805k@ruri.waseda.jp

Keita Higuchi
Preferred Networks
Tokyo, Japan
khiguchi@acm.org

Shigeo Morishima
Waseda Research Institute for Science and Engineering
Tokyo, Japan
shigeo@waseda.jp

Ken Sakurada
Artificial Intelligence Research Center, National Institute
of Advanced Industrial Science and Technology (AIST)
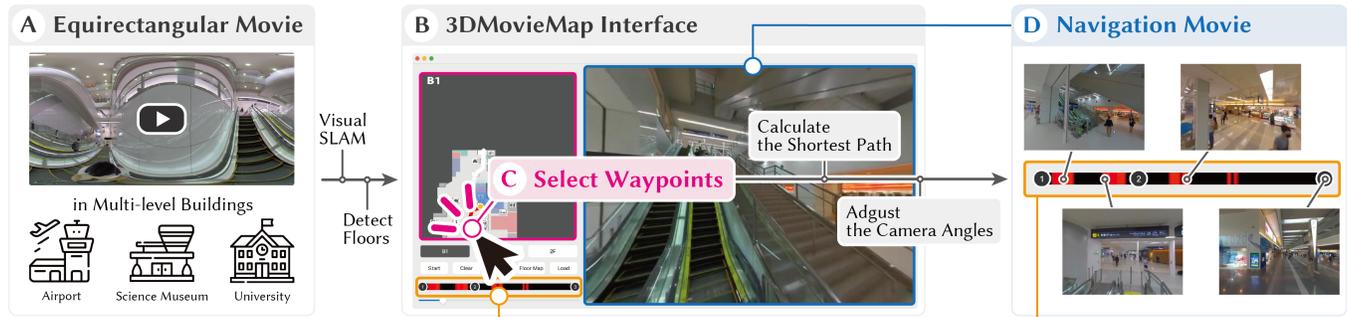Tokyo, Japan
k.sakurada@aist.go.jp

Figure 1: 3DMovieMap is an interactive route viewer system for multi-level buildings. A) The system uses an equirectangular movie recorded in a multi-level building. B) The system constructs a movie map by estimating camera positions via visual SLAM and detecting floors. C) Users select some waypoints on the floor maps. D) The system then calculates the shortest path that visits these points and generates a navigation movie along the path.

ABSTRACT

We present an interactive route viewer system, 3DMovieMap, which generates and shows navigation movies walking through multi-level buildings, such as a science museum, airport, and university building. Movie map systems can provide users with visual cues by synthesizing navigation movies based on their inputs of routes. However, existing systems are limited to flat areas such as city areas. We aim to extend Movie Map to generate navigation movies for multi-level buildings. The 3DMovieMap system generates a movie map from an equirectangular movie via a visual Simultaneous Localization and Mapping technology. Users select waypoints on the floor maps. 3DMovieMap calculates the shortest path that visits these points and generates a navigation movie along the route. We created four movie maps of buildings and asked two participants to use our system and provide feedback for further improvements. We will be releasing an open dataset of equirectangular movies captured in a science museum.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**.

KEYWORDS

movie map, equirectangular movie dataset, navigation, indoor environment

1 INTRODUCTION

Multi-level buildings such as shopping malls, airports, and museums have large and complex structures. Thus visitors often get lost in such buildings and have problems finding their way around public buildings [3, 12]. When people learn a route, they recognize visual landmarks around the route and/or create a cognitive map of the environment [2, 24]. Thus, previous works have proposed route viewer systems that present the users with both a map and visual cues that consist of photo streams [6, 10, 11, 14, 31] or a navigation movie captured along the route [4, 14, 34].

As a route viewer system, researchers have presented *Movie map* systems, which allow users to watch navigation movies along any route they choose [5, 17, 18, 27, 28, 32, 35]. Movie map systems generate a navigation movie along the user-selected route by connecting and switching multiple movie sequences captured in the environment. These systems map movie sequences on a two-dimensional (2D) map and estimate the positions of intersections where switch movie sequences by using metadata (e.g.,Global Positioning System (GPS) information [17]) or computer vision technologies (e.g.,visual Simultaneous Localization and Mapping (SLAM) [27, 28] or feature matching [18]). Unlike 2D environments such as city areas, which were the target of the existing movie map systems, multi-level buildings have overlapping floors. In such three-dimensional (3D) environments, if a system maps movie sequences on a 2D map like existing systems, the system can not detect intersections properly and fail to connect movie sequences. Thus, existing systems can not be introduced in multi-level buildings.

This work presents *3DMovieMap*, an interactive route viewer for multi-level indoor buildings. As a setup, the system first estimates the camera positions and orientations from an equirectangular movie by using Visual SLAM. The system then identifies the captured floor and generates a movie map by judging whether the camera paths intersect on each floor. Finally, we manually align the generated movie map on the existing floor maps. When visitors use the system, they select some waypoints they want to visit on the movie map. The system then automatically generates the shortest path that visits all selected points by using Dijkstra's algorithm. Finally, the system generates a navigation movie along the path by connecting multiple equirectangular sequences and extracting their perspective views. For generating smooth turning views, the system adjusts the orientations of the perspective views so that the camera orientations of the sequences at the connection point are matched.

In this work, we prepared four movie maps of public buildings, an international airport (terminals one and two), a science museum, and a university building. We asked two participants, including a user and a staff of the science museum, to use our system and collected their feedback about our system. They generally agreed that 3DMovieMap allowed users to easily learn their path even in a multi-level building, and the quality of turning views was enough to grasp the path. They also provided suggestions for improving the system and introducing it to the science museum. In addition, we will be releasing an open dataset of 8K equirectangular movies captured in the science museum.

2 RELATED WORK

2.1 Movie Map

The movie-map system is a route viewer system that generates navigation movies by connecting and switching multiple movie sequences at the correct position [5, 17, 18, 27, 28, 32, 35]. While many other route viewer systems can provide users with navigation movies along pre-captured routes only, movie-map systems can provide movies along any route that users choose on the map. For generating a navigation movie, these systems used panorama images [17], first-person view movies [18], equirectangular movies [27, 28], or photo streams from Google Street View [5, 32, 35]. Existing systems aligned multiple images or movies on a 2D map using GPS meta data [5, 17, 32, 35], feature matching [18], or visual SLAM [27, 28]. Sugimoto et al. proposed a visual SLAM-based Movie Map system [27, 28]. The system generates a movie map from equirectangular movies taken along streets in a city. The system localized frames of the movie on the 2D map based on the results of a visual SLAM [29] and then detected intersections. The system synthesizes turning views at detected intersections by blending the same oriented perspective view extracted from the two equirectangular movies. Since these systems mapped movie sequences on a 2D map, these systems can not detect intersections correctly in 3D structured environments such as multi-level buildings. To introduce a movie map in such 3D structured environments, we propose 3DMovieMap, which identifies the floor on which each movie sequence was captured based on the results of visual SLAM and detects intersections for each floor.

2.2 Difficulties in Multi-level Environments

When people learn a route, they often construct a cognitive map of the environment to avoid getting lost [21, 30]. When the environment has a 3D structure, people construct multi-level cognitive maps [15]. However, such a multi-level cognitive map potentially has errors, especially when floors overlap and the route contains z-directional movement (e.g.,going up/down a stair or riding an elevator) [3, 8, 12, 15, 16, 20, 22]. In such situations, people often get lost their position and orientation. One way to learn a route, besides constructing a cognitive map, is to recognize visual landmarks [2, 24]. Hirtle et al. showed that presenting visual landmarks is useful for understanding the spatial structure of the environment [24]. In this study, we target multi-level buildings (e.g.,an international airport, a science museum, and a university building), where it is difficult to learn a route. We propose a movie map system that can provide a navigation movie and visual landmarks in 3D-structured environments.

3 3DMOVIEMAP

Our main goal is to introduce a movie map system in multi-level buildings, such as airports and museums. The use case scenarios of our system are as follows:

- Building staff members capture an equirectangular movie while walking through as many routes as possible in the building. Then, they generate a movie map of the building by using our method.
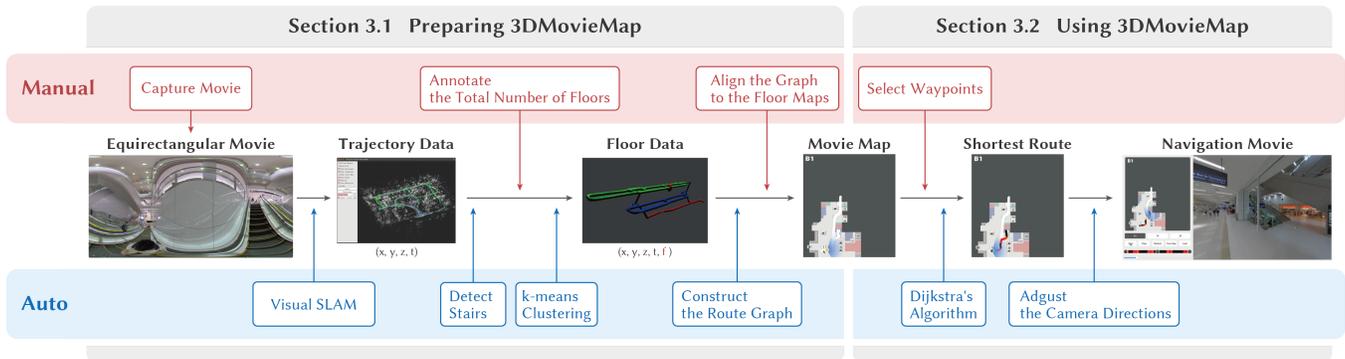
Figure 2: Overview of 3DMovieMap.

Table 1: Buildings where we captured equirectangular movies.

	video length	total numbers of floor
Science Museum	18:21	5
International Airport, Terminal One	12:16	3
International Airport, Terminal Two	14:50	3
University Building	14:58	4

- The movie map system is placed in the building (e.g.,at the entrance and near escalators or elevators). Visitors select some waypoints of interest on the interface and see the navigation movie generated by the system for learning their route.

Figure 2 shows an overview of 3DMovieMap. In the following sections, we will first describe how to set up 3DMovieMap from an equirectangular movie captured in a multi-level building (Section 3.1). Then, we will describe how to use the system and how to generate a navigation movie (Section 3.2).

3.1 Preparing 3DMovieMap

3.1.1 Input: Equirectangular Movie. We captured equirectangular movies (8K, 60FPS) to construct a movie map and generate a navigation movie. We attached a monopod to a 360-degree camera (Insta 360 Pro2) and held it above our heads while we walked around environments. In this work, we captured equirectangular movies in four environments (Table 1). Fig. 7–Fig. 10 show the floor maps of each environment with the scale. To allow the system to generate smooth and natural navigation movies, we considered the following points while capturing equirectangular movies: 1: the paths of multiple movie sequences should cross at each intersection; 2: the paths are as orthogonal as possible at intersections; and 3: the camera is at a constant height.

3.1.2 Estimating Camera Pose and Detecting Floors. We estimate the relative camera positions ($\mathbf{p}_n = (p_{xn}, p_{yn}, p_{zn})$) and orientations ($\mathbf{o}_n$) for each frame ($F_n$) of the equirectangular movie using a visual SLAM [29]. We then calculate the average of $\Delta p_{zn} = |p_{zn} - p_{z(n-1)}|$ as $\mu_{\Delta p_z}$. When $\Delta p_{zn} > \mu_{\Delta p_z}$, we judge that the frame (F_n) was captured while going up/down a stair, otherwise while walking a floor. For frames determined to be walking a floor,
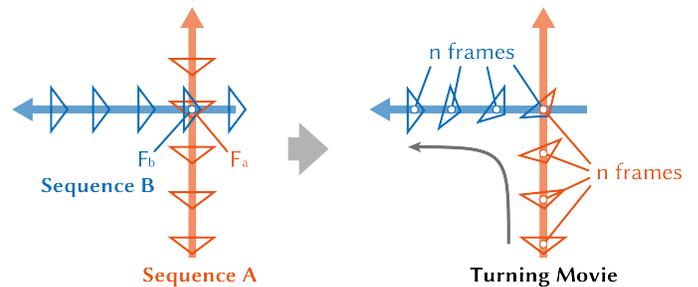

Figure 3: Synthesize turning views.

we apply k-means clustering to estimate the number of floors for each frame. We manually give the value of k (i.e.,total number of floors) for each environment.

3.1.3 Constructing and Aligning a Graph. For each floor, we construct a graph. We first map the 2D positions of each frame ((p_{xn}, p_{yn})) as nodes (N_n) and connect N_n and N_{n-1} as edges. Then, we insert a new node where the edges intersect. Finally, since visual SLAM estimates only relative camera positions, we manually map the graph onto the existing floor maps of the environment by aligning some key nodes, such as nodes of intersections or corners.

3.2 Using 3DMovieMap

3.2.1 Generating a Navigation Movie. After users select some waypoints they want to visit on the graph (Section 3.2.3), the system calculates the shortest path around each point in the order selected by the user by using Dijkstra's algorithm. For each frame along the shortest path, the system extracts perspective images in the walking direction from the equirectangular movie. Since the system uses equirectangular movies, the system can extract perspective images looking in the opposite direction from the walking direction. The system generates a navigation movie that moves in the opposite direction by playing perspective images looking in the opposite direction in reverse. The field of view (FoV) of the perspective images can be changed by users (default values are horizontal FOV = 120 degrees and vertical FOV = 90 degrees).

The system synthesizes the turning views for each intersection as shown in Fig. 3. When sequences A and B cross at an intersection,

the orientation of the perspective view at the intersection is in between the walking directions at A and B. For generating smooth turning views, the system adjusts the orientation of the perspective view using n frames approaching the intersection. In this work, we set $n = 180$.

3.2.2 Adaptively Fast-forwarded Navigation Movie. The system plays a fast-forwarded navigation movie at speed specified by the user. Previous works have presented route viewer systems that fast-forwarded navigation movies while keeping scenes turning intersections played at the original speed [4, 13]. Such adaptively fast-forwarded navigation movies are effective for recognizing the route quickly. Thus, the system also adaptively fast-forwards navigation movies as follows: Let o_n be the degree of walking direction at F_n. The system plays F_n at the original speed if $|o_{n+90} - o_{n-90}| > 60$, otherwise fast-forwards at speed specified by the user.

3.2.3 Details of User Interface. Fig. 4 shows the layout of 3DMovieMap. The system has two modes: 1) Main Mode, which shows a navigation movie (Fig. 4–9) and the map of the floor (Fig. 4–1) selected by the user (Fig. 4–2); and 2) Floor Map Mode, which shows all floor maps of the environment (Fig. 4–1).

Users select some waypoints they want to visit by clicking nodes on the floor map with a mouse. The selected waypoints are also displayed on the timeline (Fig. 4–8), and the user can change the order of the selected waypoints by dragging and dropping the icons on the timeline. As shown in Fig. 4–8, the timeline highlights sequences that are played at the original speed. The highlighted timeline is also visualized on the floor map (Fig. 4–1).

After the system generates a navigation movie, users can start (and pause) the navigation movie by pushing a button (Fig. 4–3). As shown in Fig. 4–4 to 4–7, the system contains some buttons for clearing selected waypoints, changing playback speed at three levels (original, five times, and ten times faster), changing two modes (Main mode and Floor map mode), and loading another map.

4 EQUIRECTANGULAR VIDEO DATASET OF A SCIENCE MUSEUM

This work captured equirectangular movies in four environments (Table 1). Among them, we will release equirectangular movies taken at the science museum as an open dataset on the web site of the science museum[1]. This dataset is freely available to anyone for research purposes. This dataset consists of three types of equirectangular movies. **1) Exploring the whole museum:** This movie was captured while walking through the whole area of the museum that is accessible to visitors (the first, third, fifth, and seventh floors). This video was used for our 3DMovieMap (Table 1, Fig. 11–A). **2) Exploring each exhibition area:** These movies were captured while walking through the exhibition area on the third and fifth floors (Fig. 11–B). **3) Exploring each exhibition:** These movies were captured while walking through the inside of each exhibit (Fig. 11–C). The total number of exhibitions is 16 (the third floor 7 and the fifth floor: 9). For all equirectangular movies, we prepared four types of movies featured combinations of with/without stabilization and two resolutions (4K and 8K).

[1]https://www.miraikan.jst.go.jp/en/research/AccessibilityLab/

This dataset has two unique points. First, The science museum has distinctive architecture as a part of its exhibitions. For example, a circular walkway (Fig. 6–B) goes around a globe-like display (Fig. 6–A), and escalators are in an open-air stairwell that goes from the first floor to the seventh floor (Fig. 6–D). Equirectangular movies that capture such distinctive architecture might be an interesting resource for computer vision research such as visual SLAM [29]. Second, the movies captured various objects such as a dome-shaped theater (Fig. 6–C), a rocket engine (Fig 12–A), and a scale model of the ISS living quarters (Fig. 12–B). High-resolution movies that capture these unique objects can be used for computer vision research such as 3D reconstructions and view synthesis [19].

5 PRELIMINARY STUDY

5.1 Procedure

As a preliminary study to collect feedback on the system for improving our system, we asked a user (P1, Female, 30 years old) to use the system. We also asked a museum staff member of the science museum (P2, Female, 31 years old) to provide feedback on our system from the perspective of facility managers. P1 has not visited all of the buildings for which the 3DMovieMap was created. P2 is familiar with the science museum, but she has not visited the international airport buildings and the university building. We first provided an overview of the study and described the interface (10 minutes). Then, we asked participants to watch navigation movies by using the system (about 20 minutes). We did not specify paths for viewing navigation movies, and participants were free to operate the system and watch navigation movies. Finally, we conducted a semi-structured interview session for about 30 minutes to receive qualitative feedback. Specifically, we first asked the participants about the advantages and disadvantages of our system. Then, we asked for suggestions to improve our system.

5.2 Results

Participants generally agreed that the system allowed users to learn their path in a multi-level building easily: A1:*"Since I may lose my orientation in public buildings, I may not be able to find the direction to the destination even if I look at the map of the buildings. For example, when I reach the floor where my destination is located by escalator or stairs, I may not know in which direction my destination is located. The system allowed me to learn the path visually. So I can understand the correct direction easily. The quality of the turning views was enough to grasp the path."* (P1); and A2:*"Our museum has introduced google map street view service. When comparing the service, the system allowed us to generate a course to walk over most spots in the museum more easily and grasp the overview in a shorter time. In addition, it was easier to understand the 3D and vertical structures such as the connection between floors (the circular walkway)."* (P2) In addition, a museum staff member also agreed that the system could be helpful for museum visitors: A3:*"The simulated walk is helpful to find new places that interest visitors."*

P1 (a user) suggested enabling the system to highlight visual landmarks in the generated navigation movies to improve the system: A4:*"When I walk in public buildings, I look at some visual landmarks, such as signage, stores, and objects, to determine which direction to go. So, when the system shows a turning view, how about highlighting*
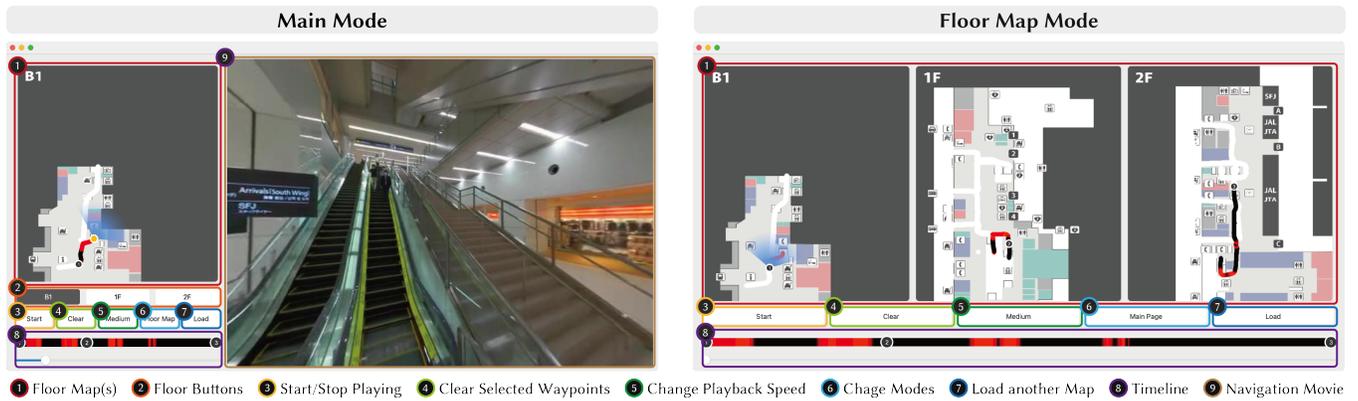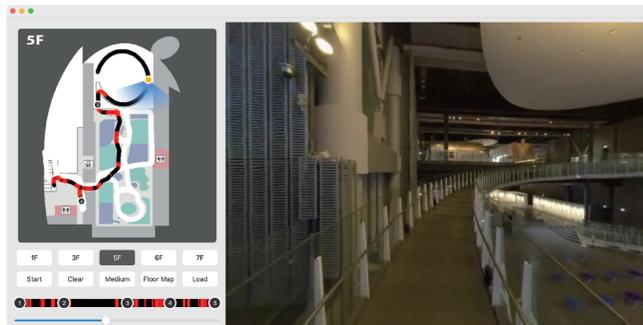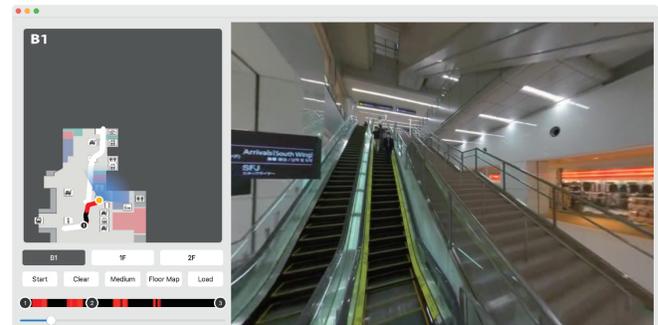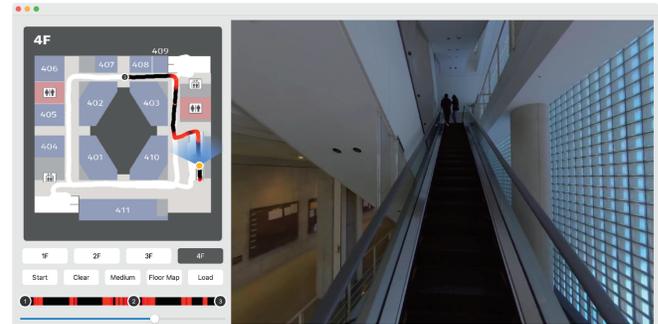
Figure 4: 3DMovieMap interface.

Figure 5: 3DMovieMap interface (examples).

such visual landmarks in the video, rather than simply slowing the video down? For example, the system can zoom in on a landmark or draw a rectangle to emphasize it." P2 (a museum staff member) provided the following suggestions for introducing the system in their museum: A5: *"The current system generates the shortest path, but it would be good to prioritize routes we would like visitors to walk through. For example, in our museum, visitors can move between the third and fifth floors via the circular walkway or an escalator. Unless it is a very long way, we want them to walk the circular walkway. In addition, I think it would be good to avoid taking the route that users*

have already watched so that visitors can see more paths and contents of the museum."; and A6: *"It would be nice if the system slowed the video down when moving near an exhibit and generated a movie with the camera pointed in the direction of the exhibit (rather than in the walking direction)."*
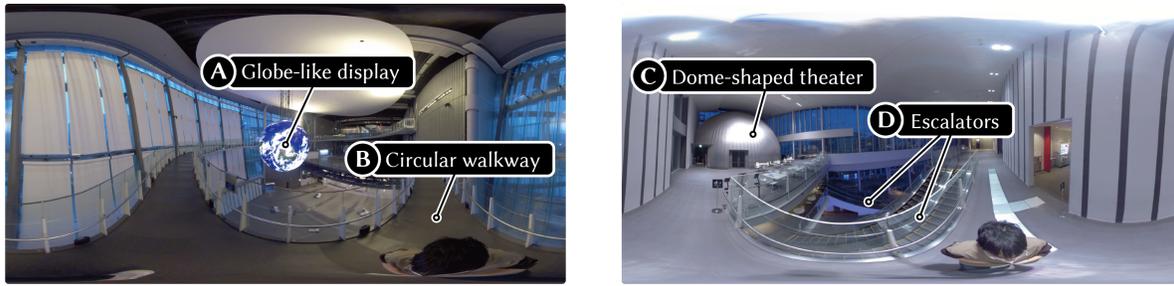
Figure 6: Examples of equirectangular movies in our dataset. The science museum has distinctive architecture and exhibitions such as a circular walkway (B) that goes around a globe-like display (A), a dome-shaped theater (C), and escalators in an open-air stairwell (D).

6 DISCUSSION

6.1 Effectiveness of 3DMovieMap in Multi-level Buildings

Participants reported that 3DMovieMap were useful for learning their path (A1) and grasping the 3D/vertical structures of the environment (A2). They also agreed that the quality of turning views generated by our system was enough to grasp the path (A1) and allowed us to grasp the overview of the path shorter time than Google Street View, an interactive panorama images viewer system (A2). The museum staff member (P2) expressed a favorable opinion of introducing this system to the museum because the simulated walk with the system is helpful in finding new places that interest visitors (A3).

6.2 Future Work: for more Practical System

In the experiments, we received some suggestions for improving the system. First, although the system adaptively fast-forwards the movie while slowing the video down at the intersection, P1 suggested that the system can highlight visual landmarks, such as signage, stores, and objects, in the movie at the same time as slowing down (A4). For example, visitors at public buildings can find their way by referencing signage that shows directions toward points of interest [7, 9]. In addition, the museum staff member (P2) suggested generating a navigation movie with the camera pointed in the direction of the exhibit (A6). Future solutions may consider combining visual landmarks information presented by building managers' annotation or signage recognition technologies [1, 23, 33] and movie trimming algorithms that automatically control a virtual camera within an equirectangular movie [25, 26].

Second, P2 commented that the path generation process that calculates the shortest path could be improved to provide users with a path that walks that the building manager would like visitors to walk through (A5). While the system defines the cost of the edges of the movie map based on the relative position estimated via visual SLAM and calculates the shortest path by using Dijkstra's algorithm, the cost can be set adaptively. For example, we can set the cost of the building's main streets lower and boring paths with nothing to see higher. In addition, we plan to automatically update the costs of the routes based on the user's past navigation video viewing history (i.e.,the cost is set higher for routes that users have

viewed in the past and lower for routes that they have not viewed yet).

7 CONCLUSTION

We proposed an interactive route viewer system, 3DMovieMap, that aims to provide users with navigation movies along any route they choose in multi-level buildings such as a science museum, airport, and university building. The system first estimates the camera positions and orientations of an equirectangular movie using visual SLAM and detects the floors of each frame based on the results of the SLAM. The system then generates a 3D structured movie map by detecting intersections for each floor. When users select some waypoints they want to visit on the interface, the system calculates the shortest path that visits all selected points. Then, the system generates a smooth navigation movie along the path by connecting multiple equirectangular movie sequences and extracting perspective views while controlling the camera angle to match the angles of movie sequences at the connection point. We constructed four movie maps of public buildings (a science museum two international airport terminals, and a university building) and asked two participants, a user and a science museum staff member, to use the system. They generally agreed that 3DMovieMap allowed users to easily learn their path even in a multi-level building, and the quality of turning views was enough to grasp the path. In the future, we will extend our system based on the feedback from the participants and conduct a user study to evaluate its effectiveness in learning routes in multi-floor buildings. For example, we have plans to highlight visual landmarks in the navigation movie and recommend a path based on the contents of the floor or the user's past navigation video viewing history, except the shortest path. In addition, we will be releasing an open dataset of 8K equirectangular movies captured in the science museum.

REFERENCES

[1] Mouna Afif, Yahia Said, Edwige Pissaloux, Mohamed Atri, et al. 2020. Recognizing signs and doors for Indoor Wayfinding for Blind and Visually Impaired Persons. In *Proceedings of the 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP '20)*. IEEE, Los Alamitos, CA, USA, 1–4. https://doi.org/10.1109/ATSIP49331.2020.9231933

[2] Vlada Aginsky, Catherine Harris, Ronald Rensink, and Jack Beusmans. 1997. Two strategies for learning a route in a driving simulator. *Journal of Environmental Psychology* 17 (12 1997), 317–331. https://doi.org/10.1006/jevp.1997.0070

[3] Laura A Carlson, Christoph Hölscher, Thomas F Shipley, and Ruth Conroy Dalton. 2010. Getting lost in buildings. *Current Directions in Psychological Science* 19, 5 (2010), 284–289. https://doi.org/10.1177/0963721410383243

[4] Billy Chen, Boris Neubert, Eyal Ofek, Oliver Deussen, and Michael F. Cohen. 2009. Integrated Videos and Maps for Driving Directions. In *Proceedings of of the 22nd Annual ACM Symposium on User Interface Software and Technology (UIST '09)* (Victoria, BC, Canada) *(UIST '09)*. ACM, New York, NY, USA, 223–232. https://doi.org/10.1145/1622176.1622218

[5] Peng Chi, Chen Bing-Yu, and TsaiC Chi-Hung. 2010. Integrated Google Maps and smooth street view videos for route planning. In *Proceedings of 2010 International Computer Symposium (ICS '10)*. IEEE, Washington, DC, USA, 319–324. https://doi.org/10.1109/COMPSYM.2010.5685494

[6] Hadziq Fabroyir, Wei-Chung Teng, Shu-Ling Wang, and Rayi Yanu Tara. 2013. MapXplorer Handy: An Immersive Map Exploration System Using Handheld Device. In *Proceedings of of the 2013 International Conference on Cyberworlds (CW '13)*. IEEE, Washington, DC, USA, 101–107. https://doi.org/10.1109/CW.2013.64

[7] Rodney Fewings. 2001. Wayfinding and airport terminal design. *The journal of navigation* 54, 2 (2001), 177–184.

[8] Nicholas A Giudice and Hengshan Li. 2012. The effects of visual granularity on indoor spatial learning assisted by mobile 3D information displays. In *Proceedings of International Conference on Spatial Cognition*. Springer, New York, NY, USA, 163–172. https://doi.org/10.1007/978-3-642-32732-2_10

[9] International Health Facility Guidelines. 2016. Wayfinding Guidelines International Health Facility Guidelines. http://www.healthfacilityguidelines.com/GuidelineIndex/Index/Wayfinding-Guidelines

[10] Harlan Hile, Radek Grzeszczuk, Alan Liu, Ramakrishna Vedantham, Jana Košecka, and Gaetano Borriello. 2009. Landmark-Based Pedestrian Navigation with Enhanced Spatial Reasoning. In *Proceedings of of the 7th International Conference on Pervasive Computing (Pervasive '09)* (Nara, Japan). Springer-Verlag, Berlin, Heidelberg, 59–76. https://doi.org/10.1007/978-3-642-01516-8_6

[11] Harlan Hile, Ramakrishna Vedantham, Gregory Cuellar, Alan Liu, Natasha Gelfand, Radek Grzeszczuk, and Gaetano Borriello. 2008. Landmark-Based Pedestrian Navigation from Collections of Geotagged Photos. In *Proceedings of the 7th International Conference on Mobile and Ubiquitous Multimedia (MUM '08)* (Umeå, Sweden). ACM, New York, NY, USA, 145–152. https://doi.org/10.1145/1543137.1543167

[12] Christoph Hölscher, Tobias Meilinger, Georg Vrachliotis, Martin Brösamle, and Markus Knauff. 2006. Up the down staircase: Wayfinding strategies in multi-level buildings. *Journal of Environmental Psychology* 26, 4 (2006), 284–299. https://doi.org/10.1016/j.jenvp.2006.09.002

[13] Hui Huang, Dani Lischinski, Zhuming Hao, Minglun Gong, Marc Christie, and Daniel Cohen-Or. 2016. Trip Synopsis: 60km in 60sec. *Computer Graphics Forum* 35, 7 (2016), 107–116. https://doi.org/10.1111/cgf.13008

[14] Thomas H Kolbe. 2004. Augmented videos and panoramas for pedestrian navigation. In *Proceedings of the 2nd Symposium on Location Based Services & TeleCartography 2004, 28-29th of January 2004 in Vienna*. Springer, New York, NY, USA, 10 pages. https://doi.org/10.1.1.112.1759

[15] Hengshan Li and Nicholas A Giudice. 2012. Using mobile 3D visualization techniques to facilitate multi-level cognitive map development of complex indoor spaces. *Spatial Knowledge Acquisition with Limited Information Displays* 21 (2012), 31–36.

[16] Hengshan Li and Nicholas A. Giudice. 2013. The Effects of 2D and 3D Maps on Learning Virtual Multi-Level Indoor Environments. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on MapInteraction (MapInteract '13)* (Orlando, Florida) *(MapInteract '13)*. ACM, New York, NY, USA, 7–12. https://doi.org/10.1145/2534931.2534941

[17] Andrew Lippman. 1980. Movie-Maps: An Application of the Optical Videodisc to Computer Graphics. *ACM Trans. Graph.* 14, 3 (July 1980), 32–42. https://doi.org/10.1145/965105.807465

[18] Wang Miao, Liang Jun-Bang, Zhang Song-Hai, Lu Shao-Ping, Shamir Ariel, and Hu Shi-Min. 2018. Hyper-Lapse From Multiple Spatially-Overlapping Videos. *IEEE Transactions on Image Processing* 27, 4 (2018), 1735–1747. https://doi.org/10.1109/TIP.2017.2749143

[19] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. 2020. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In *Proceedings of the 16th European Conference on Computer Vision (ECCV '20)*, Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm (Eds.). Springer International Publishing, Cham, 405–421.

[20] Daniel R Montello and Herbert L Pick Jr. 1993. Integrating knowledge of vertically aligned large-scale spaces. *Environment and Behavior* 25, 3 (1993), 457–484. https://doi.org/10.1177/0013916593253002

[21] John O'Keefe. 1991. An allocentric spatial model for the hippocampal cognitive map. *Hippocampus* 1, 3 (1991), 230–235. https://doi.org/10.1002/hipo.450010303

[22] Romedi Passini. 1984. Wayfinding in architecture. *Environmental Design* 4 (1984), 57 pages.

[23] Manaswi Saha, Alexander J Fiannaca, Melanie Kneisel, Edward Cutrell, and Meredith Ringel Morris. 2019. Closing the gap: Designing for the last-few-meters wayfinding problem for people with visual impairments. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19)*. ACM, New York, NY, USA, 222–235. https://doi.org/10.1145/3308561.3353776

[24] Hirtle Stephen C. 2019. . Morgan & Claypool Publishers, Oak View Drive San Rafael, CA, USA. https://doi.org/10.2200/S00921ED2V01Y201904HCI043

[25] Yu-Chuan Su and Kristen Grauman. 2017. Making 360° Video Watchable in 2D: Learning Videography for Click Free Viewing. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR '17)*. IEEE, Washington, DC, USA, 1368–1376. https://doi.org/10.1109/CVPR.2017.150

[26] Yu-Chuan Su, Dinesh Jayaraman, and Kristen Grauman. 2017. Pano2Vid: Automatic Cinematography for Watching 360° Videos. In *Proceedings of the 13th Asian Conference on Computer Vision (ACCV '17)*, Shang-Hong Lai, Vincent Lepetit, Ko Nishino, and Yoichi Sato (Eds.). Springer International Publishing, Cham, 154–171. https://doi.org/10.1007/978-3-319-54190-7_10

[27] Naoki Sugimoto, Yoshihito Ebine, and Kiyoharu Aizawa. 2020. Building Movie Map - A Tool for Exploring Areas in a City - and Its Evaluations. In *Proceedings of the 28th ACM International Conference on Multimedia (ACMMM '20)*. ACM, New York, NY, USA, 3330–3338. https://doi.org/10.1145/3394171.3413881

[28] Naoki Sugimoto, Toru Okubo, and Kiyoharu Aizawa. 2020. Urban Movie Map for Walkers: Route View Synthesis Using 360° Videos. In *Proceedings of the 2020 International Conference on Multimedia Retrieval (ICMR '20)* (Dublin, Ireland). ACM, New York, NY, USA, 502–508. https://doi.org/10.1145/3372278.3390707

[29] Shinya Sumikura, Mikiya Shibuya, and Ken Sakurada. 2019. OpenVSLAM: A Versatile Visual SLAM Framework. In *Proceedings of the 27th ACM International Conference on Multimedia (ACMMM '19)*. ACM, New York, NY, USA, 2292–2295. https://doi.org/10.1145/3343031.3350539

[30] Edward C Tolman. 1948. Cognitive maps in rats and men. *Psychological review* 55, 4 (1948), 189. https://doi.org/10.1037/h0061626

[31] Benjamin Walther-Franks and Rainer Malaka. 2008. Evaluation of an Augmented Photograph-Based Pedestrian Navigation System. In *Proceedings of of the 9th International Symposium on Smart Graphics (SG '08)* (Rennes, France). Springer-Verlag, Berlin, Heidelberg, 94–105. https://doi.org/10.1007/978-3-540-85412-8_9

[32] Quanfa Xiu, Xiaorong Shen, Tianlong Zhang, and Luodi Zhao. 2018. An Immersive Roaming Method Based on Panoramic Video. In *Proceedings of of 2017 Chinese Intelligent Systems Conference (CICS '17)*, Yingmin Jia, Junping Du, and Weicun Zhang (Eds.). Springer Singapore, Singapore, 431–441. https://doi.org/10.1007/978-981-10-6499-9_41

[33] Yutaro Yamanaka, Seita Kayukawa, Hironobu Takagi, Yuichi Nagaoka, Yoshimune Hiratsuka, and Satoshi Kurihara. 2021. One-Shot Wayfinding Method for Blind People via OCR and Arrow Analysis with a 360-degree Smartphone Camera. In *Proceedings of the 18th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '21)*. Springer, New York, NY, USA, 19 pages. https://doi.org/10.1007/978-3-030-94822-1_9

[34] Guangqiang Zhao, Mingjin Zhang, Tao Li, Shu-ching Chen, and Naphtali Rishe. 2015. City recorder: Virtual city tour using geo-referenced videos. In *Proceedings of 2015 IEEE International Conference on Information Reuse and Integration (IRI '15)*. IEEE, Washington, DC, USA, 281–286. https://doi.org/10.1109/IRI.2015.53

[35] Qiang Zhao, Liang Wan, Wei Feng, Jiawan Zhang, and Tien-Tsin Wong. 2013. Cube2Video: Navigate Between Cubic Panoramas in Real-Time. *Trans. Multi.* 15, 8 (2013), 1745–1754. https://doi.org/10.1109/TMM.2013.2280249
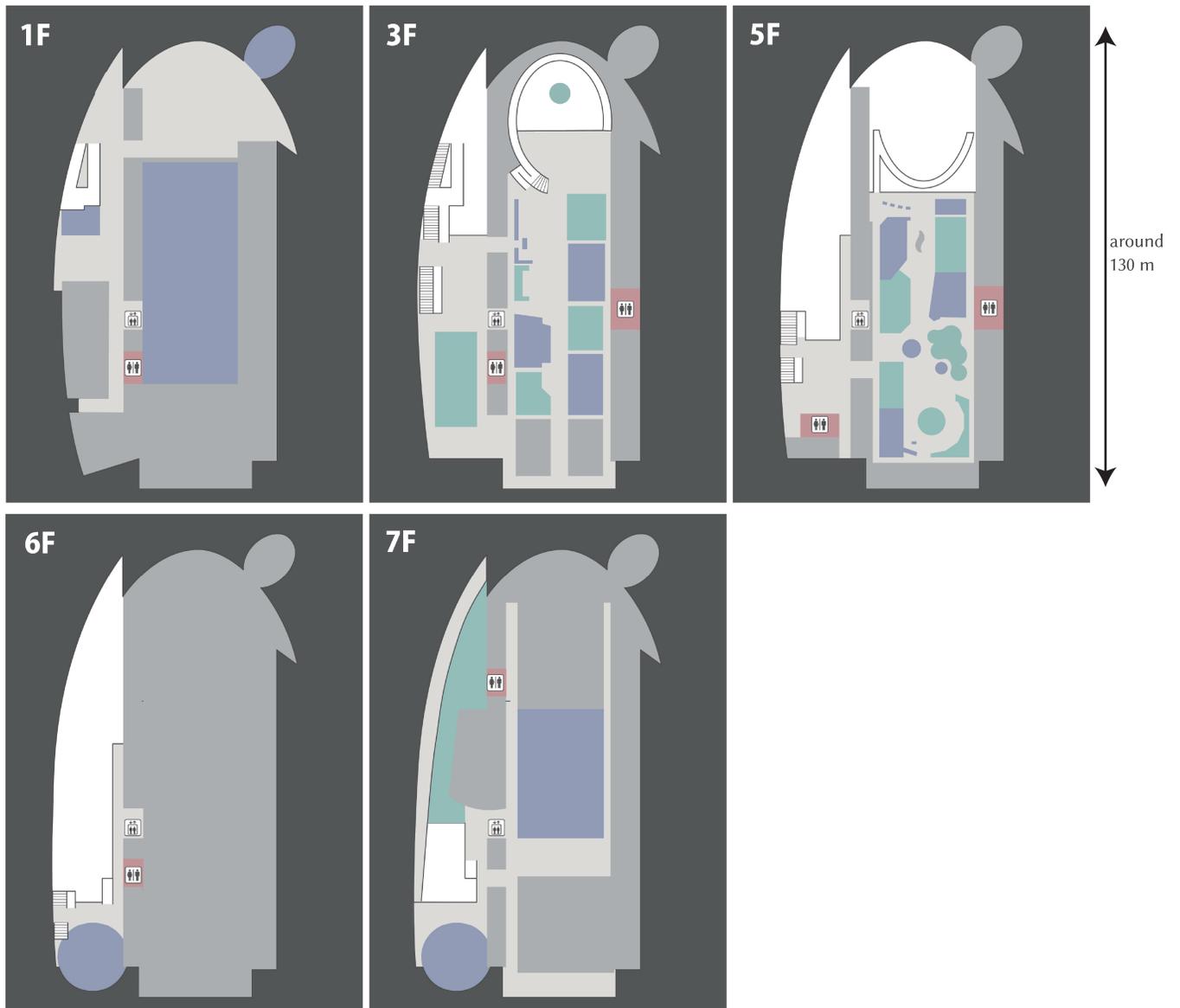
Science Museum

Figure 7: Floor maps of the science museum.

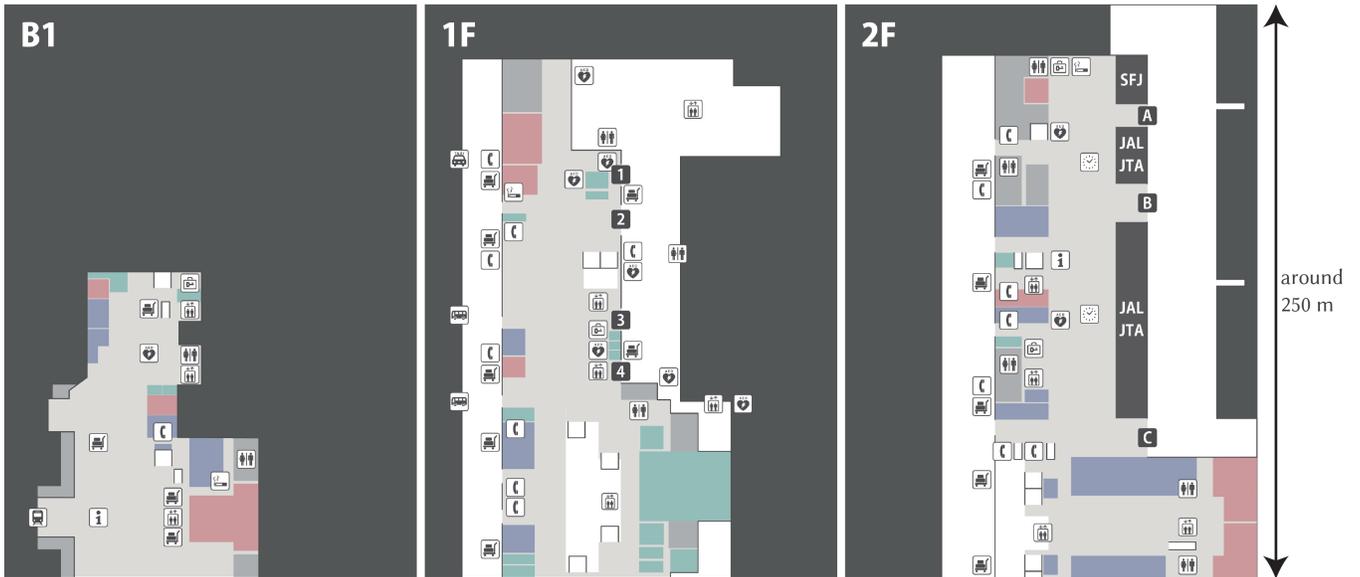**International Airport, Terminal 1**

Figure 8: Floor maps of the international airport, terminal 1.
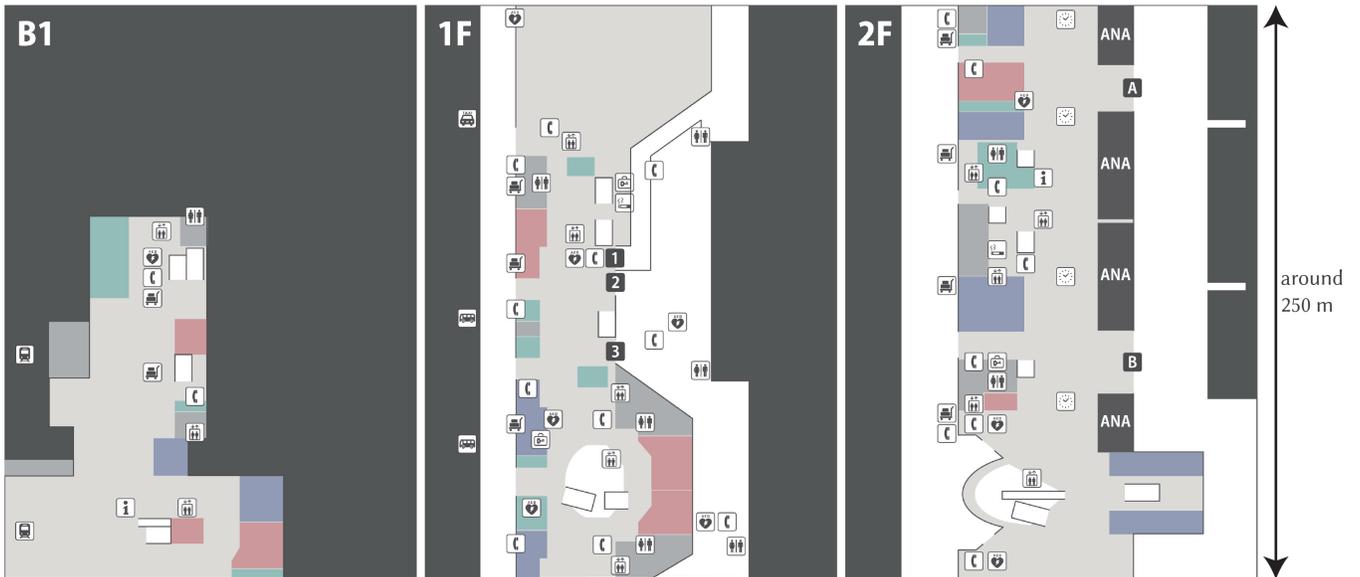
International Airport, Terminal 2

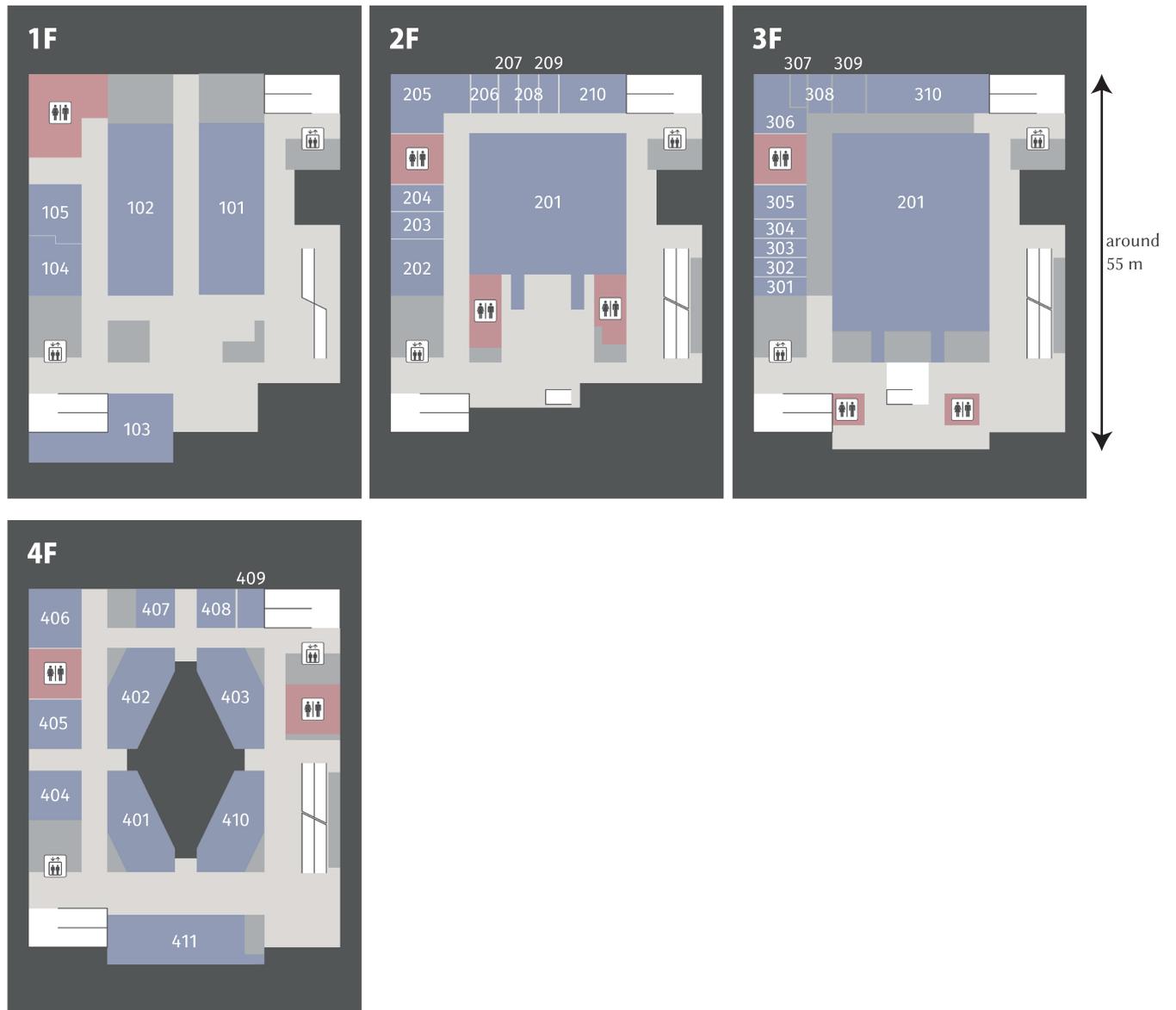**Figure 9: Floor maps of the international airport, terminal 2.**

University Building

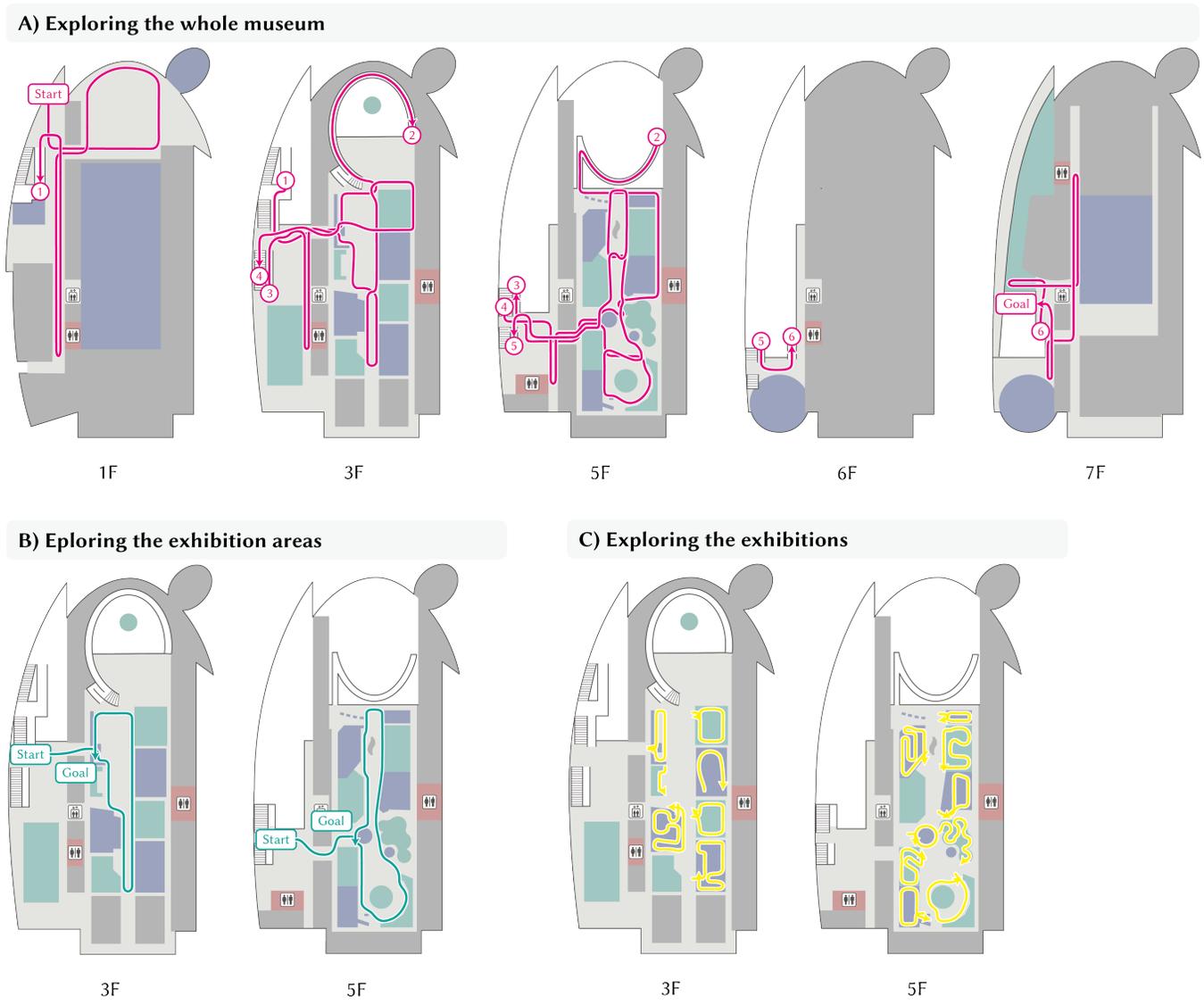Figure 10: Floor maps of the international university building.

A) Exploring the whole museum

1F   3F   5F   6F   7F

B) Eploring the exhibition areas

3F   5F

C) Exploring the exhibitions

3F   5F

Figure 11: Routes where we recorded equirectangular movies in the science museum.

A) Rocket engine   **B) Scale model of the ISS living quarters**

Figure 12: Examples of the exhibitions captured in our dataset.