UNIVERSITY OF CALIFORNIA
SANTA CRUZ

SEMANTIC INTERIOR MAPOLOGY: AN END-TO-END SYSTEMATIC TOOL FOR REPRESENTING THE SPATIAL STRUCTURE OF AN INDOOR ENVIRONMENT

A dissertation submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

COMPUTER ENGINEERING

by

Viet Q. Trinh

December 2020

The Dissertation of Viet Q. Trinh is approved by:

Professor Roberto Manduchi, Chair

Professor Luca de Alfaro

Professor David Lee

Quentin Williams
Interim Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Semantic Interior Mapology: An End-to-end Systematic Tool for Representing the Spatial Structure of an Indoor Environment

by

Viet Q. Trinh

Due to the lack of visual cues and access to detailed environmental information, many visually impaired people are reluctant to travel independently in unfamiliar locations. Several lines of research have addressed the problem of wayfinding for blind individuals, mainly focusing on localization and guidance systems. Even when these systems are available, pre-journey learning is a valuable resource, helping blind people to create a mental image of their surrounding and to maintain their orientation in the case of system malfunction.

In this dissertation, I introduce an end-to-end systematic tool for a collective spatial mapping of an indoor environment consisting of small-scale features (i.e., building fixtures, room furnitures, types of floor-covering, etc), and for the production of tactile maps at multiple scales, promoting pre-journey spatial awareness. This systematic application, named Semantic Interior Mapology or SIM, is a three-fold. The first component, Map Conversion Tool, allows one to quickly and accurately trace a floor plan from an architectural image of it. Next, the Map Population Tool component segments out small-scale elements of interest from a 3D scan of a room, and then geo-registers them within the building's spatial layout. These initial semantic spatial relationship is stored as a vectorized format that is amenable to reproduction in multiple modalities. The last component, Map Authoring Tool, produces on demand the tactile maps of indoor environments from the building's structural layouts and its 3D-scanned interior spatial contents that are captured and vectorized previously. Such maps are embossed at different spatial scales, representing the building's general structure, a zoomed-in of a specific area, or an interior of a room, with specific constraints on the density and distances of tactile features. My end-to-end systematic application described in this dissertation minimizes the time and effort required to acquire a detailed description of

an indoor space, and produces accurate results even in the case of complex building layouts.

Dedicated to my grandfather, Major Ngoc Quang Trinh, who sacrificed his life for my achievements today ...

Acknowledgments

I would like to express my deepest appreciation to my advisor, Dr. Roberto Manduchi. His guidance and support during my Ph.D. study has shaped the researcher in me substantially: being curious for the unknown, questioning the existing knowledge, and applying new findings into serving the underrepresented segments of our society. I would also like to extend my sincere thanks to Dr. Luca de Alfaro, Dr. David Lee, and Dr. Jim Whitehead for being members of my committee, especially during this unprecedented time of Covid outbreak. Many thanks to my Computer Vision's lab mates for fruitful discussions and moments of inspiration. Finally, I'm deeply indebted to my parents, sister, and brother for their unconditional love, amazing encouragement, and never-ending support.

Chapter 1

Introduction

1.1 Spatial Awareness for Visually Impaired Individuals Before a Self-travelling

Wayfinding for people with visual impairments (VI) has mainly concentrated on localization and guidance to known destinations. Current technical approaches incorporate a rich set of inertial sensor readings and radio-frequency fingerprint matching (e.g., Wifi access points, iBeacon nodes) to track a navigation trace [11]. Some localize users from a photo of their current location, through features mapping with labeled key visual elements such as signs, doors, windows, and fixtures [89, 30]. Others employ advancement in audio guidance systems to provide acoustic or verbal directions to a desired destination from the current user whereabout, which can be tracked via Wifi triangulation [20], Bluetooth Low Energy beacons [2], RFIDs [24], or computer vision systems [57]. Even when these systems are available, a prior detailed and location-specific environmental information, which is useful for blind travelers create a global mental image of the environment and maintain orientation in the case of system malfunctioning, is a valuable resource.

Pre-journey learning is the process of familiarizing oneself to a spatial environment, or planning a travel route, prior to an actual travel. Multiple studies have shown that, when visually impaired travelers are given the opportunity to preview a route, they are able to follow it more accurately and with fewer errors [35, 6, 36]. Such route overviews often come in the form of either a verbal or a tactile description. In

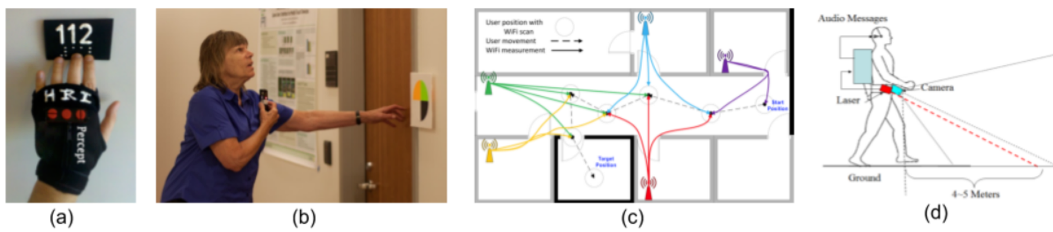


Figure 1.1: Wayfinding technologies for VI people: (a) Indoor navigation system recognizes room numbers in Braille that are pre-installed in the passive RFID tags, and then provides a verbal instruction [24], (b) A sign-based wayfinding application uses a camera cell phone to detect specific color marker [57], (c) WiFi-guided system instructs a loop closure for indoor navigation using mobile devices [66], and (d) A context-aware-based audio system integrates the laser data and image edge profiles for guidance [47].

the first case, people with visual impairments employ their problem solving skills and their Orientation and Mobility (O&M) training to contextualize a planned verbal route description. These verbal routes are either prepared by online communities of O&M professionals or generated based on similar existing ones [40]. In the latter case, tactile description, or tactile maps, give readers the layout of a venue and the spatial relationship between its landmarks, enhancing one's ability to self-orient and facilitating a safer and more confident travel.

Aside from knowing how to get to places, knowledge of one's surrounding is critical for environmental awareness that helps with correct and safe decision making. Sighted people rely heavily on visual landmarks to orient themselves and to avoid obstacles as they travel. Unfortunately, access to such visual cues can be vexing for blind individuals, forcing them to navigate using non-visual sensing that is cognitively taxing and inaccurate. Current maps and floor plans accessible to visually impaired people generally do not contain spatial representation at the level of detail that is useful for blind navigators.

1.2 Semantic Interior Mapology

In this dissertation, I introduce an end-to-end systematic tool for a collective spatial mapping of an indoor environment consisting of small-scale features (i.e., building fixtures, room furnitures, etc), and for the production of tactile maps at different levels of details, serving the purpose of promoting pre-journey spatial awareness. My proposed system is a composition of three separate toolboxes, namely *Map Conversion Tool*, *Map Population Tool*, and *Map Authoring Tool*. These toolboxes can be operated independently at different stages, and beneficially, their outcomes can be joined together in representing the semantic spatial layout of an environment in different modalities (figure 1.2).

The Map Conversion Tool allows one to easily trace an architectural floor plan to produce the semantic information at the building level. Such information is then converted into a vectorized map, hierarchically organized in terms of spaces. Each space is characterized by a set of wall corners and possibly entrance corners, where pairs of adjacent wall corners may or may not be joined by a wall. This vectorized map is stored in a file format, named *sim*, that is inspired by the Polygon File Format (PLY). This toolbox lets users define a grid of horizontal and vertical lines, where each line overlaps with a segment in the floor plan representing a wall. A single line may contain multiple disjoint wall segments, and by generating a line, it reduces the number of required input selections for wall segments that are co-planar. This strategy exploits the fact that most floor plans have walls intersecting at 90 degrees. The Map Conversion Tool is a web application with an intuitive interface designed to quickly and accurately convert a floor plan into a digital format amenable to interactive visualization.

The Map Population Tool is a semi-automatic procedure that segments 3D scans of indoor spaces inside a building and annotates individual objects of interest, resulting a semantic representation of an interior at the room level. The 3D-scanned scenes are obtained using off-the-shelf Occipital's Structure Sensor[1], which has a RGB-D camera and software for registration and stitching of multiple 3D point clouds into a mesh. The segmentation process of a scanned scene is organized into a sequence of stages: scene collection, orientation, rectification, super-pixelation, and segmentation. Results of this is then fed into the annotation process that groups all super-pixels with similar orientation into a corresponding object. Such spatial data is then geo-registered with the *sim* format of the same environment, generated by the Map Conversion Tool,

[1]Structure Sensor. https://structure.io/

and used to populate the interior by producing the JSON file of the desired items. The semantic representation yielded from these two toolboxes can be served as advantageous complements to pre-journey learning. For instance, they can be converted into the GeoJSON format for populating an accurate 3D model of a floor plan onto OpenStreetMap[2]. By using assistive technology such as a screen reader or a screen magnifier, visually-impaired individuals can learn the spatial layout of a building, as well as its interior contents.

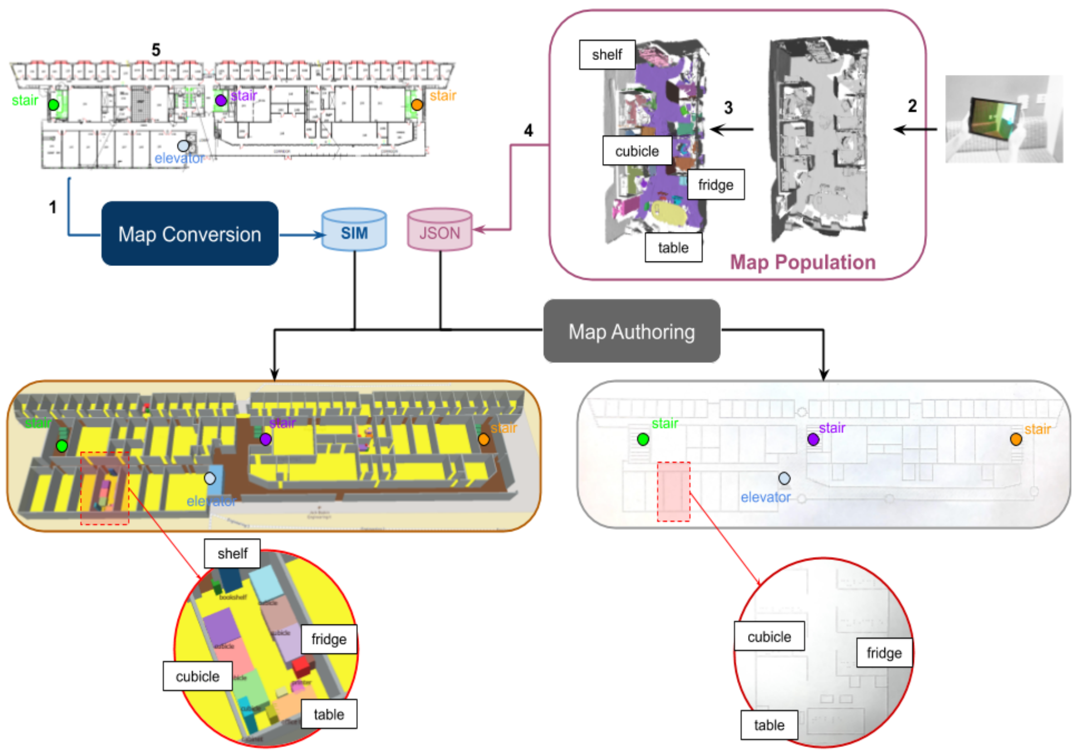


Figure 1.2: The components of Semantic Interior Mapology: (1) Map Conversion Tool for tracing a floor plan and encoding its spatial features representation into the *sim* format; Map Population Tool for (2) importing a 3D mesh of a space, (3) segmenting elements of interest, and (4) embedding small-scaled objects of interest into a JSON map; (5) Map Authoring Tool rendering spatial information using visual and non-visual modalities: a pop-up 3D model on OpenStreetMap and tactile maps at different scales.

To serve the purpose of pre-journey learning, my system includes the Map Authoring Tool that automatically generates tactile maps at different spatial scales,

[2]OpenStreetMap. https://www.openstreetmap.org

representing a building's structure, a zoomed-in of a specific area, or an interior of a room. This toolbox produces digital tactile map files at desired scales, based on the building's structure encoded in sim and the semantic layout of an interior derived from its 3D scan. Additionally, it also incorporates a simple editor, only available at the room level, that allows users to edit or delete annotated objects of interest, thus, facilitating a clean-cut map production while maintaining essential spatial information. These tactile map files are then available to download and can be printed on a 292 mm × 279 mm embosser sheet, with resolution of 20 DPI and a 12 mm margin on all sides. My end-to-end systematic tool for encoding and representing the spatial structure of an indoor environment is available to use at `https://sim.soe.ucsc.edu`.

1.3 What Should Map Authoring Tool Produce and Is Sim Really Useful for Pre-journey Learning Process?

Although my systematic application is able to help encoding a building's spatial layout at the desired resolution, decision on what level of detail should be contained at a given scale is not a trivial task. This *generalization* problem is common to all types of map design [55, 63], especially in the case of tactile maps. In fact, tactile sensing affords relatively low spatial resolution, which reduces the density of details reproducible on a standard embosser sheet (292 mm × 279 mm). For example, the minimum center-to-center spacing between two Braille dots must be 2.28 mm [38], or any embossed symbol must have the minimum diameter of 6 mm [78]. Which details can be part of the map and which can be removed are usually decided by following guidelines developed in [18, 3], or just simply weighing several tradeoffs. During my study, I and my advisor, Dr. Roberto Manduchi, conducted a focus group to gain some insights into this challenge. The goal of the focus group is to understand: *Which type of features should be represented on tactile maps at a given scale? What is the appropriate set of distinctly discriminable tactile graphics to be used for representing these features? And what is the optimal selection of map scales that can facilitate pre-journey spatial awareness?*

Findings from the focus group indicate that embossing maps of a building at different levels of detail could provide useful spatial information for the pre-journey learning. Specifically, maps at the structure and section scales give readers a general picture of the building layout, and contain the wayfinding information to reach specific destinations. A participant mentioned a situation of taking an elevator and expressed her confusion of where to turn after getting off it. Another participant considered a situation of being dropped off in front of a specific entrance at a bus station, and with the information provided on the tactile map, he can quickly navigate himself to the front desk. Whereas, some think that map at the room scale is secondary and less useful because the position or the presence of interior furnitures is an unknown variable. One mutual agreement emerged from the focus group discussion is that map needs to be universal, in which tactile symbols must be standard, understandable and discriminable. The Map Authoring Tool takes in such insights and produces a tactile representation, consisting of architectural features (e.g, wall, entrance, staircase, elevator, water fountain) and essential interior features (e.g, door, table, shelf, etc), that serves pre-journey spatial knowledge acquisition (figure 1.3).

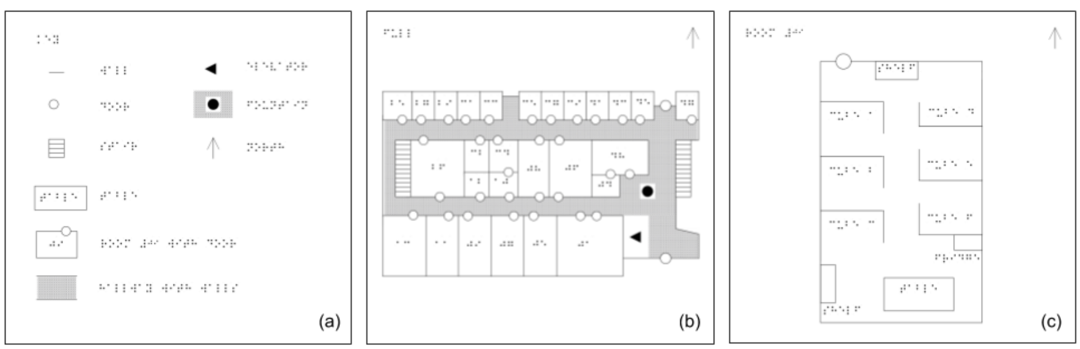


Figure 1.3: Tactile maps generated by Map Authoring Tool include both architectural features and interior features: (a) the map key, (b) the section-scale map represents the architectural layout of a specific zoom-in of a building with features of walls, entrances, doors, staircases, elevators, water fountain, rooms and corridors, (c) the room-scale map represent the interior layout of a room with features of cubicle, shelf, and table

As the natural sequel, I and Dr. Roberto Manduchi conducted a user study to experiment with the use of multi-scale tactile maps that are automatically generated from the Map Authoring Tool. Our study evaluates whether access to spatial information represented in these maps could lead to increased spatial awareness, well-maintained self-orientation, and efficient navigation from one place to another. Given a set of tactile maps embossed at the section and room levels, the participants were asked to describe the graphical symbols, to locate furnitures (i.e, cubicles, tables, shelves) appearing within a room, and to explore the general layout of an area within a building. They were also tasked to imaginatively orient their bodies in a certain direction, and then point to various location inside the building. Once the participants were familiar with the maps, Dr. Manduchi asked them to construct and describe a path, starting from the building's entrance, to a very specific position inside a room. This path must be expressed in terms of turns, intersections, doors passed to the left or to the right, and any landmark along the way (e.i., staircase, water fountain, elevetor). As a result, the majority of participants found no trouble in completing the required tasks. In fact, they give a high compliment on the provided tactile representation: *"the maps can definitely be used and helpful"*, *"the symbols are easy to identify"*, and *"looking at the maps give me the general layout of the place, and how one room connects to another"*. It is proven that my end-to-end systematic tool is able to collectively encode semantic information of an indoor environment in great details, and to produce cutting-edge embossed tactile maps at different resolution, thus, promoting spatial awareness for visually impaired travelers prior to a journey.

Chapter 2

Vectorizing an Architectural Floor Plan

Interactive 3D visualization of building interiors provides enhanced experience of spatial exploration with respect to the traditional static maps. Using mapping platforms such as Mapbox[1], 3D pop-up environments can be easily rendered on top of generic 2D maps from web applications such as Google Map or OpenStreetMap. This type of 3D rendering may afford more intuitive and engaging access to complex building layouts, and may enable interactive features such as displaying a selected floor of a building, or activating groups of features on different levels of detail.

In order to display building interiors by means of 3D interactive tools, it is first necessary to convert available spatial data into an appropriate vectorized format. While most modern building have detailed CAD floor plans (e.g. in dwg or dwf format), this data is normally not accessible to the mapper. Floor plans, when they are available, are only accessible in an image (e.g., JPEG) or PDF format. Computer vision algorithms for the automatic conversion, from the raster to the vectorized form of floor plan images, have been demonstrated, but these algorithms are not universally applicable because of the wide variety of graphical representations used to draw the floor plans. And while companies such Google and Apple are actively acquiring digital representation of interiors of public spaces, and some of these are already available for visualization in their map applications. Such data is proprietary and not available to the public. In this chapter, I introduce the first component of my end-to-end application, the *Map Conversion Tool*, which is designed to easily convert floor plans into a digital format

[1]Mapbox. https://www.mapbox.com/

amenable to interactive visualization.

2.1 Floor Plan Analysis

Early works in floor plan analysis focused on the interactive conversion of a 2D image into a 3D model [16] [45]. The *ScanPlan* project [54] used the Hough transform for the detection of walls and doors, under the assumption of convex room shape. The algorithm of Ahmed *et al.* [1] detected and labeled rooms based on geometric reasoning involving analysis of the line thickness in a high-resolution image of a small floor plan, typically containing 4 to 5 rooms. De las Heras *et al.* [33] proposed to detect walls based on specific assumptions (i.e. walls are drawn as parallel lines in repetitive patterns that are well distributed in the floor plan). Similarly, the algorithm of [27] recognized walls by determining parallel lines separated by a defined distance. Jang *et al.* [37] used a neural network (U-net) to pre-process the floor plan image and extract a skeleton of walls.

Other recent work employs graph-based algorithms to detect a room's boundary [69] [56], trains a neural network for the task of pixel-wise wall segmentation [14] [49] [67], or reasons about a floor plan's structure from a mobile device's inertial data through crowdsourcing [90] [39]. Specifically, [25] [26] leveraged crowd-sensed data from mobile users to obtain the spatial relationship between adjacent objects to complete a floor plan reconstruction. Liu *et al.* [48] proposed a neural network, called *FloorNet*, that reasons local spatial information based on the point density captured from smartphones.

Unfortunately, automatic methods for the extraction of room layouts often fail in the case of complex floor plan images. For example, the state-of-the-art algorithm of Liu *et al.* [49] only reaches an accuracy of 85% for room segmentation. In addition, these algorithms usually break down in the case of large and complex layouts such as those considered in this paper (e.g., office buildings), and are unable to correctly detect diagonal walls or nested rooms.

2.2 Map Conversion Tool

Map Conversion Tool allows one to quickly and accurately trace a floor plan from an architectural image of it, produces a vectorized map encoding the building's structural layout, and then geo-registers this spatial information as a 3D pop-up on top of the 2D web application. This toolbox's pipeline contains three sub-components: Floor Plan Tracer, GeoJSON Generator, and MapboxGL JS (figure 3.2). Once a tracing is completed, the Floor Plan Tracer produces a spatial semantic map of a building, organized in terms of spaces, wall segments, and doors. This information is stored in *sim*, a file format inspired by the Polygon File Format. The GeoJSON Generator converts *sim* into GeoJSON which is a popular format for representing spatial information on a web platform. To render a floor plan's 3D map view, I employ the MapboxGL JS[2] engine, a location data platform. Geodetic features stored in a GeoJSON file are shown as extruded 3D objects on OpenStreetMap, which can be accessed and interacted with from a regular web browser.

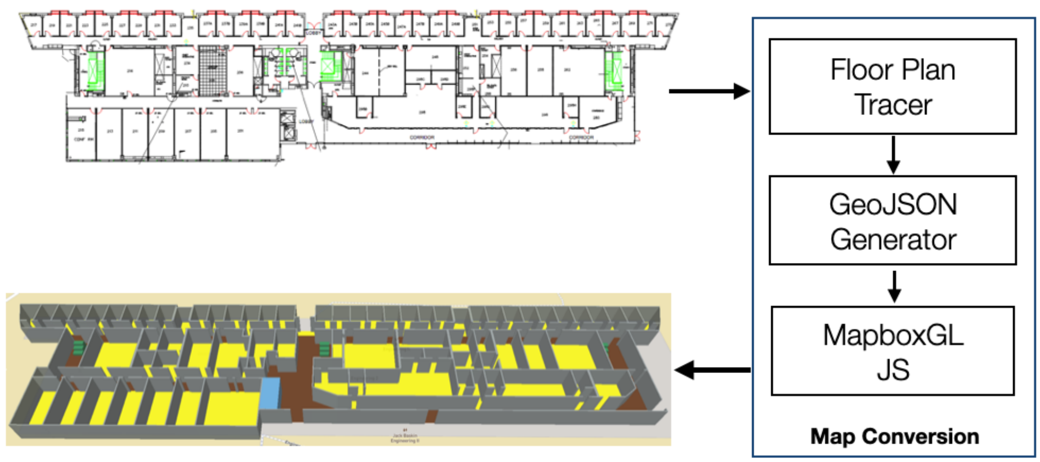


Figure 2.1: Map Conversion Tool's pipeline.

[2]MapboxGL JS. https://docs.mapbox.com/mapbox-gl-js/

2.2.1 Floor Plan Tracing

Floor Plan Tracing is a web application, consisting of the control panel on the left and the drawing canvas on the right. The tracing interface exploits the fact that most floor plans have straight walls that intersect at 90 degrees, meaning that most walls can only have one of two orientations. Note that the toolbox also supports less common situations with walls at arbitrary orientation. The floor plan displayed on the drawing canvas should be oriented such that the main wall orientations are parallel to the screen axes (figure 2.2).

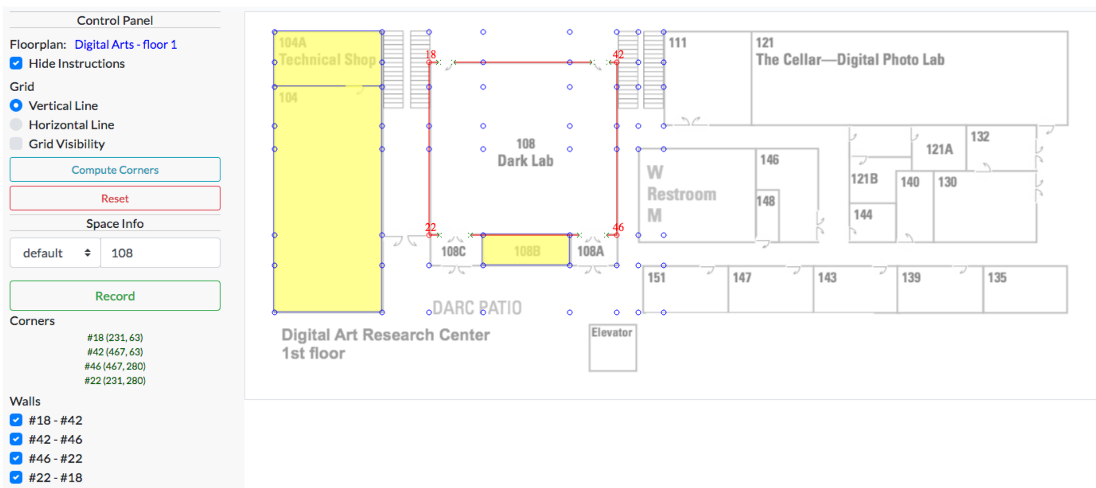


Figure 2.2: The tool GUI includes a control panel and a canvas. Possible wall corners (line intersections) are rendered as blue circles. Actual wall corners (line intersections) are rendered as red circles. Red segments represent wall segments. Gaps with endpoints marked by a green x denote entrances within wall segments. Yellow polygons with blue border show spaces for which a tracing was completed. Gaps in a blue border represent entrances to a space.

Rather than tracing wall segments by selecting endpoints (as with other web applications such as Google My Map or Mapbox Studio[3]), the Map Conversion Tool lets the user define a grid of horizontal and vertical lines, where each line overlaps with a segment in the floor plan representing a wall. The user simply *Shift*-clicks on a segment to generate a line with the desired orientation. Note that, in typical layouts, the same line may contain multiple disjoint wall segments that happen to be co-planar. This

[3]Mapbox Studio. https://www.mapbox.com/mapbox-studio

strategy is very convenient in the case of repetitive layouts, as it reduces the number of required input selections, and ensures that co-planar walls are traced by segments that are correctly aligned with each other (figure 2.3 (a)(b)). In the case of diagonal (but still planar) wall segments, the user should add two properly oriented ”ghost wall” lines, crossing an actual wall segment at the desired corner (i.e., at an endpoint of the diagonal wall segment). Non-planar walls are not currently supported by the toolkit. Lines can be added with simple *Shift*-clicks, and removed (in the case of a mistake) with *Alt/Cmd*-clicks. Once all visible wall segments have been covered by lines, the user may click on the *Compute Corners* button. This triggers computation and display, in the form of small blue circles, of all line intersections (the lines are automatically hidden from the display for an ease of view, as shown in the figure 2.3 (c)). Each intersection is assigned an unique numerical ID. Some, but not all, of these intersections correspond to physical wall corners.

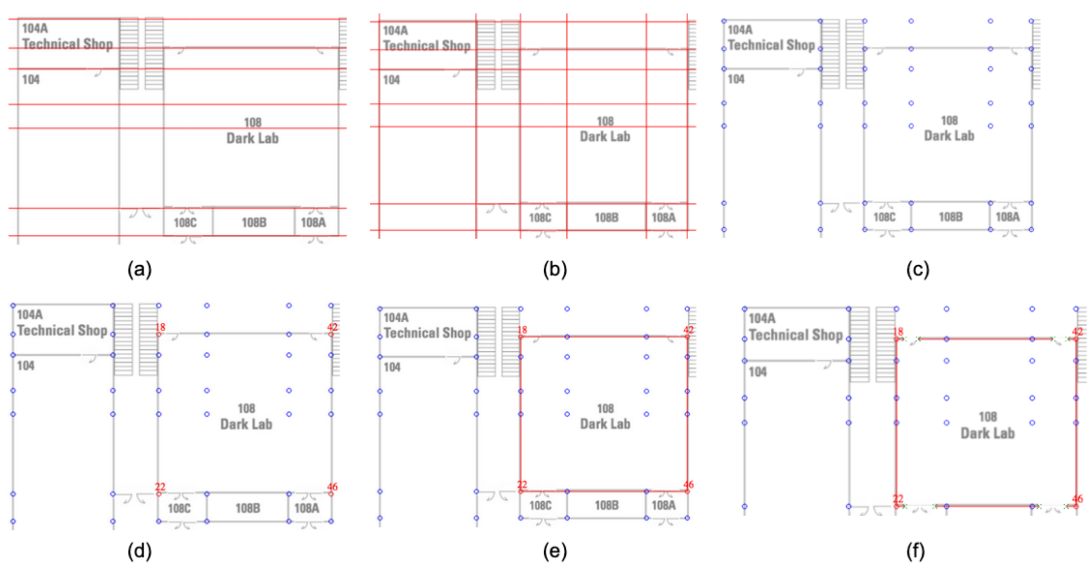


Figure 2.3: A typical Map Conversion workflow. (a) First, horizontal grid lines are generated via Shift-clicks. (b) Vertical grid lines are generated next. (c) All line intersections (possible wall corners) are automatically computed and displayed. (d) The user selects corners #18, #22, #42, #46 for the boundary of a space (Room 108). (e) The user then selects the walls connecting the corner pairs (#18, #42), (#42, #46), (#46, #22), and (#22, #18). (f) Finally, the user defines two entrances along the wall of (#18, #42) and two entrances along the wall of (#46, #22).

The next step is for the user to (1) select which ones of the intersections do correspond to wall corners, (2) select whether two nearby corners are joined by a wall, and (3) associate corners and walls to individual spaces (rooms or open areas such as corridors or halls). Note that complex spaces can be conveniently subdivided into smaller spaces as shown in the figure 2.4 (a). An example with a diagonal wall is shown in the figure 2.4 (b). Note that corners #32 and #53 are generated as the intersections of proper wall lines (i.e., lines containing actual wall segments) with "ghost wall" lines, as explained above. For example, a horizontal ghost wall line (created by a Shift-click on the map) intersects the vertical wall line at #32.

The corner selection and wall association step is accomplished as follows. Each space is visited in turn. At each space, the user clicks on the line intersections (the small blue circles) that correspond to physical wall corners within that space. The color of the selected wall corners turns red, and their associated IDs are displayed on the map (figure 2.3 (d)). These wall corners are sorted in the clockwise order, and listed in the control panel of the interface. In addition, all possible walls joining adjacent corners are also listed in the same panel. For example, in the figure 2.3 (e), after the user selects corners #18, #22, #42, #46, they are ordered as (#18, #42, #46, #22), and all possible walls connecting adjacent corner pairs are displayed. These are: (#18, #42), (#42, #46), (#46, #22), and (#22, #18) (not shown in the figure due to space limitation). The user then simply clicks on the corner pairs that correspond to actual walls, which are then displayed as red segments. In this case, the room has a closed contour (except for door openings), hence all corner pairs are selected.

As another example, consider the open space ('LOBBY') shown in the figure 2.4 (c). Its fairly complex layout is divided into a number of smaller spaces, one of which is defined by the wall corners (#33, #47, #46, #53, #56, #35). Only the following corner pairs are joined by a wall: (#33, #47), (#47,#46), (#53,#56). Note that the remaining corner pairs ((#46, #53), (#56, #35), (#35, #33)) are not selected, signifying that the spaces between them are open.

In order to trace an entrance door of a space, the user first defines the whole wall containing the door as described above (as opposed, for example, to defining two wall segments at either side of the door). Once the wall segment has been determined,

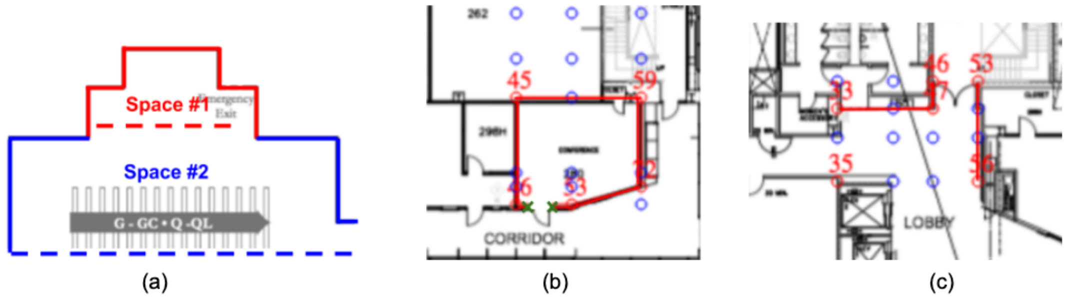


Figure 2.4: (a) A complex space can be divided into multiple spaces to facilitate tracing. (b) An example of a traced room with a diagonal wall. (c) Tracing a "segment" of an open space. Note that several wall corners are not linked by walls.

one can define the endpoints of the door segment, by *Shift*-clicking on the appropriate locations on the wall segment. For example, in the figure 2.3 (f), the user specifies two entrances along the wall connecting the corner-pair (#18, #42), as well as two entrances along the wall of (#46, #22). Two endpoints are automatically stored in a list (separate from the wall corners list), and a new entity ("entrance") is defined, joining the two endpoints. The user concludes the task of tracing a space by providing its name (e.g., a room number), and by selecting its space type from a pull-down menu. The currently supported types include: room (default), corridor, restroom, staircase, elevator.

2.2.1.1 Spatial Features Representation

The floor plan tracing process described above produces a spatial information hierarchically organized in terms of *spaces*. Each space is characterized by a set of *wall corners* and possibly *entrance corners*, where pairs of adjacent wall corners may or may not be joined by a wall. I store this information in a *sim* file. My sim format is inspired by the Polygon File Format (PLY), which is used to represent 3D objects as lists of flat polygons. A PLY file contains a list of vertices and a list of polygons, where each polygon is defined as an ordered list of vertex IDs. A sim file contains a list of wall corners and a list of entrance corners. Each space is assigned a list of wall corner IDs and a (possibly empty) list of entrance corner IDs. Additionally, sim allows one to specify whether two wall corners in the list should be connected by a wall, or not (implying an empty space between these corners).

A space is represented in the following format:

`{id type name num_corner wall_corner_indices walls entrances}`.

For example, `{s2 0 217 5 1 3 27 19 12 1 1 1 1 0 e2 3 1 2}` means that the space's ID is s2, its space type is 0 (meaning a room by default), its name is 217, and it has 5 corners whose indices, sorted in clockwise order, are (#1, #3, #27, #19, #12). The next sequence of binary values (`1 1 1 1 0`) indicates that there are walls connecting the corner-pairs (#1, #3), (#3, #27), (#27, #19), and (#19, #12); but there is no connection for (#12, #1). The last sequence with four entries (`e2 3 1 2`) denotes that there is an entrance with an identifier of e2 along the wall with index 3 (i.e., the third wall in the list: (#27, #19)). The endpoints of this entrance are (#1, #2), where these IDs refer to the list of entrance corners. Additional entrances to the same space can be listed as additional quadruplets of entries at the end of the list. Note that wall corners and wall segments can be re-used for different adjacent spaces.

2.2.2 GeoJSON Generator

A sim file can be easily converted into other formats. The Map Conversion Tool contains a converter into GeoJSON, a popular format for representing spatial information [8]. The GeoJSON *features*, representing segmented spaces, consist of a set of *properties* represented as a (key, value) mapping and a geographical *geometry* represented as a polygon (one of the geometric primitives in the GeoJSON format). A feature's properties include name, encoded color, height, and distance from the ground level. The geometry information contains the coordinates (lat, long) of the polygon's vertices. Multiple features are hierarchically grouped into a *Feature Collection* object. The GeoJSON generator in the Map Conversion Tool generates the *Feature Collection* object automatically from the *sim* structure. The figure 2.5 shows an example of conversion from sim spaces to GeoJSON Feature Collection. The Mapbox GL JS engine renders a 3D map view by extruding each feature in such collection.

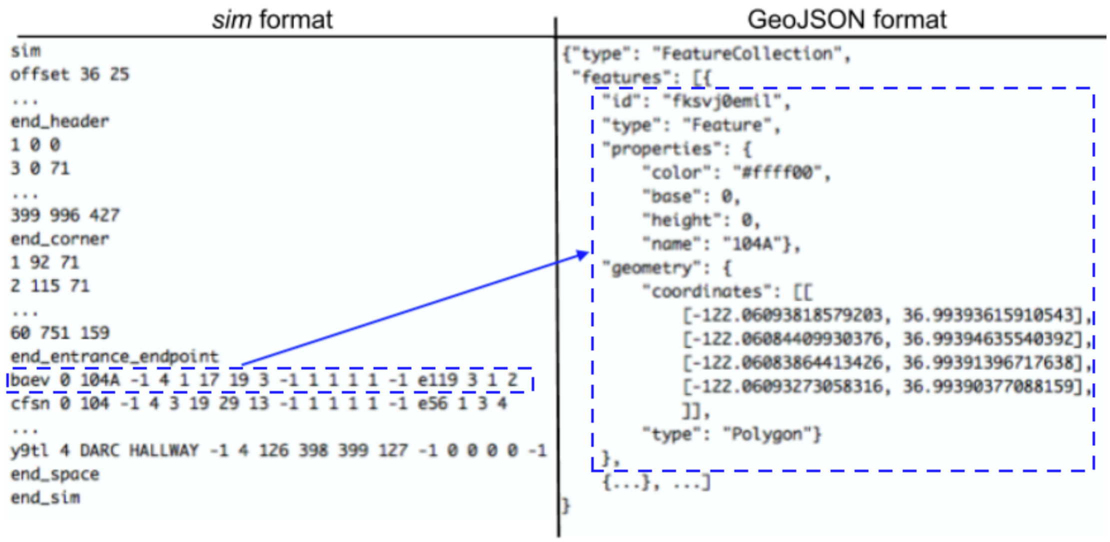


Figure 2.5: An example of conversion from sim to GeoJSON. A room, described by one row in the sim file, is represented as a feature in GeoJSON.

2.2.2.1 Geo-registration

Corners in sim are defined in terms of (x, y) screen coordinates. The conversion to (lat, long) geodetic coordinates for GeoJSON representation is performed as follows. First, I determine the geodetic coordinates of at least four corners, chosen from the building's external walls. This is easy to do if, for example, the contour of the building under consideration is visible in a web application such as Google Map or Apple Map, and the location of the selected corners can be identified in this contour (note that these applications return the WGS84 geodetic coordinates of selected locations). The geodetic coordinates of these points are then converted to Universal Transverse Mercator (UTM) coordinates using standard formulas. The UTM system is based on a conformal projection, and thus produces little distortion for small areas. Next, I determine a collineation (homographic) transformation between the (x, y) screen coordinates of the wall corners and the UTM coordinates of the same points. The collineation matrix can be found using Direct Linear Transformation [80]. The same collineation is then used to transform the (x, y) coordinates of all remaining corners into UTM coordinates, which are then converted into geodetic coordinates (figure 2.6).

Some examples of an end-to-end conversion from floor plan images to GeoJ-

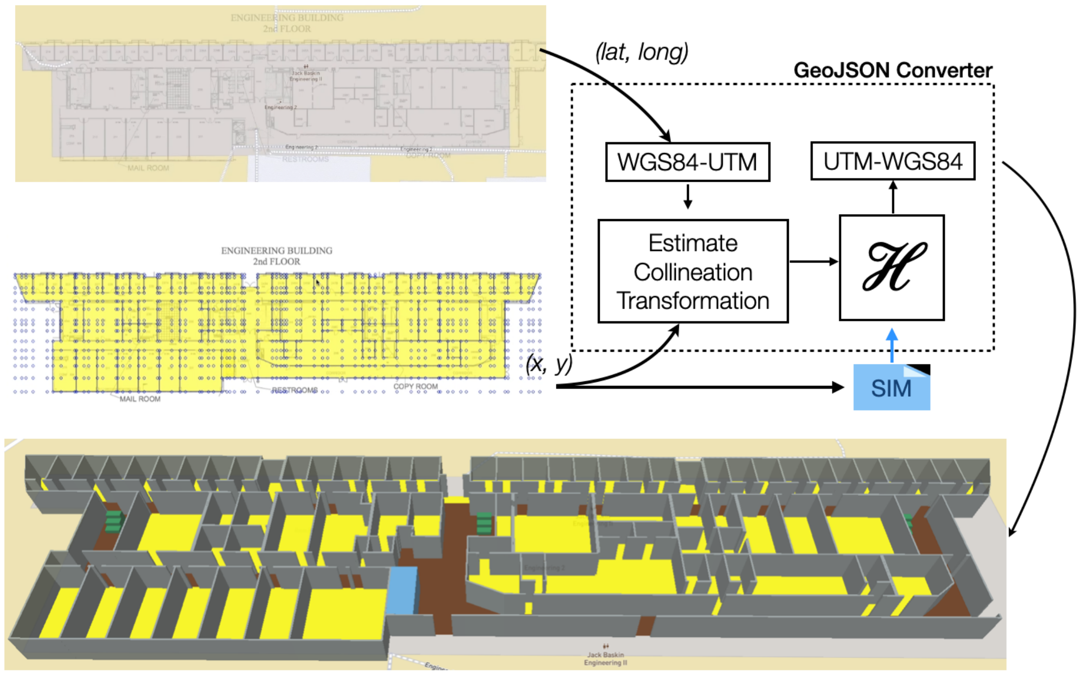


Figure 2.6: Geo-register corners of a building encoded in sim with their corresponding geodetic coordinates. Notes: WGS84 stands for World Geodetic System 1984, and UTM stands for Universal Traverse Mercator.

SON files, shown as pop-ups over OpenStreetMap using the MapboxGL JS engine, are presented in the figure 2.7. Note that in the conversion from sim to GeoJSON, staircases have been represented as three adjacent rectangles of different heights, colored in green, while elevators are shown colored in blue. Unlike Map Conversion Tool, other drawing interfaces such as Mapbox Studio or Google My Map doesn't allow one to trace a floor plan image. Individual walls need to be copied by hand, often resulting in geometric errors such as incorrect spacing or orientation. In addition, when shapes are drawn manually by hand, connectivity errors may occur. My strategy of first defining a line grid, and then selecting corners from the line intersections, ensures that co-planar walls are represented by collinear segments, and that connected wall corners remain connected.

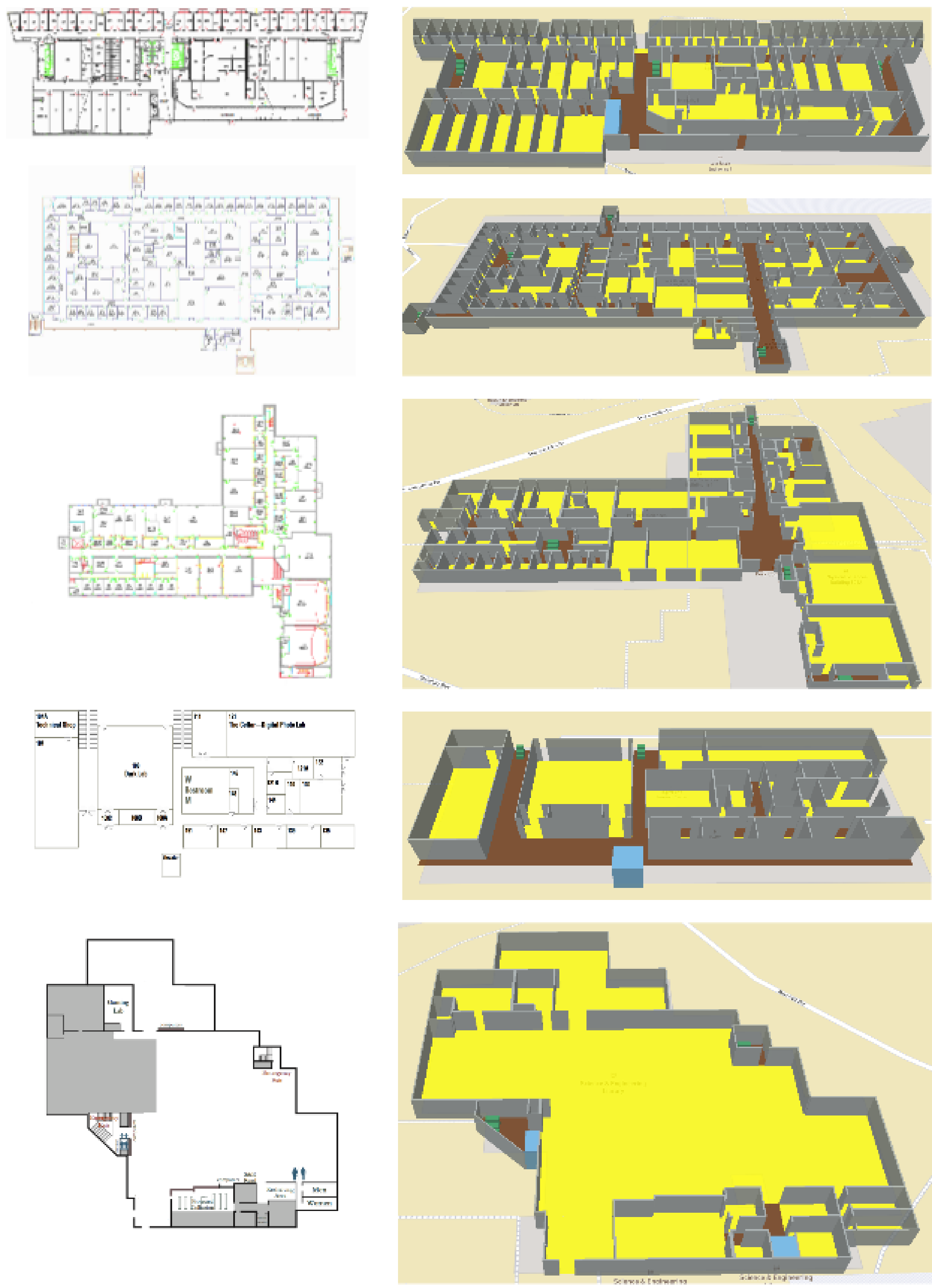


Figure 2.7: Examples of application of the Map Conversion Tool.

Chapter 3

From 3D Scans of an Interior to a Fine-grained Spatial Representation of a Building

Consider an example of Sebas, a blind front-end web developer has to attend a user interface meeting scheduled in a conference room that he has never visited before. He decides to leave his cubicle early enough to give himself a plenty of time, in case he gets lost along the way. Following verbal directions that he received from a colleague, Sebas is able to reach the room. Once he arrives, he finds himself at a loss. What is the layout of the furniture inside this room? Is there a large central table which chairs are around or are there rows of tables and chairs as in a classroom setting? On which wall is the project screen located? Maps of a building, regardless of the modality and format, typically contain information only at the level of walls and openings, such as doors. Rarely do they represent smaller-scale features such as fixtures or furnitures. Yet, when available, these features could make for a richer visualization, and could convey useful spatial information.

In the previous chapter, I described the Map Conversion Tool that allows tracing of a floor plan to produce the spatial information at the building level, consisting of walls, doors, staircases, elevators, offices and corridors. In this chapter, I introduce the second part of my end-to-end systematic application, Map Population Tool, that helps encoding spatial relationships between objects appearing within a space inside a building. Map Population Tool is a semi-automatic procedure that parses the 3D scan of a room, segments out elements of interest, and then geo-registers these elements within the building's spatial layout. The 3D-scanned scenes are obtained using off-the-shelf Occipital's Structure Sensor[1] (figure 3.1) which is an RGB-D camera attached to an iPad Air, scanning an environment into a 3D mesh. Such mesh are passed through a semantic segmentation and annotation process in which objects of interest are then registered with the vectorized *sim* map of the building, resulting in the spatial information at the room level. This information can later be represented in multiple modalities at an appropriate spatial scale.

Figure 3.1: Occipital Structure Sensor is an iPad-powered RGB-D camera which scan and import 3D images of rooms, objects, and people.

3.1 Semantic Segmentation

In the field of computer vision, convolutional neural networks (CNNs, or ConvNets) [41] are widely employed to analyze visual imagery, by learning an image's features. These features, also known as image descriptors, are learned directly and automatically in a network's layers. This greatly outperform any other manually selected ones for a variety of tasks, such as object classification [31], object detection [28], and

[1]Structure Sensor. https://structure.io/

object recognization [15]. Advancement in the problem of semantic segmentation has adapted CNNs, which originally trained to classify objects, to perform a classification at the pixel level. Some notable examples are Fully Convolutional Network (FCN) [51], U-Net [74], SegNet [5], and DeepLab [10].

Unlike the standard neural networks designed to analyze RGB inputs, 3D convolutional neural networks aim to segment and detect objects from RGB-D images or 3D point clouds. Several approaches have been proposed, ranging from voxel-based representation [86] [59] to feature 2D pooling from multiple viewpoints [79] [72]. Early works on semantic segmentation of RGB-D images relied on interactive user input (usually a stroke) to perform a segmentation [84] [65] . McCormac *et al.* [60] proposed to transfer semantic segmentation from 2D predictions to the 3D domain. Such method is able to produce a high-resolution segmentation, but none of the predictions happens directly on 3D inputs. Later approaches applied CNNs to a 3D volumetric representation to classify each voxel in the scene [77]. For instance, *PointNet*, designed by Qi *et al.* [71], is the pioneer in handling 3D point cloud data. It consumes point clouds directly, without voxelization or rendering, while preserving the permutation invariance of points in the input. *PointCNN* [46] challenged *PointNet*'s performance by weighting the input features associated with each point. In response, Qi *et al.* [73] suggested to build a graph neural network for semantic segmentation on the point cloud, where each node is a group of points and graph edges are constructed by nearest neighbor search on the point cloud.

Recent trend in semantic segmentation is to design an end-to-end neural network that optimizes the network weights, also known as the learning ability, by considering both inputs and outputs of it. A Point Global Context Reasoning proposed by Ma *et. al.* describes the contextual dependencies among 3D point cloud using a graph representation and a self-attention network model [53]. This is a plug-and-play model that can be easily integrated into any point cloud segmentation architeture. Inspired by the ability of remembering values over a time interval in the work of Long Short-term Memory by Hochreiter and Schmidhuber [34], Du and his team proposed a long-short-term context framework that exploits both the local features within a block of 3D point clouds and the "long-range" features residing in other blocks of the same point clouds to improve the segmentation performance [17]. In another approach, Zhao introduced an end-to-end deep neural network, built on top of PointNet [73], that improves the ability of extracting local feature by enlarging the receptive field of convolutional kernel [88]. This network increases the precision of point cloud segmentation over 85%. However, such automatic methods for the extraction of room layouts often fail in the case of complex floor plan images, due to incorrectly detect diagonal walls or nested rooms.

3.2 Map Population Tool

The Map Population Tool allows one to add small-scale items that are not present in the original floor plan. It starts from a 3D scan of an environment, segments out 3D objects of interest, and inserts these objects as cuboids into a semantic representation of a floor plan. The workflow is organized in a sequence of stages: Scene Preprocessing, Segmentation and Annotation, and GeoJSON Generator (figure 3.2). After a scan is completed, the Structure sensor stitches multiple 3D-scanned point clouds into one mesh, stored in the PLY format. In this work, I assume that an entirely scanned environment is a room with four walls; and the same mechanism can be extended to the case of partial scans, or for different types of spaces (e.g., corridors) (figure 3.3). In Scene Preprocessing, mesh facets with similar vertex normals in the scan are merged into super-pixels. The segmentation and annotation stage let users to group all super-pixels with similar orientation into an identified object. These objects is then stored in the JSON format that later are geo-registered with the GeoJSON representation of the same environment generated by the Map Conversion Tool, populating the interior or the desired items in the same GeoJSON file.

3.2.1 3D Scene Preprocessing

My first step is to orient the mesh acquired by the 3D scanner with the floor plan. The Structure sensor produces a mesh with its Y-axis vertical (as measured by the sensor's accelerometer), but with an arbitrary orientation of the X-Z plane. I would like to re-orient the mesh (rotate it around the Y-axis) such that the walls of the room are aligned with the X and Z axes. These axes will then be mapped to the axes of the 2D floor plan. I first select all vertices in the mesh whose normal vector is approximately

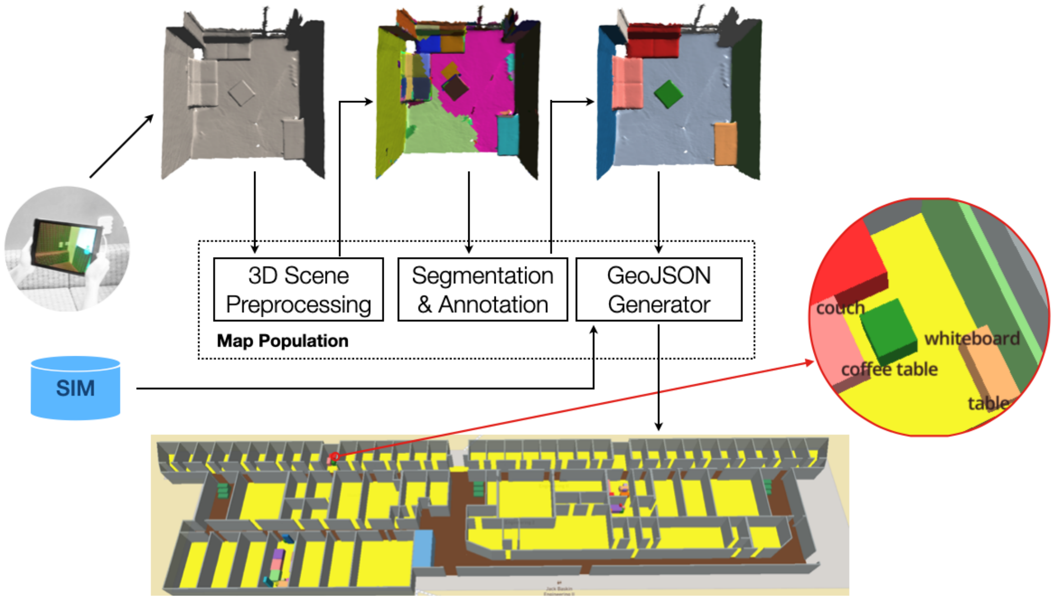


Figure 3.2: Map Population Tool's pipeline.

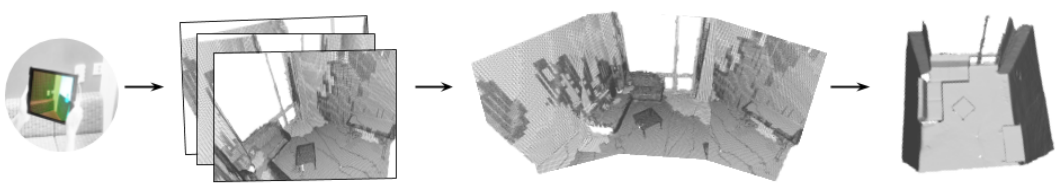


Figure 3.3: 3D Scan Process. The Occipital Structure sensor scans a room into a set of point clouds, registers and stitches these 3D point clouds into one mesh.

orthogonal to the Y-axis (i.e., corresponding to vertical surface elements). For each such vertex, I compute the angle formed by its normal and the Z-axis. Peaks in the histogram of these angles reveal the orientation of the main walls. For example, in the figure. 3.4, a peak is found at 127°, corresponding to the orientation of the longer walls. The whole mesh is then re-oriented by rotation around the Y-axis by the opposite angle.

Due to errors in data acquisition, registration or stitching, the geometry of 3D scans of environments is often inaccurate. In particular, wall scans are sometimes not planar, or walls appear not to intersect at 90°. This may affect the registration

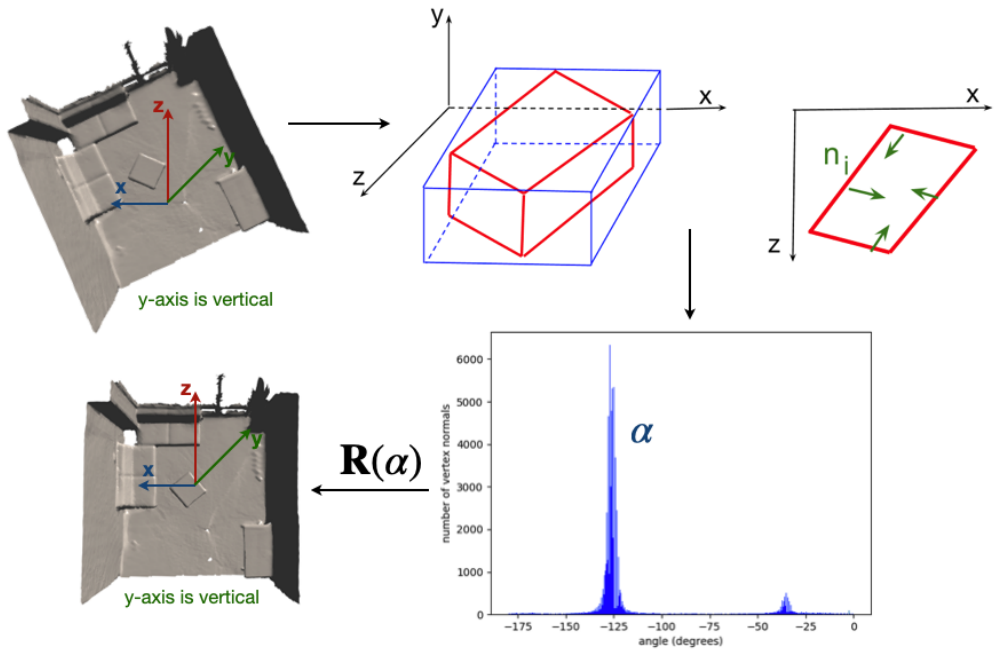


Figure 3.4: A simple example of re-orientation. The histogram of the angles between surface normals and Z-axis is calculated. After being re-oriented by an opposite of the peak angle, the longest segment is aligned with the Z-axis.

of the environment with the floor plan. I correct for global errors using the following simple procedure. First, I identify the four walls in the acquired mesh by projecting the vertices of the re-oriented mesh onto the X-Z plane (figure 3.5 (e)). I then select the vertices with Z-coordinate in the top quartile, and run the RANSAC algorithm [23] to find a robust line fitting (figure 3.5 (a)). This line represents the top wall. Repeating the same procedure for all sides (figure 3.5 (b-d)) results in four lines in the X-Z plane. An example of the result is shown in the figure 3.5 (e). From this figure, it is clear that, due to artifacts of scanning, the walls do not appear to intersect at 90°. I then find the axis-parallel rectangular box that best approximates the quadrilateral formed by the line intersections (figure 3.5 (f)). The collineation (homography) that brings this quadrilateral's vertices into the corners of this axis-parallel rectangle is computed. The mesh can then be rectified by applying the same collineation to the (X, Z) coordinates of all vertices in the mesh.

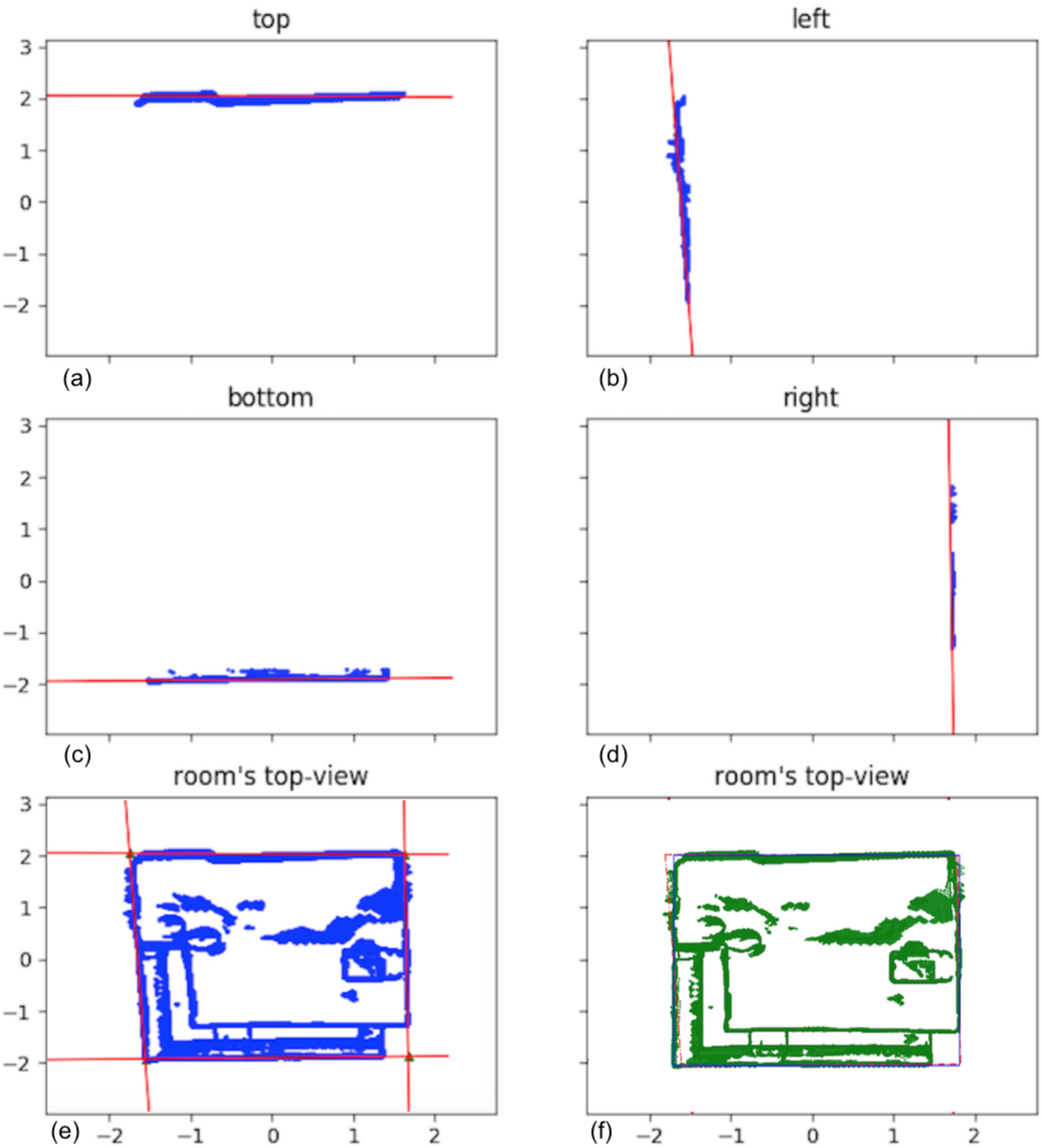


Figure 3.5: Individual walls are identified by linear fitting of the vertices of the re-oriented mesh, projected onto the X-Z plane (a - d). The resulting quadrilateral (e) is transformed into the best-fitting axis-parallel rectangle (f), and the same transformation is applied to the (X, Z) coordinates of all vertices in the mesh.

In order to register the resulting mesh with the floor plan, I first need to visually determine the correct orientation of the mesh. As mentioned previously, my re-orientation procedure aligns the longer walls with the Z-axis. However, this may not be the actual orientation of the space in the floor plan, and an additional rotation by ±90° or 180° may be required. Finally, I find the offset between the (x, y) coordinates of one corner of the room in the floor plan, and the (X, Z) coordinates of the corresponding corner of the rectangular bounding box, as well as the two scale factors that ensure that the mesh correctly fits the room in the floor plan.

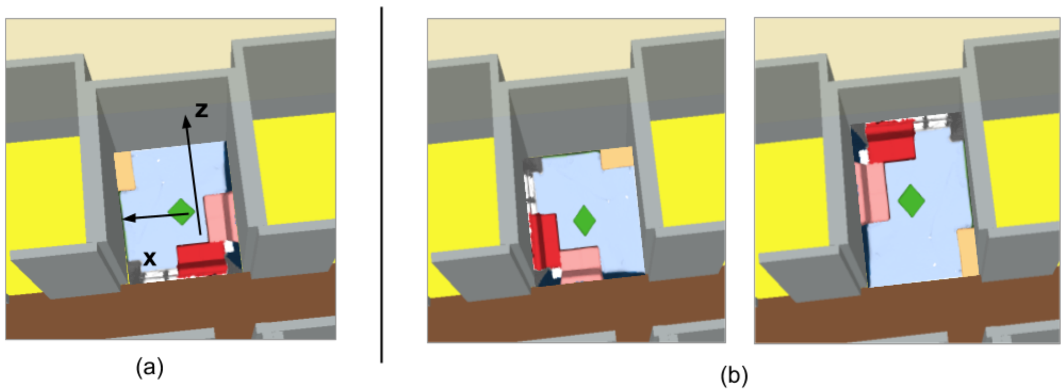


Figure 3.6: (a) Aligning the scan's longer walls with its Z-axis might not always accurately fit a room onto the floor plan. (b) Examples of an additional and manual rotation to correctly orient a room. The left image is to rotate the scan by 90°, the right one is to rotate the scan by 180°.

The last step in preprocessing a 3D scene is to generate *super-pixels*, which in this case correspond to a connected sets of mesh facets with a similar orientation. To do so, I implement the *Efficient Graph-Based Image Segmentation* algorithm described in [21] that defines a predicate for measuring the evidence for a boundary between any two facets using a graph-based representation. This algorithm helps segmenting an image, in both 2D and 3D format, by recursively selecting mesh facets based on the predicate. Additionally, the algorithm runs in a nearly linear time with respect to the number of graph edges, meaning that the size of an input scan will not affect the time executing the scene preprocessing procedure.

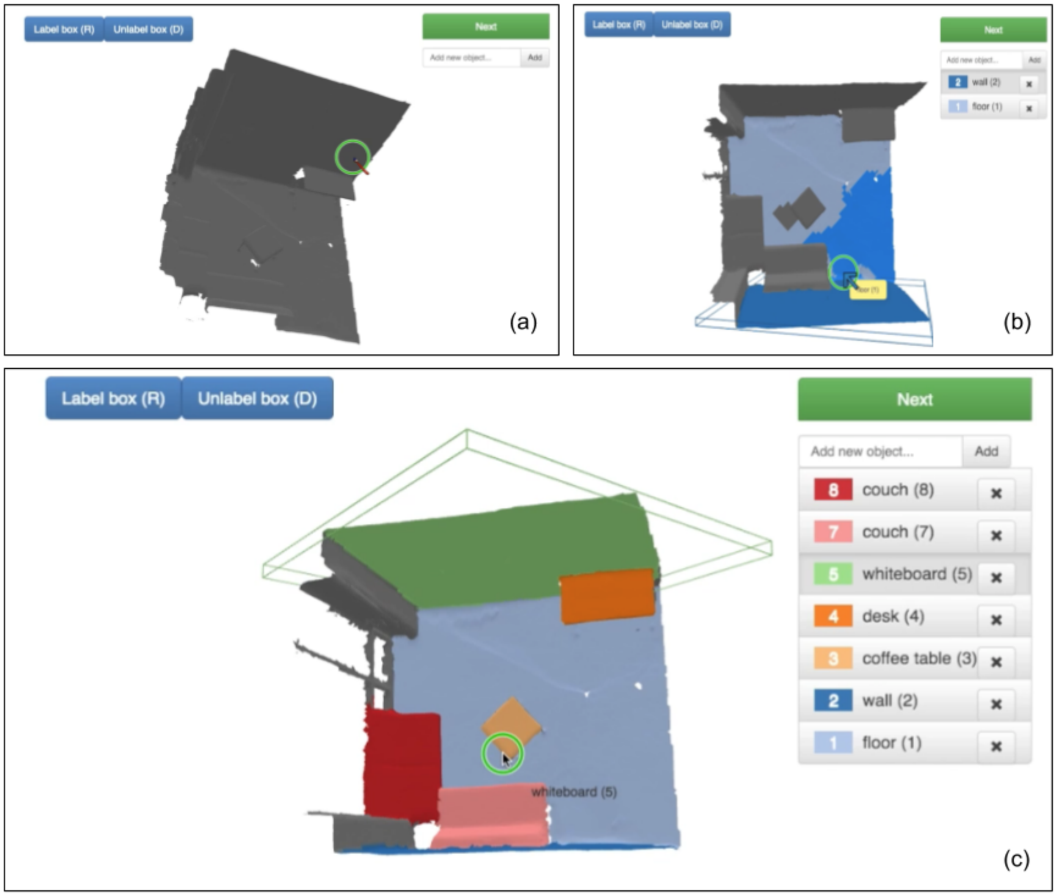


Figure 3.7: Map Population Tool's GUI for segmenting and annotating objects of interest: (a) the preprocessed 3D-scanned scene is loaded into the toolkit described in [12], (b) the on-going process of annotating a floor object, and (c) the completed segmented scene with 8 annotated objects of interest.

3.2.2 Segmentation and Annotation

The main goal of the Map Population Tool is to extract objects of interest from the 3D scans, and correctly dimension and register them within the floor plan, in the form of cuboids placed on the ground. After being oriented, rectified, and super-pixelated, the resulting 3D-scanned scene is then loaded into the web-based toolkit developed in [12] for the task of object segmentation and annotation. In this work, I integrated the toolkit into the Map Population Tool to let users manually select all super-pixels corresponding to each identified object. The interface of this toolkit is shown in the figure 3.7. Users start the segmentation and annotation process by inputting an object's name into the field *"Add new object"*, and then select any associated super-pixels (or sub-meshes) making up that object. Map Population Tool will automatically compute the bounding box surrounding these sub-meshes and assign a color for display. Note that, the calculated bounding box might also include sub-meshes that haven't been selected by users but share the same orientation as the selected ones. For example, if users select super-pixels at the upper left and lower right corners of a table object, the computed bounding box will include all super-pixels in between. Simply pressing the *R* or *D* key to include or eliminate these sub-meshes from belonging to the current annotating element. This helps speeding up the process of segmentation and annotation by reducing the number of super-pixels users have to click on. As a result, each annotated object of interest is represented as a JSON feature, which includes the overall dimension derived from the calculated bounding box, the composition of its corresponding sub-meshes, and the color specifically assigned to it. Examples of automatic super-pixelation and manual object extraction are shown in the figure 3.8.

Since the entire scanned scene is already correctly registered with the floor plan, adding the spatial description from this JSON feature into the same *sim* file is trivial. Similar as *sim*, corners of bounding boxes of extracted features are represented in terms of (x, y) screen coordinates. The conversion to (lat, long) geodetic coordinates for GeoJSON representation is performed as exactly described in the previous chapter. In fact, the GeoJSON Generator component of the Map Conversion Tool is reused in this Map Population Tool. These GeoJSON representation of elements of interest are automatically placed as extruded 3D cuboids, in their correct location on OpenStreetMap. Some examples of map population with individual objects are shown in the figure 3.9.

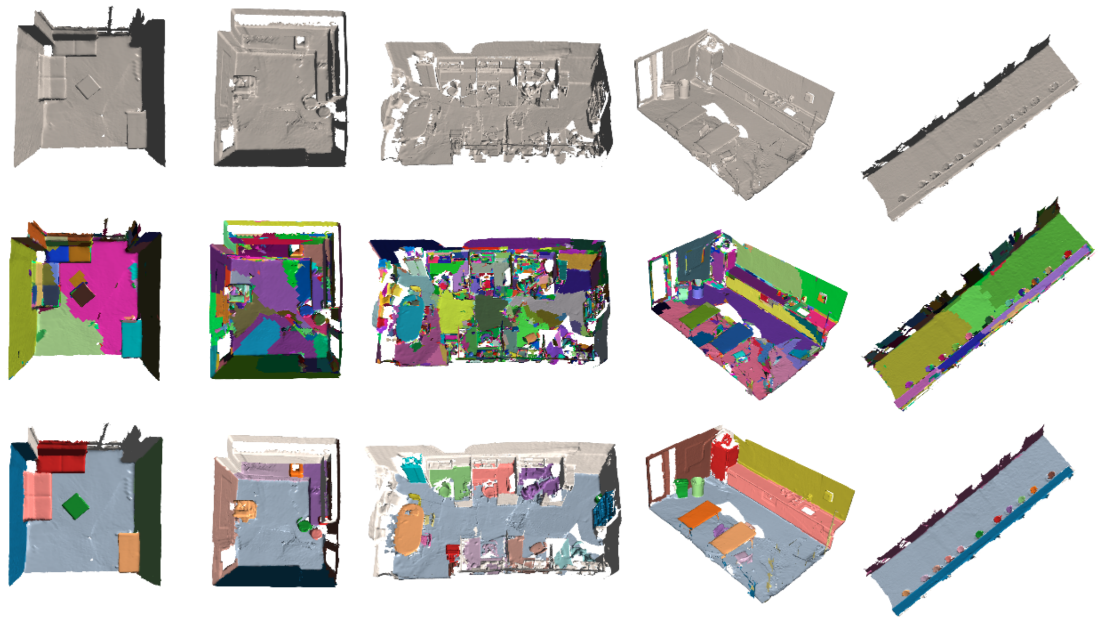


Figure 3.8: Top row: 3D scans of indoor environments using Occipital's Structure sensor. Center row: automatic super-pixelation of the meshes using the algorithm of [21]. Bottom row: segmentation of individual objects using the web toolkit described in [12].

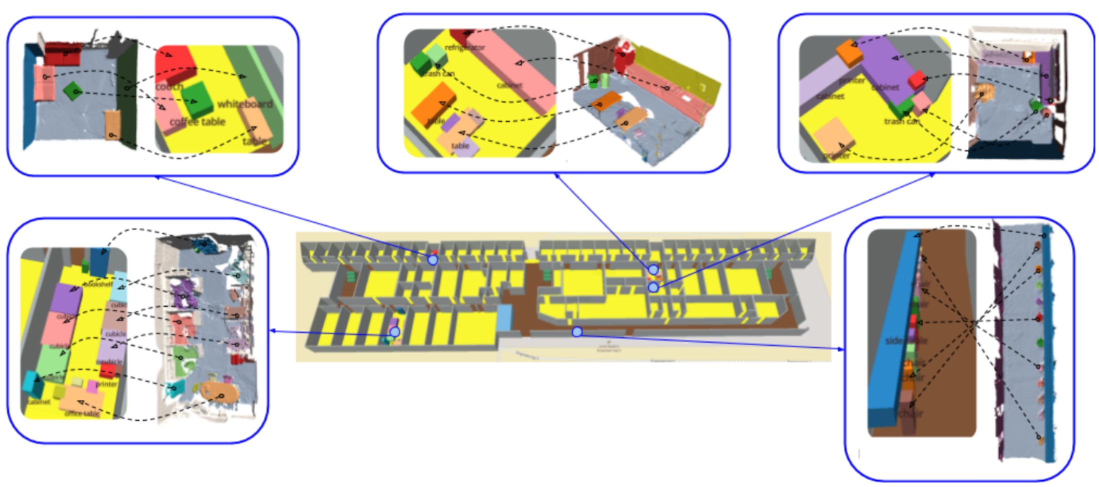


Figure 3.9: The bounding boxes of individually segmented objects are placed in the GeoJSON file containing the building's floor plan. The entire building with its objects of interest is displayed over OpenStreetMap.

Chapter 4

Multi-scale Map Authoring Tool

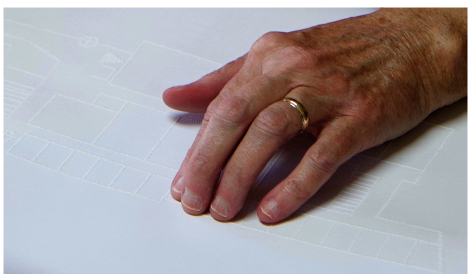


Figure 4.1: Exploring a tactile map of a building.

Access to spatial information can be vexing for people who are blind. Lacking visual input, blind individuals must rely on their own knowledge, through direct experience or otherwise, of the spatial configuration of places they are visiting. Pre-journey learning, the process of "learning a spatial environment or plan a travel route prior to actual travel" [81], is an effective way to mitigate the difficulties of independent blind travel. Multiple studies have shown that, when blind travelers are given the opportunity to "preview" an indoor route, for example using a tactile map [6, 36], they can follow a route more accurately and with fewer errors. Tactile maps give readers the layout of a venue and the spatial relationship between its landmarks, allowing them to build a prior representation of the space to be traversed in the form of an egocentric spatial image [52] or an allocentric cognitive map [64].

Regardless of different techniques in making tactile maps (i.e., embossing [9] [85], audio-tactile pairing [7] [19], 3D printing [32], [87]), it is impractical to design an one-size-fits-all tactile schema, because different users have different needs for tactile maps. For example, color-coded features on tactile maps (e.g., building, open space, bus stop, entrance, etc) are helpful for people with visual impairments but not totally blind

[83]. The ability of visually impaired people to understand a tactile format depends on many factors such as the severity of impairment, or how long they had regular sight before being impaired. Thus, automation of tactile map making tailored to the user needs is crucial in helping people with visual impairments.

Although tools such as Map Conversion and Map Population introduced in the previous chapters can help encoding a building's spatial layout for a later embossing, the type of features embossed on tactile maps to facilitate pre-journey spatial awareness and how to select distinctly discriminable tactile symbols to represent these features are challenging. My advisor, Dr. Roberto Manduchi, and I conducted a focus group to look into the importance of embossing features for pre-journey spatial knowledge acquisition, and the appropriate tactile symbols to represent these features at different resolutions. Insights from the focus group discussion have driven the decision on what level of detail should be embossed at a given scale. In this chapter, I discuss the third component of my end-to-end system, Map Authoring Tool, that produces tactile maps of indoor environments at multiple scales; and features represented on these maps at a specific scale are selected based on findings from the focus group.

4.1 Tactile Maps

Tactile maps are essential for visually impaired people to compensate for their visual loss, and to equip themselves with mental images of environments through touch. Unfortunately, creating a tactile map by hand can be time-consuming and requires specific expertise, which may be one of the reasons why tactile maps are not universally available. Current technical approaches for automatically generating tactile maps focus on either outdoor environments or indoor spaces, but not both.

4.1.1 Outdoor Tactile Map Generation

Early works in automatic generation of tactile maps relied on data from geographical information systems (GIS) to render tactile elements of outdoor environments. For instance, the *Talking TMAP1* is an extension of the existing *TMAP* framework [62], enriched with *Talking Tactile Tablets* (TTT) [61]. Users were able to use a web interface to specify the desired location and the size of a tactile map. Streets and important landmarks (e.g., parks, rivers, buildings) within the specified area, as well as their related audio information, were embedded into the map file that could be later printed offline using Braille embossers. The embossed map was then placed on a TTT, users touched down on any street shown on the map to hear its name and relevant nearby information, such as the addresses of buildings, a direction of traffic, etc (figure 4.2 (a)).

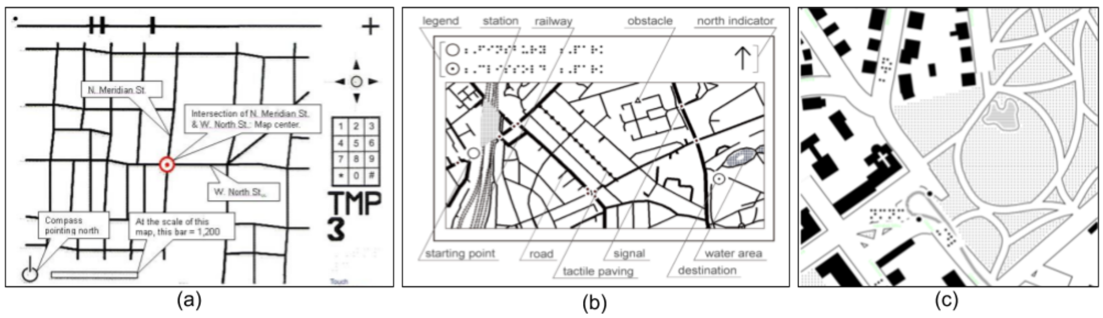


Figure 4.2: Automatic generation of outdoor tactile maps: (a) map was generated by *Talking TMAP1* described in [61], (b) map was generated by *TMAC2* described in [85], and (c) map was generated by *Mapy.cz* described in [9]

Maps generated by *TMACS2* [85] are printed on capsule papers and raised up by a heater. TMACS2, using OpenStreetMap as its underlying data source, takes an address or a point of interest as an input, generates a tactile map of any location in the world, and allows users to adjust the map's scale in the same way as Google Map. Tactile maps created by TMACS2 include roads, railways, rivers, stations, water areas, traffic signals, obstacles, and departure and destination locations (figure 4.2 (b)).

The *Mapy.cz* [9] project follows the same principles as TMACS2 by generating maps from OpenStreetMap. There are three fixed scales available, related to the standard paper sizes. The basic one was derived from the minimum width of a road in which Braille letters can be embossed. The tactile sheet for printing the basic scaled map is the A4-size swell paper. Key elements captured in Mapy.cz include building, water body, green area, industrial area, street, wall, railway, stairs, tram, and cable car (figure 4.2 (c)). These tactile maps share a common theme of displaying "important" outdoor landmarks such as buildings, parks, stations, road networks, and rivers, but not structural layouts of buildings.

4.1.2 Indoor Tactile Map Generation

Recently, there is a growing interest in technologies for automating indoor tactile map making. For example, the Audio-Tactile navigation system proposed by Papadopoulos *et. al.* generates audio-tactile maps from digital map files containing specific spatial information of a building [68]. Tang and his team introduced a hybrid method for the automatic generation of 3D indoor maps from AutoCAD architectural floor plans, which extracts semantic information from AutoCAD files (e.g, rooms, exits, etc) and constructs a topological map showing the geometric relations among different rooms in the building (figure 4.3 (a))[81]. Such information are useful for producing accessible maps that can be later used in the pre-journey learning. Similarly, Auricchio *et. al.* represented a building plan in 3D tactile graphics that allows some perception of height in [4], and Luciene Delazari designed a schema preserving the topology of an interior environment, based on indoor routings between rooms and corridors segmented from a floor plan (figure 4.3 (b)) [13]. These prior works only focused on the structural elements of a building, such as walls, doors, or staircases. Small-scale description of furniture items or floor covering, which can be useful for navigation without sight, were not considered.

(a) (b)

Figure 4.3: Automatic generation of indoor tactile maps: (a) 3D map was generated from an AutoCAD floor plan described in [81], and (b) map was generated from the indoor-routing schema described in [13]

4.2 Assessing the Perceived Utility of Multi-Scale Indoor Tactile Maps

One very practical challenge of automatic tactile map production is that maps of indoor places in digital format are often difficult to find. Even though most buildings may have detailed CAD floor plans, what is available in most cases are only pictures of these maps in JPEG or PDF format. And even when a map is available in an appropriate format, the designer, or the algorithm tasked with converting it to a tactile form, needs to decide what level of detail should be contained at a given scale. Although this *generalization* problem is common to all types of map design [55, 63], it is a particularly relevant one, and yet relatively unexplored, in the case of tactile maps. This is because tactile sensing affords relatively low spatial resolution, which reduces the achievable density of detail reproducible in a tactile map. The average spatial tactile acuity at the index finger is of about 1.2 mm [43]; and the Braille dots must have the minimum center-to-center spacing of 2.28 mm. Hence, when representing a certain portion of space (e.g., the floor plan of a building wing) on a Braille paper sheet whose standard size is 11 by 11.5 in., the designer needs to decide which details can be part of the map, and which can be removed, lest the map become too crowded, and thus difficult to read [75]. This is usually done following "tricks of the trade" or guidelines developed by expert practitioners [18, 3].

With the availability of technology to support automatic generation of indoors tactile maps, the generalization problem still remains: *What is the adequate scale, and thus the adequate level of detail, at which a map of a building should be embossed? Should maps be made available at different spatial scales? And if so, what is the optimal selection of scales so as to facilitate creation of a mental spatial representation without becoming confusing?* In order to get some insights into these questions, I and my advisor conducted a focus group with blind participants to understand the perceived utility of using multiple maps at different spatial scales, and thus different level of detail, to represent the interior of a building.

4.2.1 Method

Dr. Roberto Manduchi recruited seven participants (four identified as female, three as male) with ages ranging between 23 and 70. All participants were blind, with at most some residual light perception. They were recruited from the Vista Center for the Blind and Visually Impaired in Santa Cruz, CA[1]. All of the participants considered themselves expert independent travelers. Three of them used a guide dog for mobility, while the remaining ones used a long cane.

For this focus group, I prepared multiple copies of three tactile maps, representing the same building locations at different spatial scale (figure 4.4). These maps, embossed using a ViewPlus Max Embosser, had building name and floor number positioned at the top-center. A map scale and an arrow pointing to the North direction were embossed at the top-left and top-right corners, respectively. Building names, floor numbers, and map scales were embossed in Braille. The table 4.1 lists the tactile symbols and patterns used to represent features and spaces in the maps. These tactile graphics were determined to be distinctively discriminable. The staircase symbol was suggested in [42], while the other symbols were drawn from [50]; and the texture pattern for different spaces were proposed in [70]. I chose the following spatial scales for my maps: *structure*, *section*, and *room*.

- *Structure-level:* Due to the considered building's elongated shape, the structure-level map (figure 4.4-a) was embossed over two contiguous sheets of size 11 by 11.5 in. The other scales were embossed on a single sheet. The structure-level map displays the general layout of a building, consisting of walls, offices, corridors, building entrances, staircases and elevators. A wall was embossed as a solid line; and an office is represented by an empty untextured space enclosed by at least 4 walls. Corridors are represented as textured areas. In the structure-level map, doors and office numbers are not rendered. I believed that marking doors of each office would have led to a confusing high-density pattern. Also, there was not enough room to emboss all office numbers in Braille.

- *Section-level:* This scale represents an expanded view of a specific area inside the

[1]https://vistacenter.org/

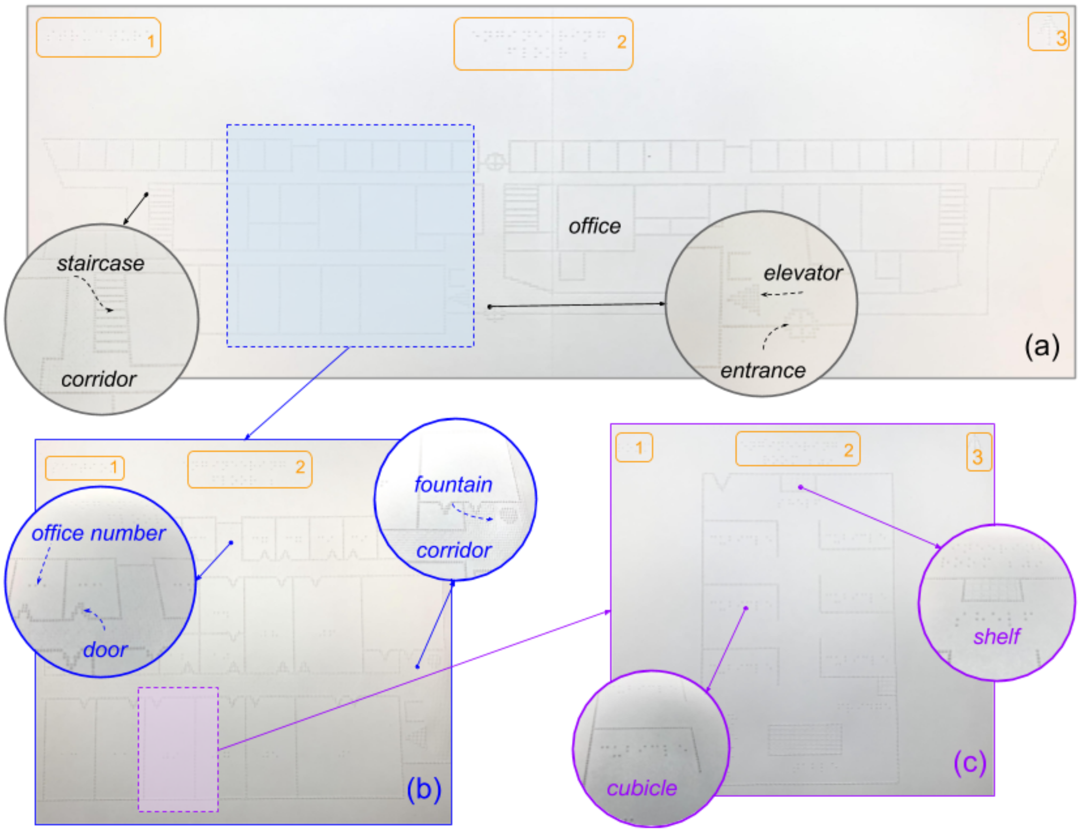


Figure 4.4: The provided tactile maps to participants in the focus group. They represent the same building at different spatial scales: (a) structure, (b) section, (c) room. The map scale, the building name, and the North direction are encoded on each map at the positions 1, 2, and 3, respectively.

building (figure 4.4-b). In addition to features already considered in the structure-level map, the section-level map displays office numbers, doors, and a water fountain. The office numbers were embossed at the center of each office, and doors were rendered as wedges along walls. The pointy top of a wedge represents the direction to enter the room.

- *Room-level:* In this scale, the map displays a room's interior in detail (figure 4.4-c). For the focus group, I mapped a laboratory, featuring a cluster of cubicles, a long table, two bookshelves, and a fridge. The names of all furniture items were annotated in Braille.

A copy of the maps at three scales was distributed to each participant at the beginning of the focus group. Participants were first asked to orient the maps such that the arrow pointing to the North direction was found at the top-right corner. Next, participants were asked to identify and locate several features in the maps: entrances, staircases, corridors, office spaces, office doors, office numbers, and furnitures (figure 4.2). These maps did not contain a legend with the symbols meaning; instead, participants were explained in words how each symbols was shaped. After an initial exploration, the focus group started in earnest. A number of questions were proposed, with the goal to elicit a discussion on the perceived utility of tactile maps for indoors in general, as well as of the multi-scale versions that were provided. The focus group was audio recorded for later transcription.

	Symbol
entrance	
staircase	
elevator	
door	
fountain	
	Pattern
corridor	
fixture	
furniture	

Table 4.1: Tactile graphics represent features and spaces at different scales.

4.2.2 Findings

The focus group transcript was analyzed independently by my advisor, acting as the moderator, and me serving as the assisted moderator at the focus group. Each of us independently identified a list of themes and issues that emerged from the conversation. Then, we met to discuss the findings and find a consensus on the set of topics used to code the relevant parts of the conversation. The resulting topics are discussed below.

4.2.2.1 Perceived utility of the maps for indoor pre-journey learning

Several participants felt that these maps would be useful, perhaps as an "add-on to the place you will like to go". A situation considered was that of a driver dropping a person in front of door B, "and you're like I wonder where I saw door B on the map. I just need to walk. Left right and I'm at the front desk." Or when visiting a medical clinic:

"The more that you can do independently having that correct information". Some asked where these maps would be kept, and how they would be made available. This concern was clearly in the mind of several participants, who mentioned past negative experiences of documents in Braille that were supposed to be available, but could not be accessed.

Table 4.2: Participants in the focus group were exploring the provided tactile maps at three different levels of scale.

Not surprisingly, the physical size of the maps was a concern for some participants – especially for the map spread over two sheets. Smaller is better, especially if maps were meant to be carried along in a trip. In this case, it would be preferable if they were embossed on a plastic material that could be rolled up. But even so, one would need to find a flat area to flatten the map on, which may be unpractical.

While a sense of independence was generally considered valuable, some of the participants noted that often there are people nearby who can offer help. This may reduce the perceived importance of maps, especially at the room-level scale. Interaction with sighted bystanders is not always easy, though, such as in crowded situations: "If there's a lot of other people around, I don't have a clue. There's too much input coming." Or, bystanders may sometimes be too eager to help: "If there's other people in there you know we hesitate just for a minute they're gonna be 'hey can I help you, you know and blah blah'".

4.2.2.2 What can be learned from a map?

One participant, who has been blind since birth, said that she felt the maps, or at least the map at largest scale, did give her a general picture of the layout, but she would not get anything "extra" from the map than if someone had just explained the scene to her (e.g. enumerating the corridors and the staircases.) This is because it was "difficult for [her] to picture a building layout from a two dimensional map". Interestingly, this participant felt that the map would be more useful post-facto – after "wander around and screw up like I screwed up a few times, I could look at a map and go 'oh that's what I did'. But I have a harder time going the other way."

Maps convey information about the size of spaces. Whether this information is easy to use when building a mental picture of a place was debated, with one participant feeling that at least the relative size of two spaces could be easily inferred from the map, while another feeling that, knowing the exact length of, say, a corridor, was not particular useful, besides going "Gosh, this is big!". Maps could be made to also contain wayfinding information to reach specific destinations, although this was not the case for the sample maps presented to this focus group. For example, a participant mentioned a frequent situation of taking an elevator, then not knowing whether to turn left or right after getting off – something she felt would be useful to have in a map.

4.2.2.3 One or more scales?

The need for multiple scale levels was appreciated by several participants, in particular for the first two scale factors. Some agreement emerged on the structure-level map being the most useful one, provided that it could contain room numbers, or that it could somehow be combined with the section-level map. One interesting observation was that different scale maps may be useful for different experiential levels. As one participant put it: "...really these maps are useful during different phases of your familiarity with the building. So the first time you go into a building, the big map is really useful. After you've been there a few times the big map will be less useful than it used to be...The smaller or the middle map might be more useful but once you've been there a few times the usefulness of that map falls off as well around."

For what concerns the room-level map, there was substantial disagreement on its utility. Some participants felt that all three scale levels are useful; for example, if "I walk in the door and I know where the front desk is and I know whether there would be chairs off to my right or my left or whatever I find, and wait for my name to be called". Others thought that the room-level map was '"kind of secondary", and that "the time and effort it takes to make that is less useful because the variables are too high", or that, due to the possible presence of movable objects of furniture, "when you have to get down to there, the map is out of date before you finish drawing". The way blind individuals negotiate a room-level space may also be different than for larger spaces (e.g., corridors, halls). In one participant's words: "From a practical point of view, I would walk into a room and stand at the door, listen to get a sense of the size of the room…I would go around the perimeter and come up and and just figure it out. I wouldn't take the time to use [the room-level map]".

4.2.2.4 Universal map access

Maps need to use symbols that must be understandable. Symbol standardization is an important issue; the need for using symbols and textures that are easy to interpret also emerged in the discussion. One participant pointed out that some symbols (e.g. building entrance) should be designed such they catch the user's attention right away, so that they are easy to find in the map. Also, given that these maps are at different scales, some additional information would need to be added to specify what kind of "view" is represented in the map. Some participants commented on the trade-off of using Braille character in lieu of symbols (e.g., to label an entrance). While Braille may take more space, it is easier to interpret – but only for those who know Braille. Indeed, it was noted that many of the potential users of these maps may not be able to read Braille.

4.2.3 Discussion

Tactile maps are arguably a case of unexpressed potential. Many people believe they could be a valuable tool for pre-journey spatial learning, yet they are seldom used in practice. Part of the problem stems from very practical considerations: where to find these maps, when and how to explore them. It is possible that new refreshable display technology [44], or vibro-tactile display on commodity tablets [29], will alleviate some of this practical issues. Another major challenge is how to represent details at a wide range of levels. Lacking the ability to zoom in or out or to pan the map content (something that could theoretically be possible with refreshable or vibro-tactile display), multiple embossed maps at different scales are necessary to "see the tree and the forest". Our focus group was designed to gather feedback on this type of multi-scale tactile maps for indoor environments.

Perhaps not surprisingly, general consensus was often difficult to find on various themes. Some participants loved the idea of accessing indoor maps, other didn't see a lot of value in them. Some appreciated all three scale levels, others would just keep the structure-level map if it could contain more detail (note that, in the case of the building considered in our maps, this would be impossible to achieve due to the constraint imposed by tactile sensing resolution). This may again be driven by very practical considerations: having to manage multiple sheets of embossed paper in order to access different levels of detail for the same place is cumbersome. The room-level map received the most discordant comments. Given the wide variety of content that can be found in a room (e.g., table and chair vs. cubicles and desks vs. bathroom stalls and appliances), it may be impossible to generalize an assessment of the value of a map at this scale level from a single example.

4.3 Map Authoring Tool

The problems of cartographic generalization and automatic tactile map production have been investigated independently over the past few decades. However, the use of standard generalization techniques for the production of indoor tactile map that could be used by visually impaired travelers has received relatively little attention. Recent advances in machine learning techniques have prompted researchers to revisit the problem of cartographic generalization. Several neural-network models have been developed for the tasks of recognizing, grouping, and typifying buildings [76, 22, 82]. However, these models are only able to learn and predict a building's contour and the geographical distribution if groups of buildings, and may not generalize well for the representation of the layout of an indoor space.

Findings from the focus group discussed in the previous section indicate that embossing the map of a building at different levels of detail could be provide useful spatial information for pre-journey learning. In this section, I introduce the last component of my end-to-end systematic tool, namely Map Authoring Tool, that produces tactile maps of indoor environments at different spatial scales on demand, including the

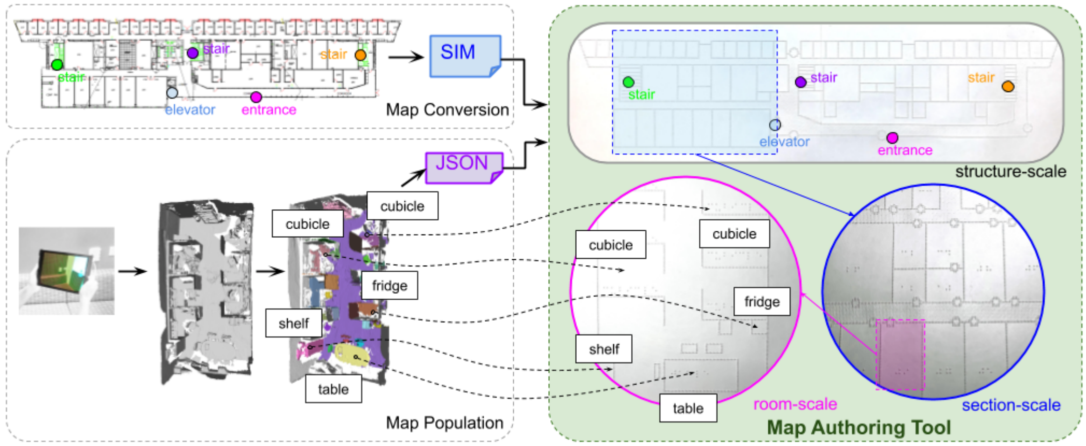


Figure 4.5: Map Authoring Tool produces tactile maps at three different scales, based on the spatial information captured in Map Conversion Tool and Map Population Tool.

building's architectural structure, zoomed-in of specific areas (*sections*), and a small-scale layout of a room, highlighting the spatial relationship among objects in the room (figure 4.5). This Map Authoring Tool produces a digital tactile map file at a desired scale, based on the building's structure represented in *sim* and the semantic layout of an interior space encoded in a JSON map that are collectively acquired from the previous toolboxes (Map Conversion Tool and Map Population Tool).

4.3.1 Tactile Graphics Resolution

Following the study described in [78], my map authoring tool renders segments with length of at least 0.5 in (12 mm), with a minimum distance of 0.2 in (5 mm) between two segments. For easy discrimination, symbols representing different features have a minimum diameter of 0.25 in (6 mm), with minimum distance between two symbols of 0.5 in (12 mm). Braille characters for annotation of objects and spaces have size of 0.16 in $\times$ 0.26 in (4 mm $\times$ 6 mm) [38].

4.3.2 Tactile Map Design

The produced digital tactile map files can be printed on a 11.5 in $\times$ 11 in embosser sheet, with resolution of 20 DPI and a 0.5 in (12 mm) margin on all sides.

It is partitioned in two sections: the *header* (10.5 in × 3 in; 267 mm × 76 mm); and the *body* (10.5 in × 7 in; 267 mm × 178 mm). The header contains on the top-left the building name, the floor number, and the map scale, as well as an arrow pointing to the North on the top-right. The body has the tactile map at the desired scale.

There are two types of embossed features in the map: ***structural*** and ***interior***. Structural features are those traced from a building's floor plan, including entrances, staircases, elevators, escalators, walls, and doors. Interior features represent objects that are segmented and annotated from 3D scans, such as tables, cubicles, shelves, and other pieces of furniture. This map authoring tool allows one to choose between three different scales:

1. ***Structure-scale***: General building layout, consisting of rooms, corridors, and *structural* features. Wall are embossed as solid segments, while rooms are represented by untextured areas, enclosed by at least 4 walls. Corridors are rendered as textured areas. In this scale, the room number or door is usually not rendered due to space constrains.

2. ***Section-scale***: Expanded view of a specific area inside a building. In addition to the features already considered in the structure scale, a section-scale map also displays room numbers, room doors, and any available *interior* features. Room numbers are embossed at the center of each room, and doors are represented as circles along walls [50].

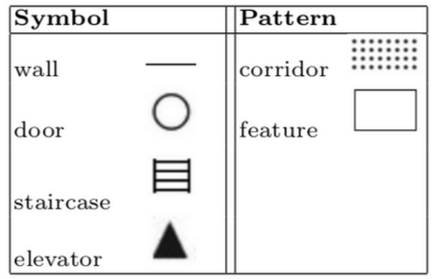
3. ***Room-scale***: The layout of a small area (typically a room), including walls, doors, and all annotated *interior* features.

The features to be embossed at different map scales when space permits are shown in the figure 4.6-a. Tactile symbols and patterns used to represent such features are also listed in the figure 4.6-b. Note that the staircase symbol embossed in our maps was suggested in [42], while the other symbols were drawn from [50]. The texture patterns were proposed in [70].

My tactile map authoring tool limits rendering of a room or a section of a building to a single embosser sheet. The building's general layout (structure-scale) can span multiple pages; this allows for rendition of very elongated buildings. The tool

Structure	entrance, staircase, elevator, escalator, wall, door
Section	wall, door, bench, couch, desk, table, shelf, chair, directory, board
Room	wall, door, cubicle, desk, couch, bench, shelf, table, cabinet, nightstand, fridge, bathtub, toilet, sink, chair, board, printer, trash can

(a)



(b)

Figure 4.6: (a) Embossed features at different map scales, and (b) tactile graphic symbols representing these features.

automatically selects the features to be rendered based on the selected scale, while adhering to tactile resolution constraints mentioned in the section 4.3.1. Figure 4.7 shows an example of Map Authoring Tool's user interface at different map scales, along with the produced tactile maps.

Figure 4.7: The top row shows the Map Authoring Tool control panels. The bottom row shows the generated tactile maps at 3 different scales: (a) structure, (b) section, and (c) room.

4.3.3 Room-scale Editor

Interior features at room-scale are represented by their bounding boxes, which are shaped as vertical-oriented cuboids. These cuboids are shown as rectangles in the tactile map, having sides parallel to room's walls. Unfortunately, poor performance in acquiring 3D-scanned scenes or segmenting objects of interest is occasionally inevitable. This produces overcrowded, unaligned, or overlapping embossed features, resulting in tactile maps at the room level to be confusing. In some cases, multiple different features appeared within a room are stacked on top of each other (e.g: a printer is on the top of a table); thus, the generated tactile map files are guaranteed to have cuboids nested, increasing the density of represented features but decreasing the perceived utility of a tactile map.

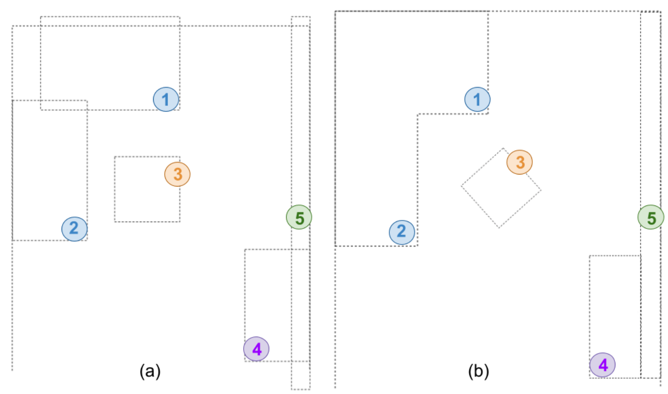


Figure 4.8: A generated tactile map at the room-scale (a) before and (b) after being edited.

In order to facilitate a clean-cut map generation and to convey essential spatial layout of an indoor environment, the Map Authoring Tool includes a simple editor (only available at the room-scale) that allows users to translate, rotate, scale, and delete features within a boundary, making the maps easily perceivable through touch. This editor also allows users to merge multiple cuboids representing the same type of object into a single polygonal feature, freeing up more map spaces for Braille annotation (figure 4.8). This can be helpful for objects with complex shapes. In this work, I employed the Union Boolean operation on polygons, developed by Martinez et al. in [58], for the task of features grouping. As shown in the figure 4.8, the feature #1 (a couch) was translated and then merged with another couch (#2); the coffee table (#3) was rotated to its correct orientation; and a desk (#4) and a whiteboard (#5) were both scaled down to their correct dimension. Note that if two objects (e.g., a table and a printer) are physically on top of each other, they will be represented as two stacked cuboids, which will be mapped as two nested rectangles. In this case, the innermost rectangle can be removed using the editor.

Figure 4.9: Indoor tactile maps generated by Map Authoring Tool at different scales.

The Map Authoring Tool's user-friendly interface allows one to select any region, room, or type of indoor features to be embossed. For example, one might choose to render only features that are close to walls (countertops, benches, or shelves) vs. furnitures positioned randomly in the middle of a room (tables, chairs, etc). Sample multi-scale indoor tactile maps generated by this tool are shown in figure 4.9. Maps labeled as (1) show an entire building's general layout at the structure scale; those labeled as (2) render specific sections, or zoom-in, inside a building; and tactile maps labeled as (3) emboss particular office spaces at the room-scale. Map Authoring Tool represents a corridor or a hallway as a textured pattern, while leaving an office's floor blank. The figure also show 3D room scans (a), their segmentation (b), and the results after manual editing (c), along with Braille annotations. These Braille annotations at the section-scale denote room numbers; whereas, in the room-scale tactile map, they annotate the interior features segmented from a 3D scan. I believe that the Map Authoring Tool is an innovative and useful tool for the automation of tactile map making, and that its simplicity of use may appeal to both practitioners and casual users.

Chapter 5

Multi-scale Embossed Tactile Maps for Pre-journey Visualization: An Experimental Study

Most often, a technological solution is discovered only to find out later that it does not fit an actual user's needs. To avoid such evitable pitfall, I and my advisor, Dr. Roberto Manduchi, had conducted a focus group discussed in the previous chapter to understand what my systematic application should produce in order to promote a pre-journey spatial awareness. Findings from that focus group motivated me to develop the Map Authoring Tool that produces tactile map files at a desired resolution. In this chapter, I describe an experimental study on the multi-scale embossed tactile maps generated from this Map Authoring Tool. The ultimate goal of this study is to ensure that accessing to the spatial information encoded in my tactile map files could lead to increased spatial knowledge acquisition, hence encouraging visually impaired people to have more self-confident travels in unfamiliar indoor environments. The observations, quantitative results, and survey outcomes from this study not only confirm the usability of these tactile maps but also shed some light on the different strategies one might take to read a map, as well as insightful suggestions on making such maps even more accessible.

5.1 Method

Dr. Roberto Manduchi recruited nine participants, five identified as female and four as male. Seven of them were born blind with at most some residual light perception, and the other two gradually lost their eyesights over time at a young age. Five participants have a little to no experience with tactile maps, while the remaining are only familiar with tactile graphics of continents, block and intersection, and college campuses. None of them has never been exposed to tactile maps of a building similar to the ones used in this study.

In this user study, I prepared nine copies of four tactile maps that were embossed on a single sheet of paper using a View Plus Embosser. Each map has the name and the arrow pointing to the North direction positioned at the top-left and top-right corners, respectively. Two of them, namely "Sample Room 24" and "Sample Full", render a fictional building at two different spatial scales (section and room), serving as the rehearsal for getting participants familiar with the actual generated tactile maps. The other two, "Full" and "Room 09", generated from the Map Authoring Tool represent a laboratory at the room level and a zoom-in of the building containing this lab at the section scale. More details, the "Sample Room 24" map is a simple version of the "Room 09" map, which has only two cubicles and one table; and the "Sample Full" version simplifies the "Full" map by containing a circular corridor with 4 rooms having doors located in different directions. Features represented in the actual generated maps (Full and Room 09) are drawn from the semantic spatial information encoded in the vectorized *sim* map and the JSON feature collection that are collectively acquired using my end-to-end systematic tool described in the previous chapters.

In addition, I also embossed the key sheet which explains tactile symbols being used to represent features on my maps. This key sheet is divided into two columns: the left one of six symbols and the right one of three symbols. Each symbol is followed by an explanation embossed in Braille. I packaged these five tactile sheets into an envelope and then mailed out to the participants. The figure 5.1 shows the key and embossed tactile maps used in our study. Once the participants received it, Dr. Roberto Manduchi scheduled a session with them over the phone or zoom to conduct the study. Before the session took place, Dr. Manduchi specifically asked them to not open the envelope to ensure the study's integrity. In order to help participants quickly and correctly identifying a map, each one has a corner cut-off at a specific location. Precisely, the "Sample Full" map was cut off at the lower left, while the "Sample Room 24" map was chipped off at the lower right corner. In the same fashion, the "Full" and "Room 09" maps were also clipped at the upper left and upper right corners, respectively (figure 5.1). Features in my generated tactile maps were embossed as following:

- lines represent walls which may have openings such as doors or cubicle entrances.
- circles are doors which may connect a room to room, a room to hallway, or a hallway to the outside of a building.
- a set of parallel horizontal segments denote a staircase.
- a rectangle with Braille annotation inside or beside it represents a fixture or furniture indicated by the annotation. In this study, it could be a table, a shelf, or a fridge.
- a rectangle with a number inside of it and a circle along its side can be interpreted as a room. Note that, a room must have at least one door, and it can connect with another room through its sharing door.
- a texture pattern enclosed by lines indicates a hallway.
- a filled triangle in a flat area represents an elevator.
- a filled circle surrounded by a texture pattern locates a position of a water fountain on a hallway.

In this user study, three experiments were carried out: *Map Reading*, *Feature Pointing*, and *Path Following*. In each experiment, the participants started with the "sample" version as a rehearsal to get familiar with the provided tactile maps, and then moved to use the generated ones for an actual study. Each experiment is followed by a set of questions, which was designed to evaluate whether the participants were able to perceive the spatial information conveyed inside the maps, including the building's general layout and the semantic relationship among objects appearing within. Each session

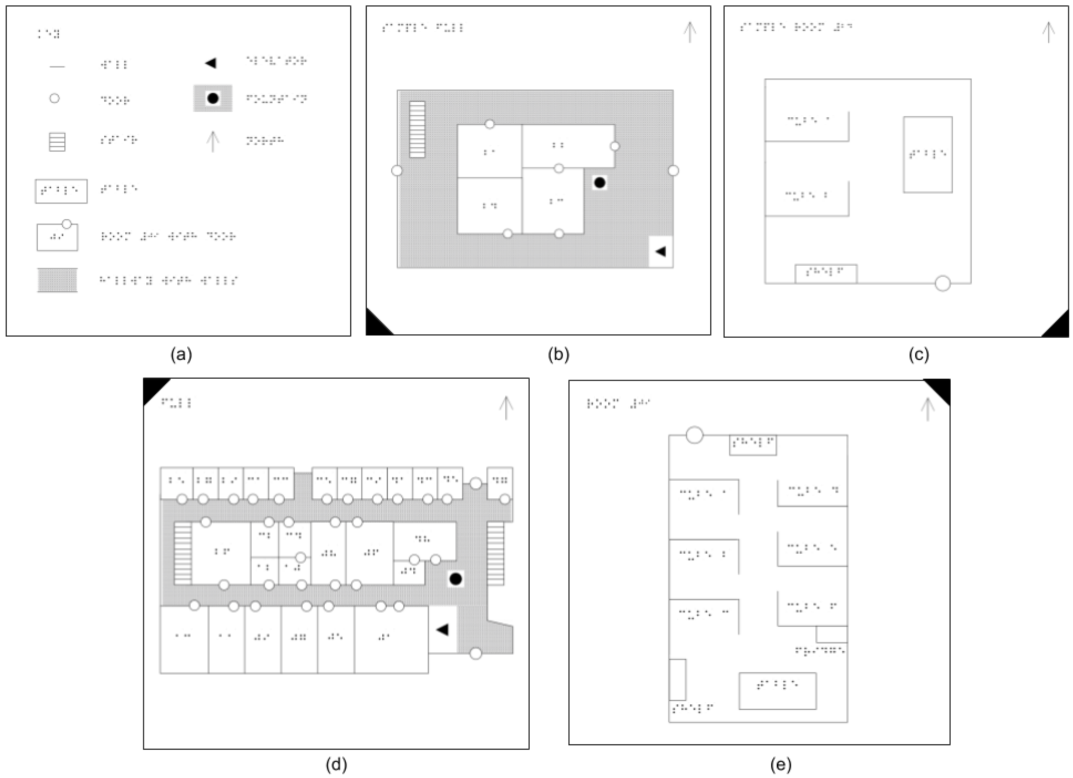


Figure 5.1: Key and multi-scale tactile maps used in our study: (a) the map key includes 9 tactile symbols representing embossed features and their explanation; (b) Sample Full map at the section level for rehearsal; (c) Sample Room 24 map at the room level for rehearsal; (d) Full map at the section level; and (e) Room 09 map at the room level. Note that the black triangle denotes the cut-off corner on each map to make it easily identified.

of this user study was scheduled for two hours, and averagely, the participants completed it an hour and a half. All sessions were audio-recorded for a later transcription. During the session, Dr. Roberto Manduchi were the moderator leading the experiment, while I played a role as the assisted moderator.

5.1.1 Experiment 1: Map Reading

5.1.1.1 Part 1: Rehearsal

Before started, Dr. Manduchi asked the participants to find a plat surface large enough to put on two maps at the same time, as this will be helpful in later experiments.

Also, the participants were asked to orient the maps so that the arrow pointing North embossed on each map is always located to the top-right position. We began our study by having Dr. Manduchi to explain the maps and their key to the participants.

First, we asked the participants to explore the "Sample Full" map by conducting an imaginary walkthrough. Specifically, they were asked to find all features available along the way, and describe in details where they are. Next, the participants had to walk the entire circular loop of the corridor (by following the textured area) to find the exit doors. Note that exit doors join a hallway (a textured area) with the outside of the building (a textureless area unbounded by walls). Once the perimeter walk ended, we asked the participants to find the room 22 and imagine to be in that room. The participants now needed to enter the room 23 starting from the room 22 via their interconnected doors, and then exited to the hallway through another door of room 23. The figure 5.2 illustrates paths the participants had to walk in this rehearsal.

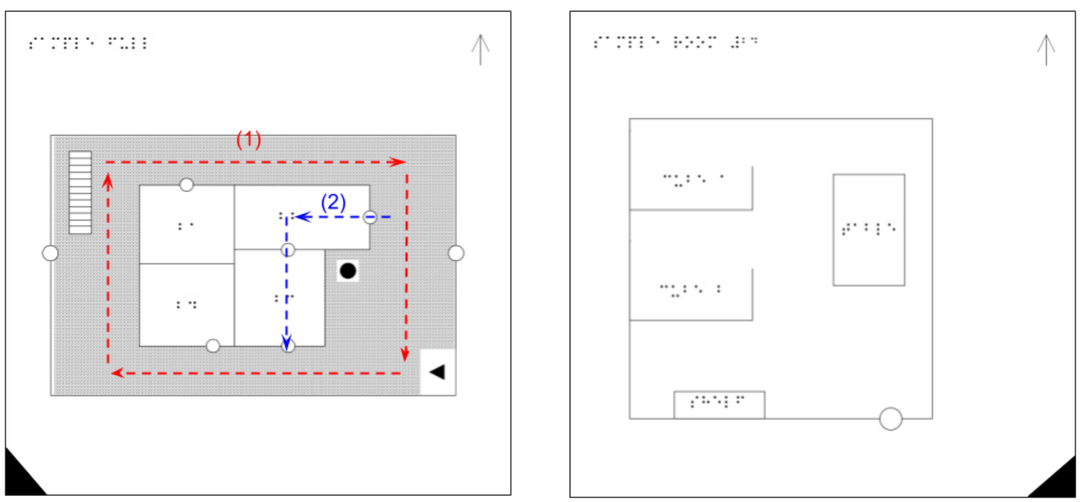


Figure 5.2: Tactile maps used in the rehearsal process: the left one is the Sample Full at the section scale, and the right one is the Sample Room 24 at the room scale. In the Sample Full map, the route (1) is the experimented perimeter walkthrough, and the route (2) is the path traveling through interconnecting rooms.

In the last part, we asked the participants to take out the "Sample Room 24" map (figure 5.2), and we explained that this is an expanded version of room 24 located in the "Sample Full" map. The participants were tasked to find the Cube A inside this map scale. Note that a cubicle is surrounded by walls and it has an opening that is not a door. In addition, we also asked the participants to find the table and shelf features represented on the map, which are rectangle with a Braille annotation embossed inside of it. During the rehearsal, we constantly reminded the participants that walls are impassable, and they can enter or leave a room (or a cubicle) through its door (or its opening).

5.1.1.2 Part 2: Trial

The trial started once the participants were familiar with the provided tactile maps and completely understood the tasks they were about to perform. In the first part of this trial, we asked the participants to explore the "Room 09" map (figure 5.1-(e)) and then to answer the following questions. We timed each question for 5 minutes or 15 minutes for the entire set.

1. Can you name all of the cubicles that are in the room?
2. How many shelves are in the room, and which cubicle is closer to each shelf?
3. Is Cube B to the East or to the West with respect to Cube E?
4. If starting from the door, would it be shorter to walk to Cube C or Cube D?
5. If starting from the table, would it be shorter to walk to Cube A or Cube F?

In the second part of the trial, we explained the general layouts of the "Full" map (figure 5.1-(d)). Specifically, the map contains a circular hallway with staircases on the both the West and East side of it. There are rooms running along both the North and South legs of the hallway, and there are also rooms joining the two legs. As in the previous task, the participants were asked to explore the map while answering the following questions. We timed each question for 5 minutes or 20 minutes for the entire set.

1. Can you name the rooms that are in the South edge of the building?
2. How many exit doors are there at the North edge of the building, and, for each such exit, what is its closest room or rooms?

3. How many exit doors are at the South edge of the building, and, for each such exit, what is its closest room?

4. Can you name all rooms that have two doors? And for those that do, which ones communicate to another room, and what this other room?

5. Is room 47 to the East or to the West of the water fountain?

6. If you were to walk from the room 47, would it be shorter to walk to the North exit or to the South exit?

7. Is it shorter to walk from the room 09 to the West or to the East staircases?

5.1.2 Experiment 2: Feature Pointing

In this experiment, the participants had to imaginatively orient their body in a certain direction and then pointed to various locations inside the building. They could express the pointing direction in term of cardinal points (North, East, South, West), clock-face direction, or just simply "up, down, left, right" with respect to their body. Similar as the previous experiment, each question is timed for 5 minutes or the entire set is timed for 15 minutes. Below is the set of questions being used in this experiment:

1. Suppose you entered the building from the North exit/entrance door. As you enter, you will be facing South (or down). *Please point at the direction of the South exit/entrance door.*

2. Suppose you are standing next to the water fountain, and that you are facing West (or left). *Please point at the direction of the Western (or left) staircase.*

3. Suppose you are exiting through one of the doors of Room 01 and you will be facing North. *Please point at the direction of the South exit/entrance door.*

4. Suppose you are in the hallway to the West (or left) of the Western staircase and you are facing North. *Please point at the North exit/entrance door.*

5.1.3 Experiment 3: Path Finding

5.1.3.1 Part 1: Rehearsal

In the last experiment, the participants were asked to keep both the "Full" map and the "Room 09" map side by side. We then instructed them to follow an imaginary path with their fingers. The path starts from the room 32 and ends at the Cube B of room 09, expressing in terms of turns, intersections, locations of rooms, doors passed by to the left or to the right, cubicles, corners, and any other landmarks (e.g., stairs, elevator, fountain, etc). Each turn should also be expressed as a left or right turn. In addition, the participants were also have to confirm their current facing orientation after each turn. The figure 5.3 illustrates such path and its verbal description is listed below:

- Start with the "Full" map, locate the room 32 and then face North.
- Walk through the door, turn left into the corridor, and then face West.
- Wal past two doors to the right, one to the left, and another two to the right; turn left at the corner to face South.
- Walk pass the stair case to arrive at another corners; turn left at the corner again to face East.
- Walk pass one door to the right, one door to the left, and another to the right. Arriving at the room 09, turn left to face South for entering the room
- Switch to the "Room 09" map, locate the door, and face South.
- Enter the room, walk South until hitting the wall of Cubicle A, and then turn left to face East
- Walk along the wall of cube A until finding its corner; turn right at the corner to face South.
- Walk pass the opening of cube E on the left, and then arriving at the opening of cube B on the right.
- Turn right to enter the cube B and end the path.

Once the participants completed a path, we then asked them to rotate their body facing North and point to (1) the door of the room (in this case, the door of the room 09), and (2) the origin of the path (in this case, it is room 32).

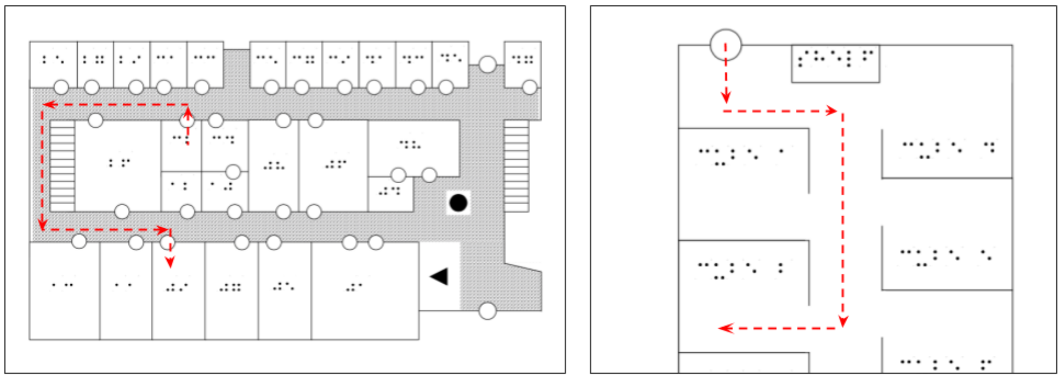


Figure 5.3: The sample route used for rehearsal in the path finding experience, starting from the room 32 and ending at cube B of room 09.

5.1.3.2 Part 2: Trial

In the second part of this experiment, the participants had to carried out two following paths independently: (1) Assuming that all doors interconnecting rooms are closed, describe the path to Cube D of the room 09, starting from the North exit, and (2) Starting from the Cube D inside room 09, describe the path to the North exit, assuming that all doors interconnecting rooms are all opened. For each path, we constantly reminded the participants to enumerate all doors and landmarks they passed on their left, and to denoted their orientation after each turn. Once a path is completed, we asked the participants to point to the entrance door of the arrival room or to point to the origin of the path.

5.1.4 Exiting Survey

The last part in this user study was a survey, which help us to understand the perceived utility of the automatically generated maps. Each statement in this survey can be rated on the scale from 1 to 5, where the score of 1 indicates that I ”strongly disagree" and the score of 5 means "I strongly agree". Additionally, the participants were encouraged to provide any comments or express any of their feeling toward the maps. The list of statements in the exiting survey is following:

1. The Full map was easy to read
2. The Room 09 map was easy to read
3. The symbols were easy to identify
4. Finding a certain room in the Full map was easy
5. Finding a path in the map was easy
6. I feel that, when imagining to follow a path in the map, I was able to correctly identify the orientation of my body at all times
7. When imagining to follow a path with a transition from the Full map to the Room 09 map, this transition was easy for me to follow

5.2 Results

5.2.1 General Observations

The participants are proficient in Braille reading, except P1. All of them were able to recognize the key sheet, to differentiate two different map scales being used in the study, and to distinguish a variety of symbols and patterns representing the embossed features. The participants' individual characters were clearly reflected in their interaction with the embossed tactile maps. Some explored the maps very methodically; upon orienting the map so that the arrow pointing North locates at the top right, they would quickly skim through the map layout by either panning their fingers from left to right and top to bottom, or tracing along the perimeter of the map in the clockwise order. Others had a more laissez-faire attitude, and just kept searching the map until they found what they were looking for. During the rehearsal, all participants were able to mentally construct the general layout of a building, including embossed features, room numbers, and doors interconnecting rooms.

Surprisingly, in the experiment of feature pointing, all participants performed generally well. They showed no trouble or difficulty when being asked to mentally locate a feature while self-orienting their body. Five of the participants expressed the pointing direction in the clock-face ordering, while the others chose to use the cardinal points of North, East, South, West. Interestingly, the participant P9 accurately reasoned the pointing direction in the degree notation. Instead of using cardinal points or clock face ordering, he applied dead reckoning process into calculating the position of the pointed feature with respect to his body in terms of degrees. Due to the time constraints, the participant P5 skipped this experiment.

Although the participants were able to recognize the general layout of a building embossed at the section scale, four of them missed the question 4 in the second part of the experiment 1 which asks to find all rooms having two doors, and for those do, which has door interconnecting to the other rooms. These participants at least successfully found a room having two doors embossed on the map, but unable to detect any door that can lead to another room. As a result, in the experiment 3, when being asked to describe a path from room 09 to the North exit that can shortcut through any interconnecting rooms, these participants couldn't find such path. Instead, they followed the reverse of their original path conducted in the first part of the same experiment.

5.2.2 Quantitative Results

Table 5.1, 5.2, and 5.3 show the quantitative results of our three experiments. Each question in each part of the trial is worth one point, except for the experiment 3. Specifically, the possible following paths in the first part of the third experiment include two routes: along the South-edge corridor and along the North-edge corridor (blue and red lines shown in the figure 5.4). Similarly, valid paths for the second part of the same experiment also include the route along the South-edge corridor and the route along the North-edge corridor after *passing* through interconnect rooms (violet lines shown in the figure 5.4). All these routes have four critical turning points; hence, the maximum point for this experiment is eight, four for each path.

	Room-scale Map					Section-scale Map						
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**
P1	1	1	1	1	1	1	1	1	1	1	1	1
P2	1	1	1	1	1	1	1	1	1	1	1	1
P3	1	1	1	0.5	0	1	1	1	0.5	1	1	1
P4	1	1	1	1	1	1	1	1	0	1	1	1
P5	1	1	1	1	1	1	1	1	0.5	1	1	1
P6	1	1	1	1	1	1	1	1	0.5	1	1	1
P7	1	1	1	1	1	1	1	1	1	1	1	1
P8	1	1	1	1	1	1	1	1	1	1	1	1
P9	1	1	1	1	1	1	1	1	1	1	1	1

Table 5.1: Experiment 1 quantitative results.

5.2.2.1 Experiment 1

The experiment 1 data show that participants were able to recognize the general layout of the building embossed on the "Full" map, and the spatial information among features being represented on the "Room 09" map. Precisely, the participants correctly named all six cubicles (cube A, cube B, cube C, cube D, cube E and cube F) available inside the room 09, identified two shelves at the North and South East corners, and reasoned a walkable distance between a door or a table to a specific cubicle. Similarly, the participants successfully listed all rooms along the South-edge and North-edge of the building, as well as estimated the spatial distance from a specific room to a staircase or an exit. Perhaps not surprisingly, the most challenging task in this experiment is to find a room having two doors, which can communicate to another room. The participants P4, P6, and P7 were struggling with this task. They were only able to find that room 48 and room 06 having two doors, but could not relate any information on the interconnecting rooms through their sharing door. In contrast, the other participants could at least identify two such rooms (room 26, 32, 12, 34, 10, 08, 06, 48, and 01) and correctly pointed out which room can communicate to another (room 34, 10, 48).

	P1	P2	P3	P4	P5	P6	P7	P8	P9
Q1	1	1	1	1	0	1	1	1	1
Q2	1	1	0	1	0	1	1	1	1
Q3	0	1	1	1	0	1	1	1	1
Q4	1	1	1	1	0	1	0	1	1

Table 5.2: Experiment 2 quantitative results.

5.2.2.2 Experiment 2

In opposition to what we expected, the participants performed extremely well in the experiment 2 which is feature pointing. They were able to mentally orient their body as instructed and accurately pointed to the asked origin. In this experiment, we allowed the participants to express the direction as "in front of me", " to my right", or "at an angle of X degrees to my left". Fortunately, the participants P2, P3, P6, P4, and P7 chose the clock-face orientation to indicate their pointing position; and the participants P1, P8, and P9 employed the cardinal points of North, East, South, West. Occasionally, the participants P4 and P7 switched back and forth between the cardinal points and "up/down/left/right" expression.

5.2.2.3 Experiment 3

	P1	P2	P3	P4	P5	P6	P7	P8	P9
Route 1	S-4	N-4	S-4	S-4	S-4	S-4	N-4	S-4	S-4
Route 2	S-4	N-4	N-3	S-4	S-4	S-4	0	N-4	S-4

Table 5.3: Experiment 3 quantitative results: route 1 is the path from the North exit to the cube D of room 09, and route 2 is the reversed path. *Notes: the N-prefix means that the route is along the North-edge corridor, and the S-prefix indicates the route is along the South-edge.*

It is no doubt that mentally following a path in an unfamiliar environment is not a trivial task. Although the participants were accustomed to the maps at this moment, some might need an extra time to reason their orientation and to determine the next navigation step. For example, the participants P3, P5, P7 initially lost their direction due to failure in self-registering their body's orientation onto the scene, resulting in a lack of confidence when making a turn. However, after mentally placing themself onto the path, these participants were able to carry out the task and to identify all features along the way. The participant P7 did not complete the second path of this experiment due to her tiredness. The other participants (P1, P2, P4, P6, P8, P9) performed extremely well on path traveling and showed no trouble or difficulty.

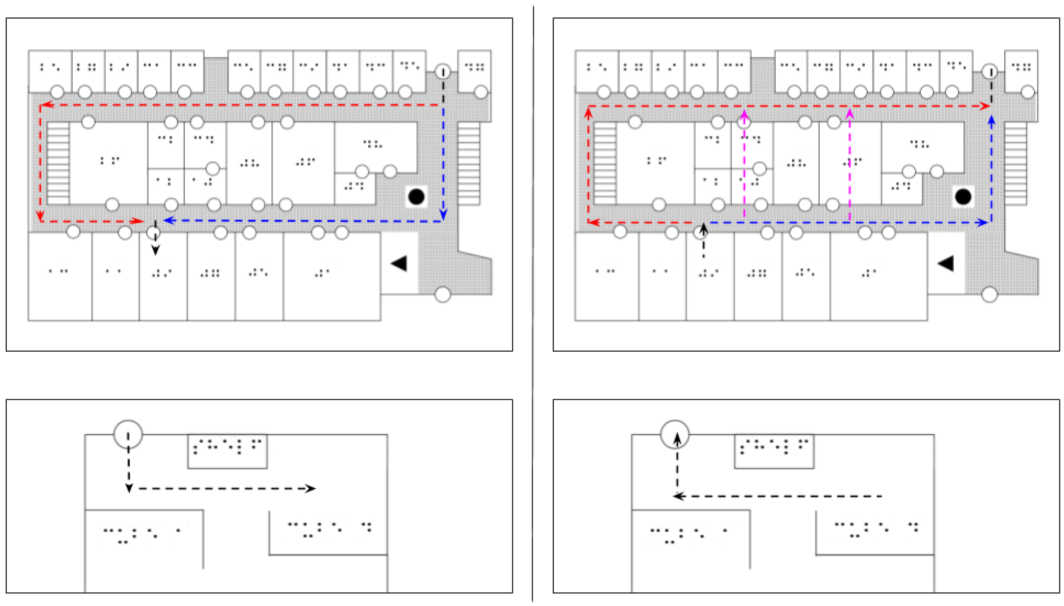


Figure 5.4: Possible conducted routes in the path finding experiment. The left column denotes paths conducted in part 1 of the experiment: starting from the North exit of the building and ending at the cube D of room 09. The right column shows paths for completing part 2: starting from cube D of room 09, and ending at the North exit. Notes: the red lines indicates paths along the North-edge corridor, the blue lines represents paths along the South-edge corridor, and the violet lines are for paths passing through interconnecting rooms.

Paths to follow in this experiment are generally categorized into: South-edge route and North-edge route (figure 5.4). For the first task of arriving at the cube D of room 09 from the North exit, all participants followed the South-edge route, except P2 and P7. For the second task of navigating to the North exit from cube D of room 09 and given that participants can travel through interconnecting rooms, most of them still chose to follow the South-edge route which does not cut through any room, except

P2, P3, and P8. Specifically, the participants P2 and P8 were able to determine the path through the rooms 10 and 34, while the participant P3 conducted a route passing through room 6 that directly connecting the South and North-edge corridor.

5.2.2.4 Exiting Survey

Table 5.4 summarizes the outcome of the exiting survey. It is clear that maps generated by my systematic tool are easy to read, their embossed symbols are easy to identify, and finding a certain room on such maps is also an easy task. However, the survey scores also show that following a specific path that spans multiple map scales is not trivial. It requires the participants to mentally register their self-orientation while constantly imagine to walk in that environment.

P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**
Q1: The Full map was easy to read								
3	5	5	5	5	5	5	5	5
Q2: The Room 09 map was easy to read								
4	5	5	5	5	4	5	5	5
Q3: The symbols were easy to identify								
5	4	4	5	5	5	5	5	5
Q4: Finding a certain room in the Full map was easy								
3	5	5	3	5	5	5	5	5
Q5: Finding a path in the map was easy								
4	5	4	4	3	5	5	5	5
Q6: I feel that, when imagining to follow a path in the map, I was able to correctly identify the orientation of my body at all times								
5	5	3	3	5	5	5	5	5
Q7: When imagining to follow a path with a transition from the Full map to the Room 09 map, this transition was easy for me to follow								
4	5	4	4	5	5	4	5	5

Table 5.4: Exiting survey outcome.

Answers to the open-ended questions at the end of the survey brought to a number of interesting issues. In general, all participants agreed that the generated tactile maps do give out the general layout of a building immediately, giving them the idea of how rooms are placing next to each other, and how to go from one room to another. They also believed that such maps can be useful for the blind and visually impaired travelers to consult before an actual travel. One major drawback is that the maps are too large to be carried around or explored while actually walking. The participant P2 suggested that the elevator symbol should be embossed to point up/down, rather than left/right as currently, to indicate the moving direction of it. The participant P6 commented that most of the tactile maps serve as points of reference since the scale does not match exactly with the real world, in terms of how many steps he has to take or how far from one point to another point. Because of this, he would only use it before a travel rather than during the travel. Regarding Braille annotation on the maps, the participant P9 realized that it would be fine for those reading Braille Grade 1, however, it is ambiguous for the Grade 2 readers. Such Braille writings embossed on the tactile maps should clearly have a number sign or a letter sign appropriately.

5.3 Discussion

This user study has confirmed that access to spatial information represented in the multi-scale tactile maps generated by my systematic tool could lead to increased spatial awareness, well-maintained self-orientation, and efficient navigation from one place to another. The users were able to understand the general layout of an area inside a building, as well as locating fixtures (e.g., staircase, elevator, entrance, etc) or furnitures (i.e, cubicles, tables, shelves) appearing within. To a certain degree, these tactile maps are proven to be useful in increasing the spatial knowledge acquisition. Once available, the users can simply *"study the map and record the description before an actual traveling, [they] might not look at the map all the time but if [they] lost, [they] will take it out to consult"*. Also, the users *"can find locations of different places in [their] head at home"*, and understand *"the general layout of the place, and how to go from one place to another"*.

The survey and final discussion at the end of this user study has shed some light on how to improve the automation tactile maps making for VI travelers, especially for an indoor navigation. Tactile maps generated by my toolboxes should be accompanied by a verbal description to describe the distance the maps are trying to convey. For example, how long or how wide the building is; or how far away from one point to another point is. In addition, tactile maps at all scales would need to be more compact while still preserving the current density of details being represented, aiming to increate the maps' mobility. Last, tactile graphics representing the vertical transportation between floors inside a building should be expressed more explicitly. For example, the staircase symbol should denote the walk-up and walk-down position, and the elevator symbol should be embossed to point up and down instead of being left and right.

Chapter 6

Conclusion

In this dissertation, I introduce an end-to-end systematic tool for a collective spatial mapping of an indoor environment consisting of small-scale features (i.e., building fixtures, room furnitures, types of floor-covering, etc), and for the production of tactile maps at multiple scales, serving the purpose of pre-journey spatial knowledge acquisition. This systematic application, named Semantic Interior Mapology or SIM, is a composition of three independent toolbox: Map Conversion Tool, Map Population Tool, and Map Authoring Tool.

In the first part, I present the novel Map Conversion Tool that allows one to quickly and accurately trace the layout of a floor plan, produces a vectorized map stored in the *sim* format, that is amenable to interactive visualization. This floor plan tracing tool organizes the collected spatial information into spaces. Each space is characterized by a set of wall corners and possibly entrance corners, where pairs of adjacent wall corners may or may not be joined by a wall.

Next, I describe the Map Population Tool that enables segmentation of the visible surfaces into objects of interest, such as furnitures, and then populates small-scale items that are not present in the original floor plan. Rather than manually measuring the size and location of these items, users can simply take a 3D scan with a RGB-D camera, and easily segment out individual objects using the toolkit. The 3D mesh is rectified and registered with the spatial representation of a floor plan so that the objects are automatically placed in their correct location on the map. Both Map Conversion Tool and Map Population Tool contain a converter from its current structure into GeoJSON, a popular format for representing spatial information. To visualize a floor plan's 3D map view, I employ the MapboxGL JS engine which renders geodetic features stored in a GeoJSON file as extruded 3D objects, which can be accessed and interacted with from a regular web browser.

While tool such as SIM can help encoding a building's layout at the desired resolution for a later embossing, the generalization problem still remains. Map makers still have to decide which features can be part of the map, and which can be removed, lest the map become too crowded, and thus difficult to read. In this work, I and my advisor, Dr. Roberto Manduchi conducted a focus group in order to obtain the general information on what the adequate scale is, and thus the adequate level of detail, at which a map of a building should be embossed. Findings from such focus group discussion suggested that embossing maps of a building at different levels of detail could provide useful spatial information. This has inspired the development of Map Authoring Tool that produces a digital tactile map file on demand, based on the building's structure represented in sim and the semantic layout of an interior space encoded in a JSON map that are collectively acquired using the previous toolboxes. Such map file is amenable for embossing at the scale specified by the user, with specific constraints on the density and distances of tactile features.

In order to understand the perceived utility of the multi-scale embossed tactile maps generated from SIM, we also performed an experimental study to evaluate whether access to spatial information represented in our maps could lead to increased spatial awareness, well-maintained self-orientation, and more confident indoor navigation. As a result, the participants all agree that *"the maps are helpful when traveling"*, and *"looking at the maps give me the general layout of the place, and how one room connects to another, or how to go from one room to another"*. In addition, the open-ended discussion at the end of the experiment shed some light on the promising improvement for my system. Specifically, the Map Authoring Tool should embed a verbal description to describe the spatial distance the maps are representing. For instance, how many steps does it take from the water fountain to the South entrance, and in which direction. Also, tactile symbols featuring vertical transportation elements such as staircase and elevator should indicate the operation direction(e.g., which end of the staircase users should take to walk up).

Generally, my end-to-end systematic tool is able to collectively encode the semantic spatial information of an indoor environment in great details, at multiple desired resolution, and produces cutting-edge embossed tactile maps that promotes spatial awareness for visually impaired travelers prior to a journey. I believe that my system, SIM, is an innovative tool for representing the spatial structure of an indoor environment at different modalities, and its simplicity of use may appeal to both practitioners and casual users. It is able to produce accurate results even in the case of complex building layouts. SIM is available for anyone to use at `https://sim.soe.ucsc.edu`,

Bibliography

[1] Sheraz Ahmed et al. "Automatic room detection and room labeling from architectural floor plans". In: *2012 10th IAPR International Workshop on Document Analysis Systems*. IEEE. 2012, pp. 339–343.

[2] Dragan Ahmetovic et al. "NavCog: turn-by-turn smartphone navigation assistant for people with visual impairments or blindness". In: *Proceedings of the 13th Web for All Conference*. ACM. 2016, p. 9.

[3] Nacny Amick, Jane Corcoran, et al. "Guidelines for the design of tactile graphics". In: *American Printing House for the Blind* (1997).

[4] Ferdinando Auricchio et al. "3D printing technology for buildings accessibility: the tactile map for MTE museum in Pavia". In: *J. Civ. Eng. Archit* 11 (2017), pp. 736–747.

[5] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. "Segnet: A deep convolutional encoder-decoder architecture for image segmentation". In: *IEEE transactions on pattern analysis and machine intelligence* 39.12 (2017), pp. 2481–2495.

[6] Mark Blades, Simon Ungar, and Christopher Spencer. "Map use by adults with visual impairments". In: *The Professional Geographer* 51.4 (1999), pp. 539–553.

[7] Anke Brock et al. "Usage of multimodal maps for blind people: why and how". In: *ACM international conference on interactive tabletops and surfaces*. 2010, pp. 247–248.

[8] Howard Butler et al. *The geojson format*. Tech. rep. 2016.

[9] Petr Červenka et al. "Blind friendly maps: Tactile maps for the blind as a part of the public map portal (mapy. cz)". In: *arXiv preprint arXiv:1603.09520* (2016).

[10] Liang-Chieh Chen et al. “Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs”. In: *IEEE transactions on pattern analysis and machine intelligence* 40.4 (2018), pp. 834–848.

[11] Lien-Wu Chen and Jun-Xian Liu. “EasyGO: A Rapid Indoor Navigation and Evacuation System Using Smartphones through Internet of Things Technologies”. In: *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. ACM. 2018, pp. 835–837.

[12] Angela Dai et al. “Scannet: Richly-annotated 3D reconstructions of indoor scenes”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 5828–5839.

[13] Luciene S Delazari et al. “Mapping Indoor Environments: Challenges Related to the Cartographic Representation and Routes”. In: *Geographical and Fingerprinting Data to Create Systems for Indoor Positioning and Indoor/Outdoor Navigation*. Elsevier, 2019, pp. 169–186.

[14] Samuel Dodge, Jiu Xu, and Björn Stenger. “Parsing floor plan images”. In: *2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA)*. IEEE. 2017, pp. 358–361.

[15] Jeff Donahue et al. “Decaf: A deep convolutional activation feature for generic visual recognition”. In: *International conference on machine learning*. 2014, pp. 647–655.

[16] Philippe Dosch and Gérald Masini. “Reconstruction of the 3D structure of a building from the 2D drawings of its floors”. In: *Proceedings of the Fifth International Conference on Document Analysis and Recognition. ICDAR’99 (Cat. No. PR00318)*. IEEE. 1999, pp. 487–490.

[17] Anan Du et al. “Exploring Long-Short-Term Context For Point Cloud Semantic Segmentation”. In: *2020 IEEE International Conference on Image Processing (ICIP)*. IEEE. 2020, pp. 2755–2759.

[18] Polly Edman. *Tactile graphics*. American Foundation for the Blind, 1992.

[19] Alistair DN Edwards, Nazatul Naquiah Abd Hamid, and Helen Petrie. “Exploring map orientation with interactive audio-tactile maps”. In: *IFIP Conference on Human-Computer Interaction*. Springer. 2015, pp. 72–79.

[20] Frédéric Evennou and François Marx. “Advanced integration of WiFi and inertial navigation systems for indoor mobile positioning”. In: *Eurasip journal on applied signal processing* 2006 (2006), pp. 164–164.

[21] Pedro F Felzenszwalb and Daniel P Huttenlocher. “Efficient graph-based image segmentation”. In: *International journal of computer vision* 59.2 (2004), pp. 167–181.

[22] Yu Feng, Frank Thiemann, and Monika Sester. “Learning cartographic building generalization with deep convolutional neural networks”. In: *ISPRS International Journal of Geo-Information* 8.6 (2019), p. 258.

[23] Martin A Fischler and Robert C Bolles. “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography”. In: *Communications of the ACM* 24.6 (1981), pp. 381–395.

[24] Aura Ganz et al. “PERCEPT: Indoor navigation for the blind and visually impaired”. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2011, pp. 856–859.

[25] Ruipeng Gao et al. “Fast and Resilient Indoor Floor Plan Construction with a Single User”. In: *IEEE Transactions on Mobile Computing* (2018).

[26] Ruipeng Gao et al. “Multi-story indoor floor plan reconstruction via mobile crowdsensing”. In: *IEEE Transactions on Mobile Computing* 15.6 (2016), pp. 1427–1442.

[27] Lucile Gimenez et al. “Automatic reconstruction of 3D building models from scanned 2D floor plans”. In: *Automation in Construction* 63 (2016), pp. 48–56.

[28] Ross Girshick et al. “Rich feature hierarchies for accurate object detection and semantic segmentation”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014, pp. 580–587.

[29] Nicholas A Giudice et al. “Learning non-visual graphical information using a touch-based vibro-audio interface”. In: *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*. ACM. 2012, pp. 103–110.

[30] Cole Gleason et al. "VizMap: Accessible visual information through crowdsourced map reconstruction". In: *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM. 2016, pp. 273–274.

[31] Xufeng Han et al. "Matchnet: Unifying feature and metric learning for patch-based matching". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015, pp. 3279–3286.

[32] Liang He et al. "TacTILE: a preliminary toolchain for creating accessible graphics with 3D-printed overlays and auditory annotations". In: *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. 2017, pp. 397–398.

[33] Lluís-Pere de las Heras et al. "Unsupervised wall detector in architectural floor plans". In: *2013 12th International Conference on Document Analysis and Recognition*. IEEE. 2013, pp. 1245–1249.

[34] Sepp Hochreiter and Jürgen Schmidhuber. "Long short-term memory". In: *Neural computation* 9.8 (1997), pp. 1735–1780.

[35] Emily Holmes, Gunnar Jansson, and Anders Jansson. "Exploring auditorily enhanced tactile maps for travel in new environments". In: *COLLOQUES-INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE COLLOQUES ET SEMINAIRES* (1996), pp. 191–196.

[36] Mihail Ivanchev, Francis Zinke, and Ulrike Lucke. "Pre-journey visualization of travel routes for the blind on refreshable interactive tactile displays". In: *International Conference on Computers for Handicapped Persons*. Springer. 2014, pp. 81–88.

[37] Hanme Jang, Jong Hyeon Yang, and Yu Kiyun. "Automatic Wall Detection and Building Topology and Property of 2D Floor Plan". In: *10th International Conference on Geographic Information Science (GIScience 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik. 2018.

[38] US DEPARTAMENT OF JUSTICE. "Guidance onADA Standards for accessible design". In: *Available in: http://www. ada. gov/regs2010/2010ADAStandards* (2010).

[39] Yukitoshi Kashimoto, Yutaka Arakawa, and Kiichi Yasumoto. "A floor plan creation tool utilizing a smartphone with an ultrasonic sensor gadget". In: *2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE. 2016, pp. 131–136.

[40] Vladimir Kulyukin and John Nicholson. "Toward blind travel support through verbal route directions: A path inference algorithm for inferring new route descriptions from existing route descriptions". In: *The Open Rehabilitation Journal* 5.22 (2012), e40.

[41] Yann LeCun et al. "Gradient-based learning applied to document recognition". In: *Proceedings of the IEEE* 86.11 (1998), pp. 2278–2324.

[42] Cheng-Lung Lee. "An evaluation of tactile symbols in public environment for the visually impaired". In: *Applied ergonomics* 75 (2019), pp. 193–200.

[43] Gordon E Legge et al. "Retention of high tactile acuity throughout the life span in blindness". In: *Perception & psychophysics* 70.8 (2008), pp. 1471–1488.

[44] Daniele Leonardis, Loconsole Claudio, and Antonio Frisoli. "A survey on innovative refreshable braille display technologies". In: *International Conference on Applied Human Factors and Ergonomics*. Springer. 2017, pp. 488–498.

[45] Rick Lewis and Carlo Séquin. "Generation of 3D building models from 2D architectural plans". In: *Computer-Aided Design* 30.10 (1998), pp. 765–779.

[46] Yangyan Li et al. "PointCNN: Convolution on X-transformed points". In: *Advances in Neural Information Processing Systems*. 2018, pp. 820–830.

[47] Qing Lin and Youngjoon Han. "A context-aware-based audio guidance system for blind people using a multimodal profile model". In: *Sensors* 14.10 (2014), pp. 18670–18700.

[48] Chen Liu, Jiaye Wu, and Yasutaka Furukawa. "Floornet: A unified framework for floorplan reconstruction from 3D scans". In: *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018, pp. 201–217.

[49] Chen Liu et al. "Raster-to-vector: Revisiting floorplan transformation". In: *Proceedings of the IEEE International Conference on Computer Vision*. 2017, pp. 2195–2203.

[50] Amy Lobben and Megan Lawrence. "The use of environmental features on tactile maps by navigators who are blind". In: *The Professional Geographer* 64.1 (2012), pp. 95–108.

[51] Jonathan Long, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 3431–3440.

[52] Jack M Loomis, Roberta L Klatzky, and Nicholas A Giudice. "Representing 3D space in working memory: Spatial images from vision, hearing, touch, and language". In: *Multisensory imagery*. Springer, 2013, pp. 131–155.

[53] Yanni Ma et al. "Global Context Reasoning for Semantic Segmentation of 3D Point Clouds". In: *The IEEE Winter Conference on Applications of Computer Vision*. 2020, pp. 2931–2940.

[54] Sébastien Macé et al. "A system to detect rooms in architectural floor plan images". In: *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems*. ACM. 2010, pp. 167–174.

[55] Alan M MacEachren. *How maps work: representation, visualization, and design*. Guilford Press, 2004.

[56] Anuradha Madugalla, Kim Marriott, and Simone Marinai. "Partitioning open plan areas in floor plans". In: *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*. Vol. 1. IEEE. 2017, pp. 47–52.

[57] Roberto Manduchi. "Mobile vision as assistive technology for the blind: An experimental study". In: *International Conference on Computers for Handicapped Persons*. Springer. 2012, pp. 9–16.

[58] Francisco Martínez, Antonio Jesús Rueda, and Francisco Ramón Feito. "A new algorithm for computing Boolean operations on polygons". In: *Computers & Geosciences* 35.6 (2009), pp. 1177–1185.

[59] Daniel Maturana and Sebastian Scherer. "Voxnet: A 3d convolutional neural network for real-time object recognition". In: *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2015, pp. 922–928.

[60] John McCormac et al. “Semanticfusion: Dense 3d semantic mapping with convolutional neural networks”. In: *2017 IEEE International Conference on Robotics and automation (ICRA)*. IEEE. 2017, pp. 4628–4635.

[61] Joshua A Miele, Steven Landau, and Deborah Gilden. “Talking TMAP: Automated generation of audio-tactile maps using Smith-Kettlewell’s TMAP software”. In: *British Journal of Visual Impairment* 24.2 (2006), pp. 93–100.

[62] Joshua A Miele and JR Marston. “Tactile map automated production (TMAP): On-demand accessible street maps for blind and visually impaired travelers”. In: *American Association of Geographers 101st Annual Meeting, Denver, CO*. 2005.

[63] Mark Monmonier. *How to lie with maps*. University of Chicago Press, 2018.

[64] Lynn Nadel and Lloyd MacDonald. “Hippocampus: Cognitive map or working memory?” In: *Behavioral and neural biology* 29.3 (1980), pp. 405–409.

[65] Duc Thanh Nguyen et al. “A robust 3d-2d interactive tool for scene segmentation and annotation”. In: *IEEE transactions on visualization and computer graphics* 24.12 (2018), pp. 3005–3018.

[66] Michał Nowicki. “WiFi-guided visual loop closure for indoor navigation using mobile devices”. In: *Journal of Automation Mobile Robotics and Intelligent Systems* 8.3 (2014), pp. 10–18.

[67] Guanghui Pan, Jia He, and Rui Fang. “Automatic floor plan detection and recognition”. In: *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*. IEEE. 2017, pp. 201–205.

[68] Konstantinos Papadopoulos, Marialena Barouti, and Konstantinos Charitakis. “A university indoors audio-tactile mobility aid for individuals with blindness”. In: *International Conference on Computers for Handicapped Persons*. Springer. 2014, pp. 108–115.

[69] Giovanni Pintore et al. “3D floor plan recovery from overlapping spherical images”. In: *Computational Visual Media* 4.4 (2018), pp. 367–383.

[70] Denise Prescher, Jens Bornschein, and Gerhard Weber. “Consistency of a Tactile Pattern Set”. In: *ACM Transactions on Accessible Computing (TACCESS)* 10.2 (2017), p. 7.

[71] Charles R Qi et al. “Pointnet: Deep learning on point sets for 3d classification and segmentation”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 652–660.

[72] Charles R Qi et al. “Volumetric and multi-view cnns for object classification on 3d data”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 5648–5656.

[73] Xiaojuan Qi et al. “3d graph neural networks for rgbd semantic segmentation”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2017, pp. 5199–5208.

[74] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-net: Convolutional networks for biomedical image segmentation”. In: *International Conference on Medical image computing and computer-assisted intervention*. Springer. 2015, pp. 234–241.

[75] Jonathan Rowell. “The end of tactile mapping or a new beginning: LBS for visually impaired people”. In: *Papers. XXIII International Cartographic Conference*. 2007, pp. 4–10.

[76] Monika Sester, Yu Feng, and F Thiemann. “Building generalization using deep learning”. In: *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-4 (2018)* 42 (2018), pp. 565–572.

[77] Shuran Song et al. “Semantic scene completion from a single depth image”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 1746–1754.

[78] Radim Štampach and Eva Mulíčková. “Automated generation of tactile maps”. In: *Journal of Maps* 12.sup1 (2016), pp. 532–540.

[79] Hang Su et al. “Multi-view convolutional neural networks for 3d shape recognition”. In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 945–953.

[80] Ivan E Sutherland. “Three-dimensional data input by tablet”. In: *Proceedings of the IEEE* 62.4 (1974), pp. 453–461.

[81] Hao Tang et al. "Automatic pre-journey indoor map generation using autocad floor plan". In: (2016).

[82] Guillaume Touya, Xiang Zhang, and Imran Lokhat. "Is deep learning the new agent for map generalization?" In: *International Journal of Cartography* 5.2-3 (2019), pp. 142–157.

[83] Guillaume Touya et al. "Automatic derivation of on-demand tactile maps for visually impaired people: first experiments and research agenda". In: *International Journal of Cartography* 5.1 (2019), pp. 67–91.

[84] Julien Valentin et al. "Semanticpaint: Interactive 3d labeling and learning at your fingertips". In: *ACM Transactions on Graphics (TOG)* 34.5 (2015), p. 154.

[85] Tetsuya Watanabe et al. "Tactile map automated creation system using openstreetmap". In: *International Conference on Computers for Handicapped Persons*. Springer. 2014, pp. 42–49.

[86] Zhirong Wu et al. "3d shapenets: A deep representation for volumetric shapes". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 1912–1920.

[87] Xiaoyi Zhang et al. "Interactiles: 3D printed tactile interfaces to enhance mobile touchscreen accessibility". In: *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. 2018, pp. 131–142.

[88] J Zhao, X Zhang, and Y Wang. "Indoor 3d Point Clouds Semantic Segmentation Bases on Modified Pointnet Network". In: *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 43 (2020), pp. 369–373.

[89] Yuanqing Zheng et al. "Travi-navi: Self-deployable indoor navigation system". In: *IEEE/ACM transactions on networking* 25.5 (2017), pp. 2655–2669.

[90] Bing Zhou et al. "BatMapper: acoustic sensing based indoor floor plan construction using smartphones". In: *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*. ACM. 2017, pp. 42–55.