# Constructing Topological Maps of Displays with 3-D Positioning Information

Donald J. Patterson

University Of California, Irvine, USA
djp3@ics.uci.edu

**Abstract.** To better coordinate information displays with moving people and the environment, software must know the locations and *three dimensional alignments* of the display hardware. In this paper we describe a technique for creating such an enhanced topological map of networked public displays using a mobile phone. The result supports a richer user experience, without the cost of maintaining a high resolution reconstruction of a smart environment.

## 1 Introduction and Related Work

Digital displays are relied upon to provide location and situation specific information to a wide variety of viewers (*e.g.*, [1]). Despite their cost, power and maintenance requirements, they are often preferred to non-digital signage in high-traffic venues [2]. The presence of multiple networked digital displays creates a class of navigation applications which can coordinate the display of dynamic content in a way that is not possible with static or isolated digital signage. This class of applications does not require complete knowledge of the physical environment to be effective.

An example of such an application is a hospital that guides patients to the appropriate office. Imagine that Martha must visit a cardiac specialist for the first time. She indicates her high-level goal to a kiosk, "Appointment with Dr. Theophilus" She is requested to follow a yellow arrow which appears on a series of cooperating displays [3] that guide her to the correct office. Other patients are given different colors to follow. Martha can follow her arrows without needing to be familiar with the layout of the hospital, or knowing that Dr. Theophilus has a complicated schedule that has him seeing patients in several locations during the course of the day. Incorporating tracking technology [4] would support showing the yellow arrows only on displays that Martha can actually see.

To realize this scenario, displays must know their location and position relative to each other. Location alone is not sufficient because two displays in the same location would show an arrow pointing in opposite directions (for the same destination) depending on which side of the hallway they were mounted. An expensive infrastructure in which displays contain a digital compass, a 3-axis accelerometer, and in which the environment contains a fine-grained indoor localization technology [5] would be sufficient to address this problem. In this paper we demonstrate that this additional heavyweight capability is not required.

Müller and A. Krüger [4] have done similar work in learning device topologies. Their goal was to identify the distance between devices to enable displays to show sequential content to a moving person. They detected and centrally analyzed passing Bluetooth signals from cooperating subjects to calculate distance estimates between displays. Our work builds on theirs by focusing on developing techniques for orienting the devices to enable richer *spatial* interactions to complement their temporal interactions. Our use of camera sensors provides useful additional spatial data for reducing errors in estimates that are caused by the limited spatial granularity at which Bluetooth devices can be detected.

## 2 System Overview

Our system requires a single cooperating administrator, using a cell-phone camera, to walk the entire floor plan of the building while the system displays unique 2-D barcodes on each display for calibration and identification. The camera is mounted in a small "image splitter" bracket containing two mirrors, angled at 45 degrees, that are aligned to the lens of the camera. The mirrors reflect images from both sides of the hallway allowing synchronized information from two streams of video to be leveraged during analysis (see Fig. 1). Utilizing techniques (from augmented reality research) that orient barcodes in space, the timing between frames, and the synchronized observations of opposite sides of a hallway, we are able to create accurate building topologies including display location and orientation. We evaluated our approach on simulated and real-world floor plans.

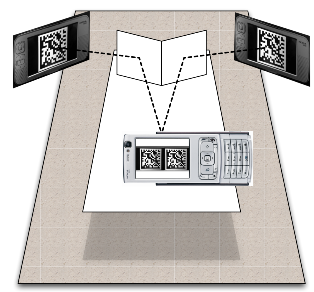

**Fig. 1.** A user's view of two mounted displays as seen through the camera and image splitter.

Determining a barcode's orientation from an image is a well understood problem in augmented reality [6] when the following two reasonable assumptions are made: First, the image splitter is being held parallel to the floor. Second, the display is mounted flush on the wall, parallel to the floor, so that the orientation of the barcode matches the orientation of the wall, allowing us to relax the assumption of orthogonal walls.

To test our system we deployed several Nokia N800's as displays. We used a Nokia N95 cell phone for video capture (640x480, 30fps) and sent the data wirelessly to a server for real-time analysis. We converted the dual video streams into time-stamped observations indicating the absence or identity of a barcode in both channels as well as the 3D orientation of the barcode. A remote server subsequently constructed candidate topologies of the floor plan that were consistent with the observations and ranked them according to a scoring model.

Alternatively, we could have used RFID tags or Bluetooth beacons to identify and calibrate the displays. However, they do not provide self-orientation nor directionality information between the reader and the tag/beacon. They also incur additional complexity and cost. In contrast, 2-D barcodes work well for calibration; they provide orientation information, camera line-of-sight provides directionality, they can be *temporarily* shown on displays during calibration, they are feasible for large deployments and id's can be dynamically assigned.

## 3 Topology Reconstruction Algorithm

To reconstruct a representation of the building with displays positioned on the topology, we present a generate-and-score style algorithm that is run in real-time as an administrator walks through the building. This algorithm maintains a list of physically feasible topologies that is updated with each new display observation. The list is ranked according to Occam's razor such that the simplest floor plan, consistent with the observations, is preferred. We assume that floor plans can be represented as planar graphs and that the administrator completes a full walk-through of the area and observes every display.

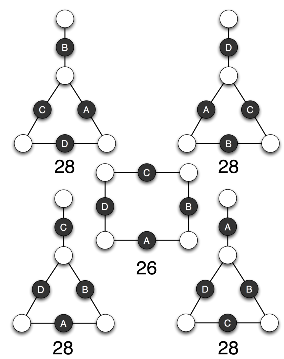


**Fig. 2.** Candidate topologies made from observing (A,B,C,D,A)

We represent topologies as a *display connectivity graph* (DCG), an undirected graph in which displays, intersections and dead-ends are represented as nodes (black, white and white respectively), and edges represent physical adjacency (roughly hallways). Display nodes contain the 3-D orientation as meta-information (See Figure 2).

***Scoring Model*** During a walk-through, when a new display is observed we inductively generate all possible candidate graphs from the graphs previously under consideration. After removing non-planar and isomorphic graphs, we score the graphs according to floor plan complexity and likely user walking behavior (see equation 1).

$$S = \alpha_1 \sum_N degree(N)^2 + \alpha_2 * |loops| + \alpha_3 * bktracks \qquad (1)$$

Our scoring function is a weighted sum of two structural penalties and one behavior penalty. The first term is a sum of the squares of the degree of the structure nodes, $N$. This term favors simpler intersections as explanations for data. The second term penalizes graphs which contain more loops. The final term is the number of backtracks that the user must have made during the walk-through. (This corresponds to leaving a structure node by the same edge by which it was entered when the degree of the structure node $> 1$). In our experiments we used weighting factors of $\alpha_1 = 1$, $\alpha_2 = 10$, and $\alpha_3 = 1$. Better topologies receive lower scores. Figure 2 shows the top 5 possible graphs after observing four displays in sequence (A, B, C, D, A).

## 4 Analysis

The first analysis we present assumes the algorithm only has knowlege of the *presence* of a display as would be the case if the displays were augmented with

RFID or Bluetooth. Later we include the additional orientation information that is afforded by our visual system.

***Simulated Floor Plan Results*** We generated 25 physically possible DCGs and a simulated walk-through as test cases.

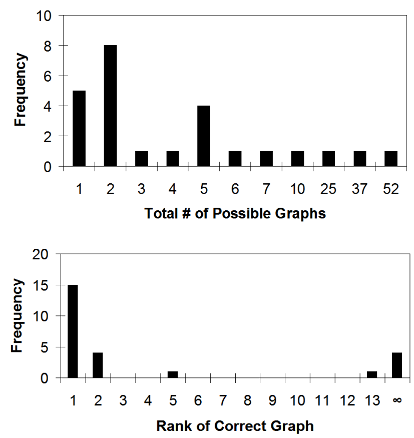


**Fig. 3.** Top: Histogram showing how many possible floor plans were generated for each of the 25 simulations. Bottom: Histogram of the rank of the true floor plan among all those generated and scored.

We measured the number of candidate graphs the algorithm created during simulated walk-throughs (see Figure 3-top).

Figure 3-bottom shows that 80% of the time the true floor plan was ranked first or second. In four cases the observations were not sufficient to close a loop in the true floor plan, so the algorithm did not have enough information to generate the correct answer. For example, in Figure 2, none of the floor plans would be generated if the walk-through observed only {A, B, C, D}. This sequence does not capture the link from D to A, and so, none of the generated floor plans would show that connection either. A "complete" walk-through is therefore necessary for optimal results, although sub-optimal results will support our motivating scenario but not produce the shortest path through the building for a user.

Removing the user model component from the scoring caused two of the correct graphs to be ranked lower.

***Video Augmented Results*** Our second analysis included orientation and dual-channel video as well. When a display is seen a second time, we compare the orientation to its previously observed state. If it is on the same side of the hallway, we know we have experienced a loop. If its orientation is reversed, we have turned around, and the previous display should be linked to the appropriate side. This is only possible because of our image splitter tool and choice of visual tags for landmarks. As an example, if we assumed that all displays in Figure 2 were on the outside of the hallways and seen on the left side of the camera splitter, the four graphs with score 28 are impossible.

Using the same simulated data, our analysis is now strongly constrained and only produces one viable, correct floor plan for each case.

***Real Floor Plans*** Finally we obtained the floor plans of 5 research institutions. We simulated the placement of displays around 4 of them and conducted a real analysis on the fifth. We asked two participants familiar with the buildings to provide us with paths through the space that would observe all the displays. We simulated 4 outcomes and empirically evaluated the fifth. The resulting number of potential graphs is reported in Table 1. In all five cases, the dual-video algorithm correctly identified the real-world topology.

## 5 Discussion and Future Work

Using split-screen video capture of 2D barcodes to locate displays provides a rapid, effective and low-cost way to calibrate a smart environment. The resulting display connectivity graph, augmented with 3D display orientation, supports rich user interface applications such as providing directional user navigation in a coordinated fashion.

| Building Name | # of Potential Graphs |
|---|---|
| Sieg (UW) | 4 |
| Microsoft(Seattle) | 6 |
| Allen (UW) | 12 |
| Bren Hall (UCI) | 32 |
| Intel (Seattle) | 100+ |

**Table 1.** Real-world buildings analyzed for complexity. The correct floor plan was identified in each case.

Although we clearly demonstrated the value of orientation as provided by the split-screen camera, there are improvements that should be made. First we assume 100% recognition of our displays. If a display is never seen, but known to be in the building, the model cannot provide any information about where the display might be located. Additionally, if a display is seen once, but later missed the floor plans will be sufficient but sub-optimal. A hybrid approach that mixed visual and RFID/Bluetooth beaconing might alleviate some of these challenges. [1]

## References


1. Huang, E.M., Koster, A., Borchers, J.: Overcoming assumptions and uncovering practices: When does the public really look at public displays? In Indulska, J., Patterson, D.J., Rodden, T., Ott, M., eds.: Pervasive. Volume 5013 of Lecture Notes in Computer Science., Springer (2008) 228–243
2. Cheverst, K., Dix, A.J., Fitton, D., Friday, A., Rouncefield, M.: Exploring the utility of remote messaging and situated office door displays. In Chittaro, L., ed.: Mobile HCI. Volume 2795 of Lecture Notes in Computer Science., Springer (2003) 336–341
3. Kray, C., Cheverst, K., Fitton, D., Sas, C., Patterson, J., Rouncefield, M., Stahl, C.: Sharing control of dispersed situated displays between nomadic and residential users. In Nieminen, M., Röykkee, M., eds.: Mobile HCI, New York, NY, USA, ACM (2006) 61–68
4. Müller, H.J., Krüger, A.: Learning topologies of situated public displays by observing implicit user interactions. In Stephanidis, C., ed.: HCI (6). Volume 4555 of Lecture Notes in Computer Science., Springer (2007) 158–167
5. Stuntebeck, E.P., Patel, S.N., Robertson, T., Reynolds, M.S., Abowd, G.D.: Wideband powerline positioning for indoor localization. In: UbiComp '08: Proceedings of the 10th international conference on Ubiquitous computing, New York, NY, USA, ACM (2008) 94–103
6. Zhou, F., Duh, H.B.L., Billinghurst, M.: Trends in augmented reality tracking, interaction and display: A review of ten years of ismar. Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on (Sept. 2008) 193–202



[1] This work was supported by NSF award: HCC:0713562