A Hybrid Indoor Navigation System

Andreas Butz, Jörg Baus, Antonio Krüger, Marco Lohse
University of Saarbrücken, Department of Computer Science
P.O. Box 15 11 50, 66041 Saarbrücken, Germany

ABSTRACT

We describe a hybrid building navigation system consisting of stationary information booths and a mobile communication infrastructure feeding small portable devices. The graphical presentations for both the booths and the mobile devices are generated from a common source and for the common task of way finding, but they use different techniques to convey possibly different subsets of the relevant information. The form of the presentations is depending on technical limitations of the output media, accuracy of location information, and cognitive restrictions of the user. We analyze what information needs to be conveyed, how limited resources influence the presentation of this information, and argue, that by generating all different presentations in a common framework, a consistent appearance across devices can be achieved and that the different device classes can complement each other in facilitating the navigation task.

Keywords:

hybrid user interfaces, navigation, resource adaptivity, user adaptivity

Introduction

Sue is rushing through the door of the huge office complex, into the empty lobby, finally in the dry, catching her breath. They managed to build 70 storey office mazes with talking lobbies, but still couldn't get hold of a simple thunderstorm. "Hello Mrs. Walker" says the building, while a small puddle is forming on the marble floor around her feet. "Can I take you to your 1PM appointment with Mr. Grey? You are well in time, so if there's anything else I can do for you..." "Coffee" she nods. "That'd be a bliss..." A holographic projection of a boy scout appears in front of her, smiles, and gestures to follow her. "So let me take you to our coffee machine first." After the first hot sip she's wondering what dubious mind of a designer had chosen the boy scout character to impersonate the building's virtual guide. But then again, she doesn't really care, because up in the hallways where there are no means to do holographic projections she will be spared of him anyway. Paper cup in her hand she follows the little guy to the elevator. "I'll send you to the 63rd floor. You'll just have to follow the arrows on your bicorder then" are his last words while the elevator is closing with a soft hiss. Upon entering the nested system of hallways, Sue takes out her bicorder and sees a red arrow appear on its display. She follows its direction which is changing every few corners. Just before she reaches Mr. Grey's office, she drops the empty cup into the tray marked on a small map on the bicorder screen. "You're well in time, Mrs. Walker" comes the voice from behind the desk. "Glad you found it!"

While they are starting their meeting inside the office, let's have a closer look at this imaginary office building of the future. Its underlying computer obviously uses various means to guide its visitors to their destination, and it does its best to do so in an adapted and unobtrusive way. The lobby, for example, was equipped with an expensive 3D projection system, while the hallways of the upper floors have to rely on simpler facilities, namely the visitor's own bicorder[1] which is fed with information by the hallway's ceiling lamp infrastructure. Now, living almost 30 years earlier, we have to put up with a comparatively small 3D graphics screen in the lobby and a little guy just running around this screen for now, but nevertheless we can get computers to design way descriptions to be conveyed by this guide and on the bicorder (which has to be a Palm Pilot for now) in a consistent way. In order to achieve smooth navigation through the whole building, presentations have to be designed coherently for different output media. The building's navigation system in fact is a hybrid one, distributed over different classes of devices, which in turn all serve the common navigation task.

Route descriptions

Up to now little research has been done on resource adaptive graphics/animation generation for different classes of presentation devices, obeying technical resources, but also taking into account the user's limited cognitive resources to decode and understand a presentation. As a domain for investigating the relation between spatial descriptions and their cognitively adequate visualization, we use the generation and presentation of multimedia route descriptions in a navigation task. Following [12] a simple model for the gen-

IUI'01 January 14-17, 2001, Sante Fe, New Mexico.

[1] The bicorder craze only really started after cellular phones came out of fashion in the late 20ies of the 21st century.

eration of route descriptions consists of three steps: The first step is the activation of a representation of spatial knowledge at the appropriate scale for the route. The second step consists of the choice of a specific route through the environment (depending in general on the mode of travel, the desired route characteristics and the user's presumed knowledge about the environment). The last step consists of a translation from the chosen route to a description, in our case a set of multimodal instructions suitable for different output devices.

Because text only descriptions are notoriously inadequate for expressing complex spatial relationships [15], our goal is to generate multimedia route descriptions for different presentation devices. Depending on the device the presentations range from a virtual walkthrough presented by a virtual human presenter (supported by spatial utterances and metagraphics to complement each other) to simple sketches of the environment or arrows indicating the direction.

A common depiction of routes are route maps. They show the route as a sequence of turning points connected by lines. In addition, maps convey survey knowledge about the environment, e.g., regions or the structure of the environment from a bird's eye view using an *allocentric* frame of reference. In contrast, verbal route descriptions consist of a sequential description of route segments, including physical elements and basic motor activities, e.g., walk, turn[2]. Verbal descriptions also differ from maps in that they describe the path of motion from an *egocentric* frame of reference, i.e. as seen by the navigating person.

Some verbal route descriptions mention regions and spatial relations between objects in the current environment. The addressee undertakes a mental journey, during which elements in the environment are localized in relation to her current position or to each other from an egocentric point of view. This *route perspective* is helpful to convey knowledge about path segments and landmarks, the so-called route knowledge ([16]).

Landmarks are elements of the environment external to the observer, serving to define the location of other objects or locations. They are memorable cues selected along a path and enable the encoding of spatial relations between objects and paths, leading to the development of a cognitive map of the environment. Landmarks are generally used in navigation tasks to identify decision or destination points or to convey route progress. They influence expectations, provide orientation cues for homing vectors and suggest regional differentiating features (see [13]).

Following [17], a path of motion can be divided into certain segments, each segment consisting at least of four parts belonging to different categories: *starting point, reorientation(orientation), path/progression* and *ending point*. Since the paths we describe are continuous, the ending point of one segment serves as the starting point for the next one. Segments are mostly separated by changes in direction, and the resulting presentation has to communicate a change in direction clearly.

The description of a path segment creates a partial view of the environment by integrating different kinds of knowledge. In our proposed model, knowledge about the possible paths through the environment is represented in a graph. Nodes of the graph correspond to turns or intersections in the real world. Edges correspond to connections between two points in the real world. Thus all paths in the graph can be divided into segments in a straightforward way.

In order to clearly describe or visualize path segments, we have to integrate knowledge about different types of landmarks from an annotated 3D model. Landmarks at decision points are needed to communicate a reorientation and/or path progression, and are located at turns along the path. Road landmarks are located along the path, but not at a specific decision point. The emphasis of potentially wrong choices at decision landmarks can be used to assure the addressee of being on the right way. Finally, we also integrate the addressee's own current position, which allows us to compute spatial relations between her and objects in the environment and to describe locations from her egocentric point of view. By integrating these different landmarks in the segment representation, we are able to compute spatial relations between objects in the scene for verbal route descriptions (see [8, 10]).

For the transfer of survey knowledge, information about regions or the structure of the environment has to be included and often another point of view is chosen. Elements of the scene are referred to in an allocentric frame of reference corresponding to a survey perspective of the environment (see [17, 18]). In this case we have to choose a suitable viewpoint to look at the scene, for example a bird's eye or top down view on a visualized map. Also, a certain amount of redundant information coding is generally useful to communicate route direction, as described in [9].

Adaptive planning of way descriptions

In order to generate structural graphical descriptions we extend an efficient hierarchical planning approach presented in [5] for the generation of 3D animation. The main assumption here is that all generated graphical presentations can be structured in the form of a tree describing parts and subparts of the graphics to a certain depth. Each part or subpart corresponds to a node in the tree. Nodes are either terminal nodes in which case they describe portions of the graphics that will be realized by one of the graphics realization techniques described in the section about graphical techniques, or they are nonterminal nodes, in which case they specify a set of subnodes and a logical, spatial or temporal interrelationship between them.

Temporal interrelationships only apply to temporal media and include the concepts *parallel*, *sequential* and *incremen-*

[2] In general it is unclear how to judge the quality of generated descriptions (see [12]), but in any case the descriptions should be easy to decode, understand and memorize, not to mention correct.

tal. An example for temporally *parallel* subparts of a graphics are, for example, a camera motion and an object motion taking place over the same timespan of a 3D animation. A *sequential* interrelationship describes a sequence of subparts taking place in a temporal order, e.g., one after each other. Specifying the subparts as having an *incremental* relationship means that after a subtree is fully expanded, this subtree can be forwarded to the graphics realization component, which is not the case with every subtree in a *sequential* list. The specification of incrementally ordered subsequences allows the graphics realization process to start its work before the structure of the graphics is fully generated, and thus greatly reduces the perceived delay from the start of the whole graphics generation process to the moment the first graphical element is shown.

Logical interrelationships include the concepts of *alternative*, *conditional* and *additional* subtrees in the structure of a graphical presentation. Both *alternative* and *conditional* subtrees specify a list of possibilities for the realization of a certain part of the graphics. *Conditional* expansion selects one alternative from this list at planning time. In this way we can, for example, specify that a certain type of graphical presentation can be generated either as a line drawing or as a 3D image of the scene depending on media restrictions. Specifying these subtrees as *alternative* postpones the decision until presentation time. The strategy here is to first expand the structurally simplest or computationally cheapest part of the tree and then – if time permits – to proceed with more complex alternatives that might be visually more appealing or clearer in the communicative sense. The resulting structure graph contains all of the various alternatives (unless planning was stopped before due to temporal restrictions) and leaves the decision which alternative is chosen to either the presentation process or even the user.

An example for an *additional* relationship between subtrees of the structure tree is the labeling of a line drawing or the creation of additional viewports for an already running 3D animation. This is assuming that already the first subtree would yield a 'working' graphical presentation, while the following subtrees contain presentation elements which will enhance the overall quality of the graphics. As it might have become obvious, the different kinds of interrelationships within the structure tree of a graphical presentation leave room for various strategies of adaptation of the generation process to limited resources either in the output medium or in the generation process.

Figure 1 shows a simplified part of a structure tree for a graphical way description. At the root of the tree we see that the way description can be presented in the form of a sketch or *alternatively* a 3D walkthrough. One of these alternatives can be chosen at presentation time based on circumstances, media restrictions or user preferences. In the case of a 3D-walkthrough the ground plan and the actual path have to be shown. Specifying these two actions as having an *incremental* interrelationship implies that the ground plan can already

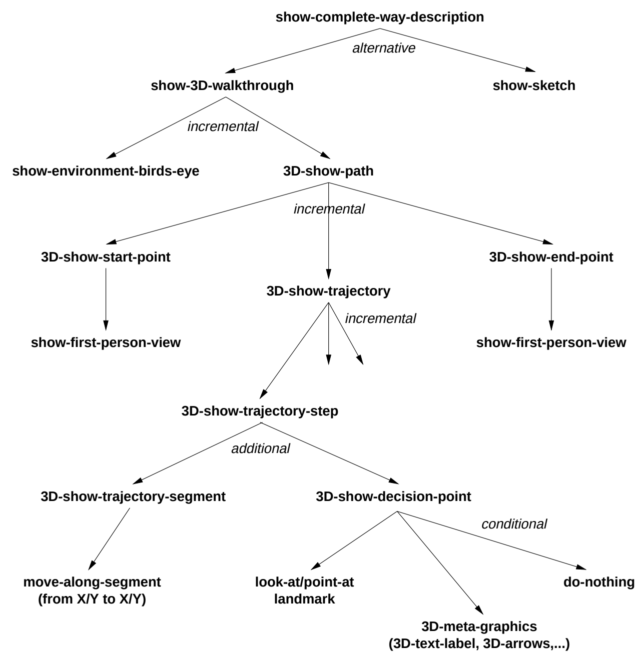

Figure 1: Simplified example of a structure tree for a graphical way description

be drawn by the realization component while planning for the path visualization continues. Had we specified the two actions as being *parallel* or *sequential*, planning would have to be finished before the realization starts.

The visualization of the path itself consists of showing its starting and ending point and the trajectory inbetween. Showing the trajectory in turn is nothing else than an incremental loop over all of the trajectory's segments (see section about route descriptions) and drawing an arrow for each of them. Optionally, after drawing each segment of the trajectory, the corresponding decision point can be shown. Depending on media restrictions this visualization can consist of a thumbnail image from the 3D world at this particular point, a text label or nothing at all.

Position, orientation and reorientation

A successful way description has to optimally support the addressee to make the right choice at every decision point. Based on knowledge about her exact *position* and *orientation*, the way description helps the user to *reorient* to the new direction and to proceed on her correct way.

In the case of a a static way description (e.g., a hand drawn sketch on a piece of paper) the user has to determine her position and orientation at every decision point herself. As mentioned before, good way descriptions support the user in this task by providing landmarks. In order to make the process of reorientation easier, one can observe that users turn the way description around until their egocentric frame of reference matches the reference frame of the map, which makes it much easier to interpret reorientation information in the way description (e.g. an arrow) correctly.

We argue that a navigation system and its user form a unit in the sense that the knowledge about the position and orientation of the user must be represented in the system. If the system knows about the exact position and orientation of the user it can provide reorientation information in a very simple manner, e.g., by means of an arrow. If in contrast the user's position and orientation are vague or missing, the system must provide information for the user to locate herself and to determine her actual orientation in space or design the user dialog in a way that helps the user to fill in the missing information. In a scenario where the quality of position and orientation information is known beforehand (e.g., inside buildings equipped with infrared transmitters), this decision can be made at planning time. In an outdoor scenario with varying GPS signal quality, the decision would have to be made at presentation time, which is reflected in the *conditional* and *alternative* constructs in our planning formalism described above. The next section will discuss how to classify the quality of location and orientation of the user.

Resolution of location and orientation measurements in buildings

In order to generate appropriate graphical presentations or interaction schemata to describe a navigation task, we need a metric for the quality of location and orientation information.

The *resolution of orientation* can simply be measured in degrees. E.g. a resolution of +/-45 degrees implies that the system cannot distinguish between two alternative choices within less than 90 degrees at a certain decision point. Different technical solutions offer different resolutions of orientation: long range magnetic tracking devices (e.g. a compass) are accurate by a few degrees of resolution. Their major disadvantages are the risk of magnetic interference and the fact that those devices have to be incorporated into the mobile information system. For solutions integrated into the environment, e.g., wall mounted infrared senders or beacons, the resolution depends on the sender's coverage area. Realistic resolutions for an infrared based system in buildings range from 20 to 180 degrees, depending on the form of transmitter and receiver. A more detailed discussion of our own approach can be found in [2, 6].

The *resolution of location* can be measured in meters and indicates the maximum deviation from the assumed position. The Global Positioning System (GPS) for example mostly works with a deviation from 1 to 50m, but only outside of buildings. Accurate radio bearing for mobile devices is technically difficult and costly and even small deviations can cause wrong implications about the position of the user (i.e. whether she is in front or behind a wall). This problem can be avoided by using optical media, such as infrared light, to track the user's position, of course at the cost of equipping the environment with special hardware. Other approaches try to determine the location with the help of computer vision [1], but they currently still lack a certain degree of accuracy and reliability. In the last section of this paper we will present

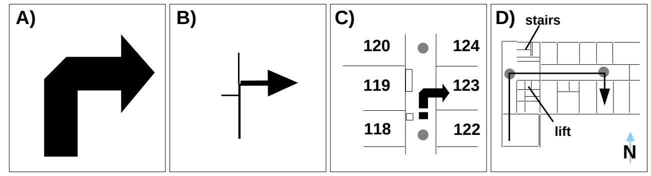

Figure 2: Four different graphical way description schemata that depend on the quality of orientation and position information.

our own solution based on infrared senders which implicitly track the user's position and orientation in a building as a side effect of broadcasting location specific data.

A third factor that is interweaved with the quality of position and location measurements is information coverage. Especially in buildings, certain areas will provide no information at all about position and location. Fortunately this is not critical if the general coverage at decision points is acceptable. Nevertheless, larger untracked areas during the navigation have to be compensated for by appropriate graphical presentations.

One important observation is that for the task of way finding, a high resolution of orientation and location information is not always necessary. Under the assumption that a user moves on a segment from one decision point to the next without changing direction, it is sufficient for the system to distinguish between these two decision points. The same holds for the resolution of orientation. It can be considered sufficient, if the system can distinguish between all the choices at a given decision point. E.g., at a T-junction the resolution of +/-45 degrees is good enough to reorient the user correctly to the new direction.

The next section will address different graphical presentation schemata that take into account the quality of position and orientation information. Furthermore in unfavorable situations of low information quality, simple user interactions are discussed that enable the user to fill in the missing information.

Graphical presentation schemata for way descriptions

We distinguish different types of situations with respect to the quality of position and orientation:

1. Sufficient position and orientation information
2. Sufficient position but insufficient orientation
3. Insufficient position and orientation information

There are also some interesting intermediate states, such as good orientation and mean position information or rough position but no orientation information.

How do these different situations cause changes in the content and style of the graphical way descriptions? In the first case the system knows enough about the actual decision point and orientation of the user to produce a simple reorientation instruction (e.g. an arrow as in figure 2A). If the quality of the orientation information decreases and the system can't exactly tell where the user is looking at, a simple arrow could mislead the user. Therefore additional information about the choices at the decision point has to be provided. Figure 2B shows such a graphical way description for an orientation resolution of +/-90 degrees. The topological diagram includes only the different choices at the current decision point, but doesn't show any additional landmarks. Please note that the map can still be roughly aligned to the user's walking direction to simplify her reorientation.

Landmarks have to be included as the quality of orientation and position information declines further. Figure 2C shows a description where the position resolution covers three potential decision points (two are indicated as grey dots). In such situations a purely topological map could cause problems and therefore an appropriately clipped area of the surrounding (here: the adjacent rooms with numbers and parts of the hallway, pillars and a locker) have to be displayed. By clicking on the grey dots the user can inform the system about her actual position and resolve the ambiguity of location, thus allowing the system to switch back to the topological presentation of figure 2B.

In the worst case there is only very rough or no information about the actual position and orientation and the system cannot align the map to the user's actual walking direction anymore. Now a greater portion of the map has to be chosen that may include several (especially already passed) turns of the user (see figure 2D). Instead of including small landmarks that are only relevant at a single decision point, global landmarks, such as stairs or elevators have to be represented in the presentation. Since it is important to explain to the user that she can't rely on the orientation of the map, the presentation contains a North Symbol to underline the external frame of reference. Again the user can communicate her position to the system by clicking on the grey dots, resulting in a closeup of that area of the building. But in order to align the map to the walking direction, the system has to ensure the user's correct orientation. This task can be accomplished by advising the user to reorient herself towards a landmark (e.g. by prompting a text: "Turn around until the stairs are to your left and the lift is to your right").

The proposed interactions during the way finding task are of course only meaningful if the user is using a mobile device. The schema used at a stationary booth will in most cases be similar to the example in figure 2D. On the other hand, as explained shortly, the system will be able to use more sophisticated techniques, such as animations or worlds in miniature[14, 7] to explain the navigation task.

A Framework for Hybrid Navigation Systems

We will now present a framework for hybrid navigation systems presented in UML notation (see figure 3) and explained

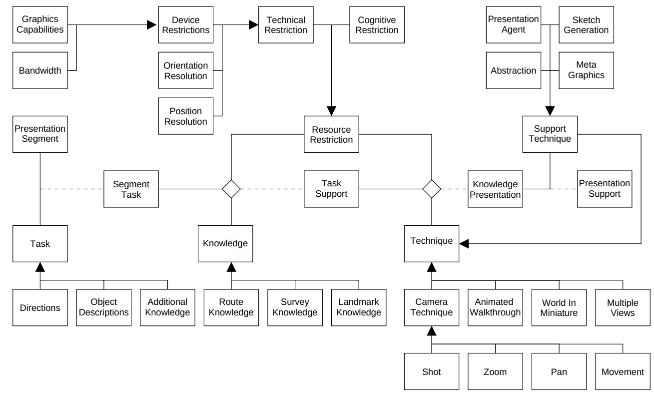

Figure 3: UML diagram of the navigation framework

in some more detail below. We will also describe graphical presentation techniques and point out their use for navigation systems.

The Generation Process

An elementary problem is the process of generating coherent explanatory presentations for a hybrid presentation system including stationary information booths with accelerated 3D graphics and mobile devices with very limited displays and communication bandwidth. The generated presentations should use appropriate techniques according to the different resource restrictions.

A presentation generally consists of one or more presentation segments corresponding to path segments as described above. Each of these presentation segments supports one or more tasks. In our context tasks are mainly directions, but also object descriptions and the presentation of additional knowledge about special locations or important objects along the way. As stated above these tasks are supported by offering different kinds of knowledge to the user: route, survey and landmark knowledge. The degree of support is mainly influenced by underlying resource restrictions. These restrictions include technical restrictions inherent to the presenting device, such as graphical capabilities and resolution, but also communication bandwidth for mobile devices. The quality of position and orientation information of the user is especially important as mentioned before. Other important restrictions are cognitive restrictions, such as the user's time constraints or presumed knowledge of the environment. All these restrictions influence the amount of information and the choice of graphical techniques to support a task.

First, the system has to determine for each segment or subset of adjacent segments the appropriate graphical way description schema. As explained in the previous section this depends on the resolution of location and orientation at each decision point. If the resolution of the location is sufficient, the presentation can be split at this decision point. In the best case (with good resolution at each decision point) the system can generate one presentation for each decision point and display it incrementally to the user along her path. If the resolution of the location at a certain decision point is bad, the system tries to recursively attach the presentation of the actual segment to a previous segment with a decision point that has a better resolution.

If, for example, the only output medium is the information booth, the only decision point with good resolution is the actual position of the user in front of the booth and all the presentations of segments have to be concatenated to one long description. Depending on the technical capabilities of the booth, the result could then be an animated walkthrough or a presentation with multiple frames.

In the presented framework both mobile devices and information booths are treated as output media underlying different technical restrictions. The mobile device can be thought of as an information booth with very limited technical capabilities, but that can be carried along and might be reused at every decision point providing a good resolution of location and orientation.

We therefore believe that this model is general enough to be used in other scenarios or for different classes of devices as well. Other output media (such as wall projection or head mounted displays) can be integrated smoothly by modeling their technical restrictions and modeling and implementing their available presentation techniques.

Graphical Techniques

While the resource restrictions are important for the decision *what* kind of knowledge to present, they will even more determine the way *how* information will be presented. The kind of generated presentation depends on the available techniques, which may include:

- camera techniques, such as static shots of the virtual environment, zooming, panning and moving.
- animated walkthroughs from an egocentric view.
- egocentric views including a World in Miniature (see [14, 7]).
- multiple views of the same scene, possibly including a bird's eye view.

These presentations can be supported by techniques such as:

- metagraphics accompanied with text to focus on landmarks or turning directions.
- graphical abstraction techniques to reduce computational load and focus attention (see [11]).
- black and white sketch generation for especially compact and flexible presentations.
- a virtual presenter showing the way and pointing to landmarks (see [3]).

Technique	Knowledge		
	Route	Survey	Landmark
Walkthrough	++	- -	+
Bird's eye view	o	++	o
World in Miniature	+	+	+
Multiple views	o	+	+
Metagraphics	++	- -	+
Abstraction	o	+	++
Sketch Generation	- -	++	o
Virtual Presenter	+	- -	++

Table 1: Different techniques and their ability to present different kinds of knowledge.

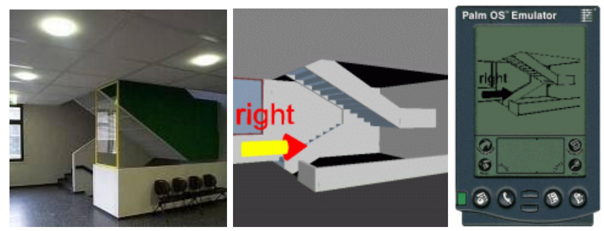

Figure 4: Picture of a hallway and two graphical presentations of the same place, one on a graphics screen, the other on the Palm Pilot

Table 1 rates different techniques for conveying different kinds of knowledge, where '+' means better, '-' means worse and 'o' stands for a neutral rating.

Coping with Resource Restrictions

An interesting aspect is the influence of resource restrictions on the choice of presentation techniques. While many different techniques may be available at a stationary system, there are less for mobile devices. For example animations are replaced with still images and due to today's monochrome displays on these appliances, black and white sketches have to be derived from a 3D model (see figure 4).

Also, bandwidth limits are hard restrictions when communicating with mobile devices irrespective of the transmission medium. In order to save bandwidth, 2D vector graphics are generated from the 3D model instead of a bitmap. These Vector graphics have two main advantages over bitmaps: they consume less bandwidth and memory and they can be scaled and rotated without loss of quality. In addition, generating vector representations after using graphical abstraction techniques can further reduce the amount of data.

Another major advantage of vector graphics is that they can be transmitted incrementally to a mobile device, thereby reducing the apparent transmission time (delay between the start of transmission and the first element displayed). Figure 5 shows an example of incremental generation and transmission: First, metagraphics are generated (the current position being indicated by a black dot) and transmitted as one pack-

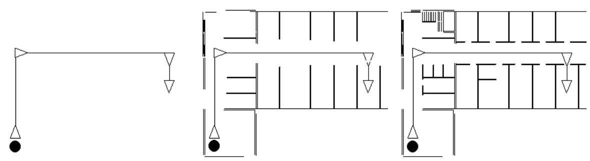

Figure 5: Incremental transmission of vector graphics: First an arrow depicting the path is shown, then the floor plan is transmitted in several steps, longest lines first.

age. Then, the rest of the scene is rendered and the corresponding vector graphics is derived. These vectors are then sorted by their size and transmitted in several packages, starting with the larger and possibly more significant line segments.

Information Booth Infrastructure

The stationary part of our system consists of a computer with a large LCD panel used as an information booth, where the user can choose a navigation goal. Then the path to this location is computed and from the list of path segments a presentation is generated. This presentation may include all techniques mentioned above: animated walkthroughs from an egocentric view possibly accompanied with a world in miniature, multiple views of the same scene including a bird's eye view, map-like black and white sketches of the building, metagraphics, graphical abstractions and a virtual presenter. The system is also used as the graphics generation server for mobile systems, including the automated generation of black and white sketches and 2D vector graphics from the original 3D model data.

Mobile infrastructure

The mobile navigation infrastructure consists of a number of strong infrared transmitters mounted to the ceiling in strategically important locations of the building, such as cross sections or landmarks. A sample installation has been done in our lab's offices and hallways, where visitors can choose from a list of employees and then are guided to the respective office. Since the mobile devices (currently Palm Pilots) had to remain unmodified and their builtin IrDA infrared ports can only transmit over distances of $1 - 2m$, the technical solution was to raise light intensity on the sender side and use the mobile devices in a passive way, i.e. just collecting and filtering data from the infrared streams.

For one user it is quite obvious how one can always send the right presentation (consisting of arrows, maps, or more generally vector graphics and text) at any given time and location and have it displayed on the device's screen. For up to ten or twenty users, presentations could be marked by user IDs and share the available infrared bandwidth in a time multiplexing scheme. In larger scenarios, however, we have to consider up to thousands of users passing the building at the same time. Since users are not being tracked, we cannot just broadcast their respective presentations only in places where they just happen to be at a certain time, but instead have to send data for all users in all locations all the time.

This at first sight seems to render the scenario absurd, but by structuring the information space differently, it can be overcome very easily. The trick here is to look at what the single senders will end up broadcasting. In the case of a rectangular cross section, for example, there are basically just four useful choices: You either have to turn left, right, walk on, or turn around because you might have gone wrong before. Every user will end up being sent one of those four possible types of information, so if we just broadcast four different arrows, each annotated with a list of some hundreds of user IDs, the necessary bandwidth will only grow very slowly with the number of users. Even broadcasting several stylistic variations of each arrow (for example with or without a schematic floor plan or a supporting text) will easily fit into this scheme.

Since the mobile devices filter out only the fitting packets from the data stream, users have the impression to receive a presentation tailored to them personally. In fact the thousands of possible route descriptions share the same building blocks, just composed in different ways. By using these techniques, we can effectively broadcast adapted navigation instructions to thousands of users in the scenario described above. More details on the system and the underlying broadcast protocol can be found in [2, 6].

Conclusions and future work

We have shown a framework for the consistent generation of a wide spectrum of different presentations on different output devices. The form and content of these presentations varies with the characteristics of the output medium, tracking accuracy and cognitive load of the user. We have implemented this framework in our own building navigation system and found that it produces adapted presentations with a high degree of coherence across devices. Also, each device is used to the best of its capabilities and different classes of devices can complement each other in conveying relevant information. In the near future we will carry out user studies to verify our assumptions on the choice of graphical techniques. Currently these assumptions are based on collected examples and on literature from the fields of graphics design and cognitive science.

Currently we are working on the extension of our system beyond buildings. By integrating a wearable computer, tracked by GPS and driving a small optical display attached to a user's glasses, we hope to be able to apply our common framework presented in this paper to yet another class of output medium. As output and interaction facilities become more complex, it seems especially important to integrate not only presentation techniques, but also user interaction schemata into our model. Depending on the device, situation and resource limitations, different types of interaction should be chosen that are appropriate for the task at hand.

In order to compare our planning approach, our group is also working on a decision-theoretic approach to the planning of navigation instructions for mobile devices [4]. This approach generates recommendation policies rather than a fixed set of instructions and by its probabilistic nature promises to yield an even more flexible user adaptation.

Acknowledgments

The research described in this paper was funded by 'Deutsche Forschungsgemeinschaft' DFG within their collaborative research program 378 on "resource-adaptive cognitive processes". Palm Pilots for the test installation were kindly provided by Deckarm & Co., St. Ingbert, Germany.

REFERENCES

1. Hisashi Aoki, Bernt Schiele, and Alex Pentland. Realtime personal positioning system for wearable computers. In *The third Symposium on Wearable Computers*, pages 37–43, Los Alamitos, California, 1999. IEEE Computer Society.
2. Jörg Baus, Andreas Butz, and Antonio Krüger. One way interaction: interactivity over unidirectional links. In *Proceedings of I3 workshop on Adaptive Design of Interactive Multimedia Presentations for Mobile Users, Sunday, March 7, 1999*. www.i3net.org, 1999.
3. Jörg Baus, Andreas Butz, and Antonio Krüger. Incorporating a virtual presenter in a resource adaptive navigational help system. In *Workshop notes of the Workshop on Guiding Users through Interactive Experiences*. C-Lab, Univ. Paderborn, 2000.
4. Thorsten Bohnenberger and Anthony Jameson. When policies are better than plans: Decision-theoretic planning of recommendation sequences. In *IUI2001: International Conference on Intelligent User Interfaces*, New York, 2001. ACM.
5. Andreas Butz. Anymation with CATHI. In *Proceedings of AAAI/IAAI '97*, pages 957–962. AAAI Press, 1997.
6. Andreas Butz, Jörg Baus, and Antonio Krüger. Augmenting buildings with infrared information. In *Proceedings of the International Symposium on Augmented Reality ISAR 2000*, 10662 Los Vaqueros Circle, P.O. Box 3014, Los Alamitos, CA 90720-1264, 2000. IEEE Computer Society Press.
7. T. Elvins, D. Nadeau, and D. Kirsh. Worldlets - 3d thumbnails for wayfinding in virtual environments. In *Proceedings of UIST'97*, pages 21–30. ACM SIGGRAPH, 1997.
8. K.-P. Gapp. Selection of Best Reference Objects in Object Localizations. In *Proceedings of the AAAI Spring Symposium on Cognitive and Comutational Models of Spatial Representations*, pages 23–34. Stanford, CA, 1996.
9. S.C. Hirtle. The Use of Maps, Images and "Gestures" for Navigation. In C. Freksa, W. Bauer, C. Habel, and K.F. Wender, editors, *Spatial Cognition II – Integrating Abstract Theories, Empirical Studies, Formal Methods and Practical Applications*, pages 31–40. Springer, 2000.
10. C. Kray and A. Blocher. Modeling the basic meanings of path relations. In *IJCAI–99: Proceedings of the 16th International Joint Conference on Artificial Intelligence*, pages 384–389. Morgan Kaufmann Publishers, Inc., 1999.
11. Antonio Krüger. Automatic graphical abstraction in intent-based 3d illustrations. In *Advanced Visual Interfaces '98*, pages 47–55. ACM Press, 1998.
12. K.L. Lovelace, M. Hegarty, and D.R. Montello. Elements of Good Route Directions in Familiar and Unfamiliar Environments. In C. Freksa and D.M. Mark, editors, *COSIT 99: Spatial information theory: cognitive and computational foundations of geographic information science*, pages 65–82. Springer, 1999.
13. M.E. Sorrows and S.C. Hirtle. The Nature of Landmarks for Real and Electronic Spaces. In C. Freksa and D.M. Mark, editors, *COSIT 99: Spatial Information theory: cognitive and computational foundations of geographic information science*, pages 37–50. Springer, 1999.
14. Richard Stoakley, Matthew J. Conway, and Randy Pausch. Virtual reality on a wim: Interactive worlds in miniature. In *Proceedings of CHI 1995, Denver, Colorado*, pages 265–272, 1995.
15. S. Towns, C. Callaway, and J. Lester. Generating Coordinated Natural Language and 3D Animations for Complex Spatial Explanations. In *AAAI–98: Proceedings of the 15th National Conference on Artificial Intelligence*, 1998.
16. B. Tversky. Cognitive Maps, Cognitive Collages, and Spatial Mental Models. In *Spatial Information Theory. A Theroretical Basis for GIS, COSIT'93*, pages 14–24, 1993.
17. B. Tversky and P. U. Lee. How Space Structures Language. In C. Freksa, C. Habel, and K.F. Wender, editors, *Spatial cognition – An interdisciplinary approach to representation and processing of spatial knowledge*, pages 157–177. Springer, 1998.
18. B. Tversky and P. U. Lee. Pictural and Verbal Tools for Conveying Routes. In C. Freksa and D.M. Mark, editors, *COSIT 99: Spatial Information theory: cognitive and computational foundations of geographic information science*, pages 37–50. Springer, 1999.