In: Proceedings of the 2nd Symposium on Location Based Services & TeleCartography 2004
from January, 28-29 in Vienna. G. Gartner (ed.), Geowissenschaftliche Mitteilungen, Schriftenreihe
der Studienrichtung Vermesungswesen und Geoinformation, TU Wien

Augmented Videos and Panoramas for Pedestrian Navigation

Thomas H. Kolbe, Bonn

Abstract

In order to support pedestrians in the tasks of orientation and wayfinding with mobile devices appropriate visualizations of their location, its surrounding, and background information are needed. In this paper we present two complementary concepts for the combination of realistic views and location specific information that can be implemented with current generation PDAs and next generation cellular phones. The first employs video clips that were taken along every path in a city which are augmented by location information. It is shown how multimedia clips consisting of videos and textual information can be stored and transferred using the SMIL standard from the W3C consortium. The second approach supports navigation by the overlay of virtual signposts in panoramic images taken for every decision point, i.e. street junctions, in a city. We argue that the location- and orientation-synchronized presentation of (previously recorded) realistic views in combination with location specific information is an intuitive complement to map based visualization on the one hand and less invasive and resource demanding than augmented reality systems with head mounted displays on the other hand.

1 Introduction

Where am I? What can I see and do here? What is important to know about the place? Where do I have to go to reach my target?

These are the most important questions that pedestrian navigation systems (PNS) should help to answer for mobile users (cf. Malaka & Zipf 2000). Although verbal output from PNS can support a user in navigation, the visual presentation of the current location, its surrounding, and location specific information is most important for orientation and wayfinding. Generally, the displayed information in PNS are composed of: 1) the visual image or abstraction of the scene and its surrounding, 2) information about the place, and 3) route / navigation information, where 2) and 3) typically augment 1). Depending on the concrete realisation of 1)-3) the difficulties for a user to match the displayed

- scene with reality (finding his position and heading),
- location specific information with the objects depicted in the scene, and
- navigation information with the streets depicted in the scene

differ a lot. For instance, maps are abstractions of the real world and are well suited to give an overview of an area and for estimation of distances and areas. Nevertheless some cognitive effort is needed to determine the own orientation from looking at maps. When using augmented reality (AR) instead, matching the scene with reality is immediate, but it may be difficult to put location specific information at the right place due to positioning and orientation errors (Azuma 1997).

Thus, there are several design criteria and trade-offs for pedestrian navigation systems that visually support orientation and wayfinding (cf. Höllerer et al. 2001, Baus et al. 2002, Gartner 2003). We believe that the following aspects are crucial wrt. the effectiveness of orientation aid, user acceptance and feasibility regarding system requirements and data availability:

- Scene visualization should be 3d and as realistic as possible to provide enough visual cues and landmark information for quick recognition (cf. Kolbe 2002).
- The visualization of the scene should be synchronized with the user's current position and heading in order to facilitate immediate matching of the displayed image with the real view perceived by the user.
- Navigation and location specific information should be displayed in terms of augmentations to the displayed scene background. However, the placement of location specific information has to take into account the available precision of the position information and spatial data (see Höllerer et al. 2001, Gartner & Uhlirz 2001, Baus et al. 2002).
- Visualization should be tailored to the small displays of mobile information devices (MID) (cf. Gartner & Uhlirz 2001, Gartner 2003). For acceptance reasons an invasive (wrt. to unobstructed vision) display apparatus like head mounted displays (HMD) should be avoided.
- The system should need as few explicit world knowledge in terms of 2d and 3d spatial data as possible (e.g. no complete 3d city model). It should be examined to what extent spatial data from existing 2d navigation systems can be exploited.
- Since the system should work with current generation Personal Digital Assistants (PDAs) and next generation cellular phones (3G), the low computational power of these MIDs has to be considered (cf. Pospischil et al. 2002).

With these issues in mind we have developed two complementary visualization approaches for pedestrian navigation. The first concentrates on the visualization of the paths along streets and crossing places in cities, for which video clips are recorded in both walking directions and augmented by location specific information. The second focuses on street junctions and integrates panoramic images with navigation information for each junction of a city.

The rest of the paper is organised as follows: In section 2 we discuss related work. Section 3 describes our approach for augmented videos, including screenshots from a protoype implementation. In section 4 the concept for augmented panoramic images is explained and screenhots from another prototype are given. Finally, section 5 presents the conclusions and directions for future work.

2 Related Work

When reviewing literature on PNS it can be observed that most of todays concepts were developed in either of two research contexts; namely location based services (LBS) and outdoor augmented reality systems. This is reflected (at least) by their different visualization approaches, especially in the way how information about the current location and the surrounding is presented to users. Some prominent systems, that were developed in the LBS context, are Cyberguide (Abowd et al. 1997), GUIDE (Cheverst et al. 2000), Deep Map (Malaka & Zipf 2000), and Lol@ (Pospischil et al. 2002). In these systems the presentation of the user's location and its surrounding is mainly based on maps, which are displayed on the screen of a mobile information device. Further multimedia information like texts and photos about landmarks and points-of-interest are linked with signatures in the map. In outdoor AR systems like Touring

Machine (Feiner et al. 1997), ARREAL (Baus et al. 2002), and Studierstube Mobile AR (Reitmayer & Schmalstieg 2003) users perceive the current location through head mounted displays (HMD) instead. Multimedia information about objects in the respective viewing area are presented as an overlay to the view of the real world.

Although both approaches differ substantially they have in common that they mix reality and virtuality to support orientation and wayfinding: the LBS approach by adding photos and movies of real objects to virtual presentations of reality (maps) and the AR approach by adding virtual objects and multimedia information to real scene views. In order to classify mixed reality applications we follow the line of (Milgram & Kishino 1994), who proposed to arrange the different approaches along the so-called *virtuality continuum*. Figure 1 reveals the big gap between the LBS approach (augmented virtuality) and AR systems (augmented reality), which we want to bridge using augmented videos and panoramas as explained in the next two sections.

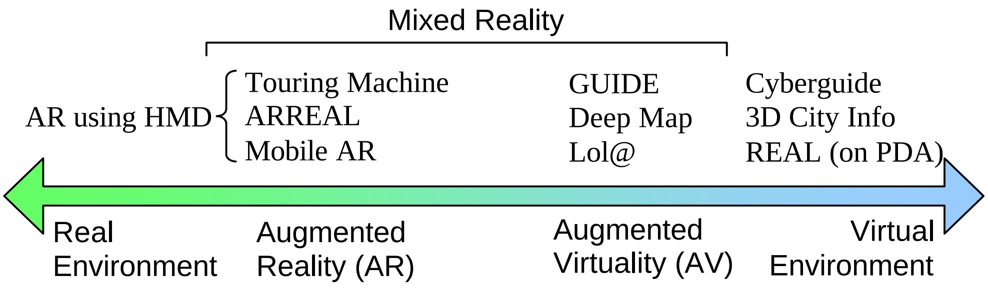

Figure 1: *Classification of PNS wrt. to the virtuality continuum (adapted from Milgram & Kishino 1994).*

Further relevant work concerns visual realism, i.e. the generation and usage of realistic views in virtual reality applications. Other PNS like 3D City Info (Rakkolainen & Vainio 2001) and IRREAL (on PDAs; Baus et al. 2002) use completely virtual presentations, i.e. 3d models to generate realistic views. However, complex and textured 3d city models are needed to reach a sufficient degree of visual realism for wayfinding (cf. Ruddle et al. 1998). An alternative approach is called *image based rendering* (IBR), where realistic views are synthesized from previously recorded photos and movies. Panoramic 360° images are one of the first and most prominent applications of IBR (Chen 1995). In recent years IBR techniques have extensively been applied for creating so-called *virtual walkthrough environments*, where users can move in realtime through a virtual space that is mainly synthesized from images and movies in a database (Lippman 1980, Chen 1995, Tanikawa et al. 1999, Kimber et al. 2001, Kolbe 2002). However, virtual walkthrough environments have not been applied in the context of mobile navigation assistance so far.

3 PNS using augmented videos

As stated above and is discussed in more detail in (Kolbe 2002), we want to provide visual support for orientation and wayfinding by combining realistic views with location specific information. Whereas outdoor AR systems already realize this concept, they have major drawbacks concerning the needed precision for position resp. orientation determination and hardware requirements. The usage of HMDs is very sensitive to orientation and position errors. Typically, small deviations already lead to inconsistent visualizations where the real view and its augmentations do not match, which quickly confuse the user.

To overcome this problem, we suggest to use previously recorded video clips of the paths in pedestrian areas. A walkthrough computer video can be considered a visual memory of the area along the route, because it is not restricted to linear replaying with constant frame rate in one time direction. Every video frame can be addressed and therefore be georeferenced individually. In contrast to holography, which memorizes the appearance of an object at a fixed position from all perspectives, walkthrough videos memorize the appearance of geographic locations in fixed perspective but arbitrary position along a trajectory. Since the videos are recorded before the system will be used, the georeferencing and synchronization of video and augmentations can be prepared off-line in advance (*Video AR:* cf. Azuma 1997, Gibson et al. 2002).

In order to obtain a complete visual representation of an area, each path is recorded on video in both walking directions. The paths are defined by the model of the respective area. Like in other PNS we model pedestrian zones by attributed graphs. Although pedestrians can walk freely over open ground, in most cases people move straight on prepared paths in the direction of their next target. Typically, a change of direction only takes place when following curved streets and at junctions. Therefore, junctions and intermediate points in a curve are represented by graph nodes and streets by graph edges. Places are modeled by 1) a ring of edges which connect all neighboured entrance points of the place and 2) edges from each entrance point to the others across the place. For big places further nodes and connecting edges are added within the surrounding polygon (see figure 2). Graph nodes are georeferenced and named. Edges are attributed by the street resp. place names and the house numbers with their positions on both sides. Furthermore, the location, the name, and the type of all points-of-interest are known. Further details on modeling (also of non-pedestrian zones) and route planning issues for PNS are given in (Kolbe 2002).

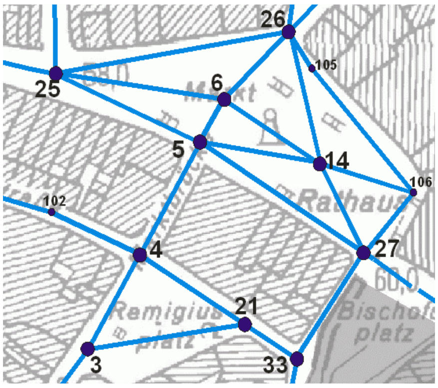

Figure 2: *Graph representation of places and streets in a pedestrian zone.*

For the operation of the PNS a position sensor like (D)GPS and an orientation sensor like a magnetic compass, a gyroscope or a combination of both is needed in the MID (cf. Retscher & Skolaut 2003, Höllerer et al. 2001, Baus et al. 2002). While walking in the modeled area, the nearest point on the next graph edge to the user is continuously determined. Depending on the user’s heading the video for either walking direction is selected. The frame of the video that is related to the current position then is presented to the user together with the augmentations. In contrast to AR using HMDs the scene views and their augmentations are always synchronized. Position errors only lead to differences between the real and the displayed scene view. If this error is small enough (less than 3m), it will be still easy for users to match the displayed view with their real view. However, this hypothesis has to be evaluated in future work.

The visual quality of the integration of location specific information with the background depends on the achievable precision of position resp. orientation information on the one hand and the quality of explicit world knowledge in terms of 3d models and coordinates of the objects depicted in the videos on the other hand. Thus, we distinguish two levels of possible augmentations to walkthrough videos which have increasing demands on precision: 1) Augmentation by annotations next to, i.e. above, below, or on either side of the movie clip. 2) Overlay of information within the movie area. For simplicity reasons we make the assumption in our current prototype, that the videos were filmed exactly along the graph edges and with constant velocity.

For the implementation one problem still had to be solved: how can videos and augmentations be integrated in advance and transferred to resp. stored in mobile devices? Above, the system should be usable on MIDs with low computational power. The solution lies in the application of the *Synchronized Media Integration Language SMIL* of the W3C consortium, a XML application which allows to integrate and relate different media with respect to a general presentation timeline (SMIL 2001, Kennedy & Slowinski 2001). SMIL browsers are available for PCs, PDAs, and cellular phones (see Real 2002). In the following we describe the structure of the SMIL representation of walkthrough videos that are augmented by location specific information (see figure 3).

```
<smil>
  <head>
    <root-layout width=”320” height=”260”/>
    <region id=”video” top=”0” left=”0” width=”320” height=”240”
            fit=”fill” z-index=”1”/>
    <region id=”streetname” top=”242” left=”25” height=”18” width=”270”/>
    <region id=”houseno_left” top=”242” left=”0” height=”18” width=”25”/>
    <region id=”houseno_right” top=”242” left=”295” height=”18” width=”25”/>
    <region id=”label1” top=”90” left=”110” height=”15” width=”105” z-index=”2”/>
  </head>
  <body>
    <par>
      <video src=”019-020.rm” region=”video” fill=”hold”/>
      <textstream src=”019-020.rt” region=”streetname”
                  dur=”7.0s” fill=”freeze”/>
      <textstream src=”019-020_links.rt” dur=”7.0s”
                  region=”houseno_left”/>
      <textstream src=”019-020_rechts.rt” dur=”7.0s”
                  region=”houseno_right”/>
      <textstream src=”019-020_label.rt” region=”label1” dur=7.0s>
        <animate targetElement=”label1” attributeName=”left”
                 to=”215” dur=”7.0s” fill=”freeze”/>
        <animate targetElement=”label1” attributeName=”top”
                 to=”40” dur=”7.0s” fill=”freeze”/>
      </textstream>
    </par>
  </body>
</smil>
```

see figure 4

Moving text label

Multimedia content for one video path

***Figure 3:** Example for the definition of a multimedia clip for one directed walking path using SMIL 2.0. The head section defines the screen layout for the different media elements. Please note, that “label1” lies in the same area as “video”, but is shown on top due to a higher z-index. The body element references and synchronizes five media elements: the walkthrough video clip, the street name, the current house numbers on both sides, and a text label.*

The head section specifies the screen layout for the multimedia clip. It consists of four non-overlapping regions for the video, the street name, and the house numbers to the left and to the right (cf. figure 4). A fifth region lies in the same screen region as the video, but will be shown on top of the video due to its higher *z-index*. This region is intended to display a text label.

The body section defines within the <par> block a group of five media elements, that will be presented simultaneously with synchronized timing. The group consists of the video of the path from a start to an end node of the route graph, the textstream with the street name, the textstream with the house numbers for both sides, and a textstream with a label for points-of-interests. Textstreams are an extension of the RealOnePlayer to the SMIL standard, that allow to show resp. hide text portions at arbitrary positions and times (see Real2002).

The position of the label field is animated by two <animate> tags. It starts at *left,top=(110,90)* (as initialised in the head section) and moves during the clip’s 7 seconds duration with constant velocity and linear interpolation to position *(215,40)*. By this mechanism label texts can be visually attached to objects depicted in the video like buildings and monuments: if the object comes nearer and moves to another screen position, the label moves accordingly. In general, we can derive from camera projection equations that linear interpolation is not sufficient, because the movement of depicted objects speeds up as the camera approaches them. SMIL2.0 offers spline animation and time manipulations (cf. SMIL 2001, Kennedy & Slowinski 2001) that would allow to reflect the acceleration of objects. However, since RealOnePlayer does not support these features (yet), we have not implemented it in the current version of the prototype.

If the route graph is attributed with house numbers and street names, the SMIL file and its referenced textstream files can be generated automatically for every graph edge in both walking directions. The duration of the <par> block is determined by the length of the corresponding path video. If labels should be placed and fitted automatically within the video area, the videos have to be georeferenced, i.e. extrinsic and intrinsic camera parameters have to be determined for each video frame. Above, 3d position information for the objects to be labeled have to be known. We are currently working on this issue and already got some promising results from experiments with automatic camera trackers (so-called *match movers*) like ICARUS (Gibson et al. 2002). In our prototype system all files for the modeled portion of Bonn city have been created manually.

Figure 4 shows screenshots from a virtual walkthrough with the prototype in the city center of Bonn. In the street “Bonngasse” the user approaches the Beethoven house, in which the famous composer was born and had lived for 22 years.

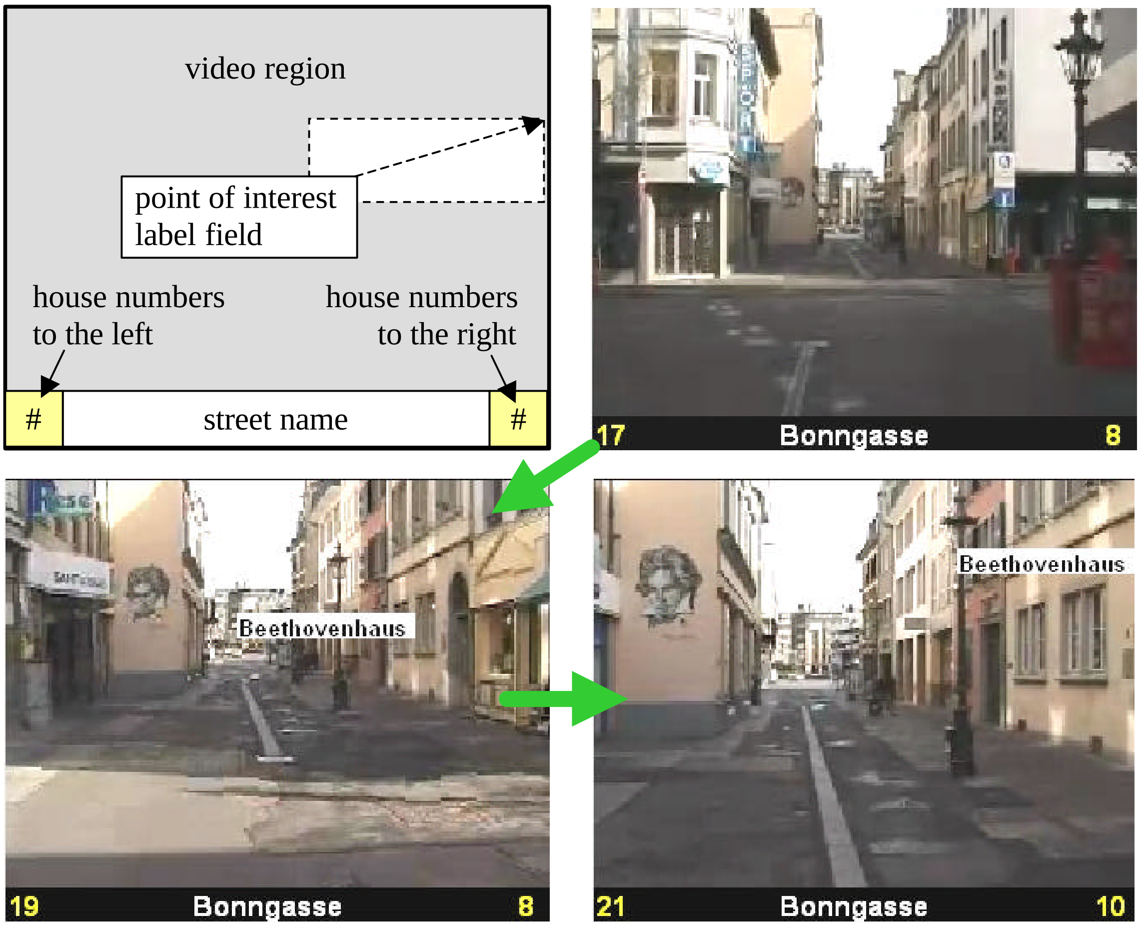

Figure 4: *Screen layout for and screenshots from the prototype for augmented video visualization using RealOnePlayer (Real 2002). The screenshots were taken at three points along the path to the “Beethovenhouse” in Bonn. Please note, that at a certain distance a label becomes visible which then moves within the video region.*

4 PNS using augmented panoramas

Our second concept for a PNS employs georeferenced panoramic images which are augmented by virtual signposts. Panoramic images memorize the visual appearance of the surrounding for a fixed location in all viewing directions (cf. Chen 1995). This makes them especially suited to support users when they are looking around at some point to find the way to their target. The idea is based on the observation that wayfinding support is especially needed at decision points, i.e. street junctions and places, where one of the different possible directions has to be chosen. Thus, for every junction of the route graph a 360° cylindrical panoramic image is captured, and the direction to north and the camera parameters are determined. Details on the generation of the panoramic images using a standard digital camera are given in (Middel 2003).

During system operation the user specifies his target (another graph node). Then the PNS continuously calculates the shortest path to the target. Like in the previous approach the MID needs to be equipped with a position and an orientation sensor. Everytime, the user is in the vicinity of a graph node, the corresponding panorama will be shown (cf. fig. 7A). The displayed portion of the panorama is always synchronized with the current heading of the user.

The following steps are processed for every change of the user's view. In order to augment panoramic views by virtual signposts, the first problem that has to be solved is the determination of areas in the projected views, in which signposts can be visualized without the occlusion of salient objects or other visual cues. This task is complicated by the lack of explicit knowledge about the world in terms of a 3d city model. In our system we assume that no 3d model is available and the edges of the route graph (cf. fig. 2) only have an additional attribute *street width*.

However, when looking at city views, one can distinguish three principal visual components: 1) building areas, 2) sky, and 3) street areas. Figure 5 shows on the left the partitioning of a city view into the three areas. Since most landmarks are found in area 1), the sky silhouette of area 2) is needed to support orientation, and the street information is the only available explicit world knowledge, we only use the street area. In order to determine the street area for the current view, the street borders for the shortest path are derived from the center line and street widths in 3d object space (see right side of fig. 5). Currently, we have not applied a digital terrain model (DTM) yet. Therefore, all heights are set to 0 in the prototype.

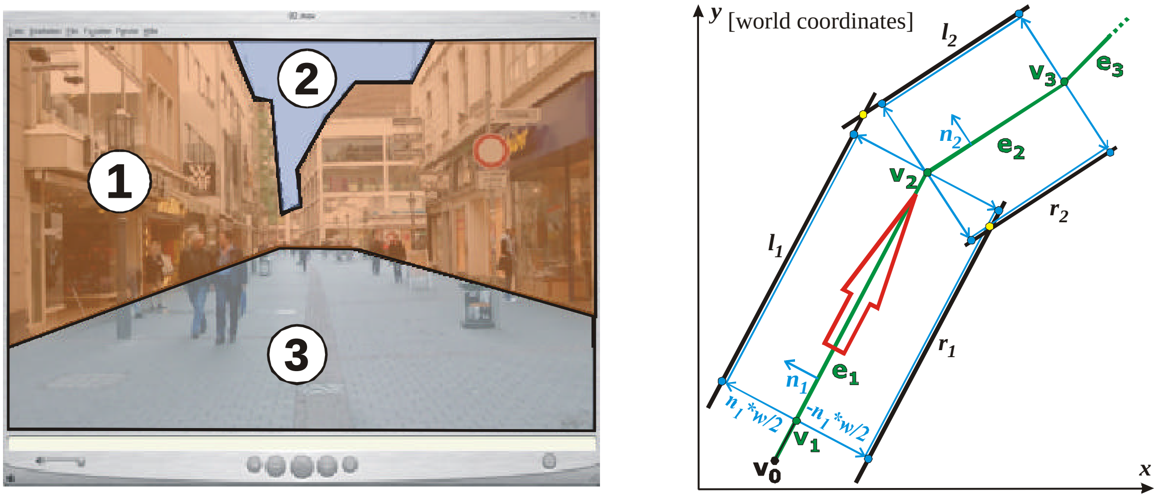

Figure 5: *Left: Principle components of a typical city scene view are 1) building areas, 2) sky, and 3) street areas. In order to minimize occlusions the location specific information like place names and waysigns are placed in the street area only. Right: Construction of the street polygons and generation of virtual waysigns in 3d object space by creation of left and right border lines from route graph edges* e_1, e_2, e_3 *and street width* w.

After the creation of the street borders, they are projected into the current panoramic view in order to compute the (2d) polygon that surrounds the visible street area in the image. In inner city areas the streets are typically bounded by buildings. Thus, the visibility of a street course ends at a curve and the projected street can be truncated accordingly. Figure 6 shows an example for the computation of the street area polygon where the street is truncated at the first left turn. Generally, several special cases for the truncation and closing of the street area polygon have to be differentiated. They are discussed in detail in (Middel 2003) and (Kolbe et al. 2004).

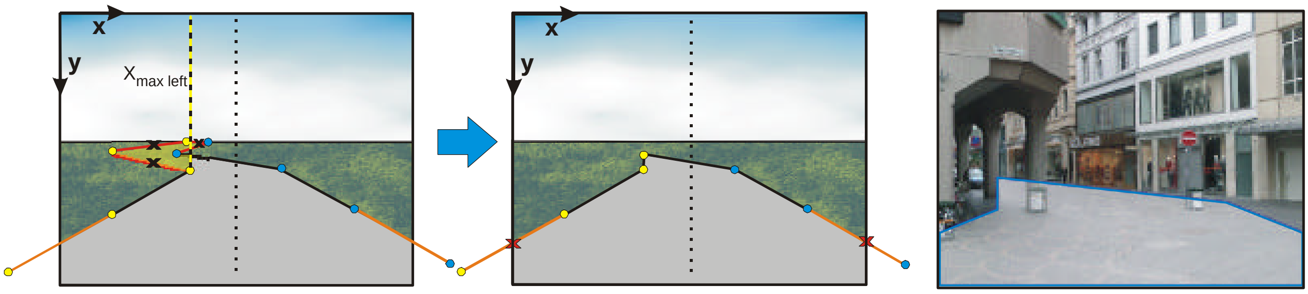

Figure 6: *Computation of the free labeling area by projection and truncation of street polygons. The left border of the street is truncated at the point where the projected x value begins to decrease. It is assumed that a building wall will occlude the rest of the path after the curve. Thus, the polygon is closed by a vertical line.*

The placement of virtual signposts is done in two steps: At first, a marker is generated in (3d) object space several meters ahead of the current viewpoint on the centerline of the graph edge, which lies on the shortest path to the target (see right side of fig. 5). In the second step the marker is projected into the current view. If it lies completely within the free labeling area, i.e. the previously computed street polygon area, it will be displayed as an overlay to the current panoramic projection. This is normally the case, when the user is looking in the direction of the shortest path to the target. If the user looks in another direction, the projected marker typically does not fit into the free area or does not lie in the polygon at all. In this case, an arrow is shown on the left resp. right side of the image, telling the user to turn in the corresponding direction to find the correct heading to the target (see figure 7B).

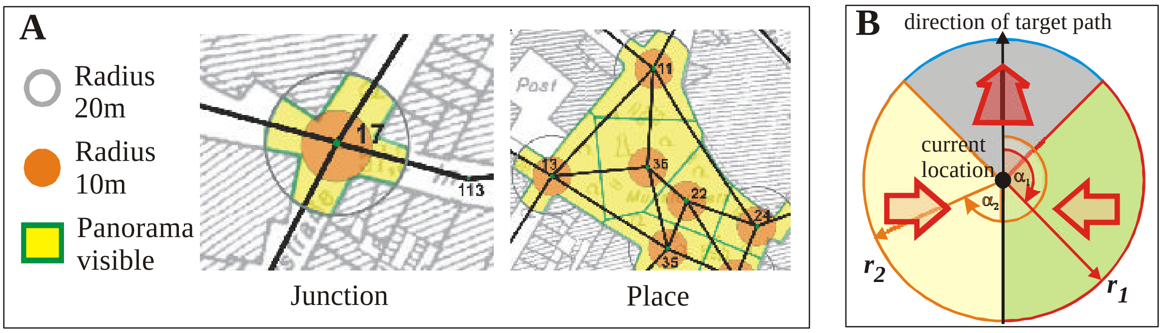

Figure 7: *A) Visibility areas for panoramas: If a user enters the 20m zone, the panorama will be displayed on the mobile device with a fixed direction to the target. Within the red zone (10m) the displayed portion of the panorama is synchronized with the users current viewing direction. B) Rotation dependent generation of signposts. When looking in target direction: 3d arrow; in other directions: simple 2d turn left resp. right signs.*

Figure 8 shows screenshots from the prototype that were made during a virtual look-around at a junction in the city center of Bonn. Since in the first and the last image the projected marker does not fit into the labeling area (orange polygon), turn right resp. left signs are displayed instead.

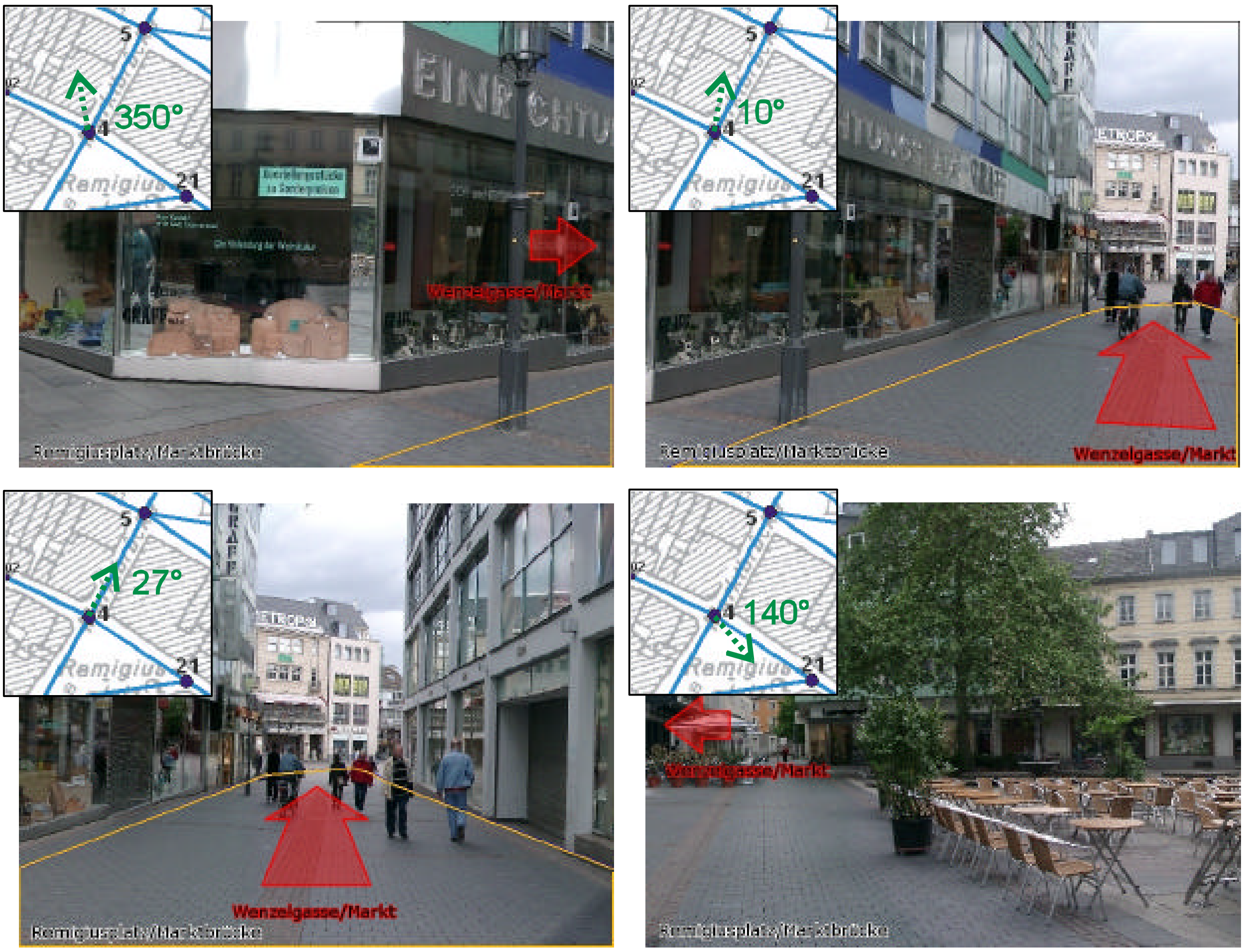

Figure 8: *Visualization of virtual signposts in different viewing directions at the same location.*

5 Conclusions and future work

Augmented videos and panoramas provide intuitive support for wayfinding and orientation and complement traditional maps in pedestrian navigation systems. They are presented on the display of a mobile device which in contrast to a head mounted display is not invasive wrt. to the user's field of vision. The concepts put only small demands on the needed resources concerning position/orientation accuracy, explicit world knowledge, and computational power and thus can be implemented and used with current generation PDAs and third generation (3G) cellular phones. Furthermore, data acquisition is relatively cheap, because videos and panoramas can be recorded resp. generated using standard off-the-shelf hardware and software. In the MIT city scanning project a system for automated acquisition of georeferenced, panoramic images is already described (Teller et al. 2001). Whereas SMIL seems to be an appropriate framework for building multimedia clips for walkthrough videos, it has to be examined in the future how panoramic images could be handled resp. how both concepts can be integrated in general.

One drawback is that videos only memorize the world's appearance at the time of recording. Thus, changes in the recorded area will not be reflected and could lead to some confusion when the user tries to match video and reality. A related problem is day and night view. However, this can be overcome when videos and panoramas are recorded both for daytime and nighttime.

Cartographic challenges in future work concern the non-overlapping placement of simultaneously shown labels and signposts in object and image space (cf. Azuma & Furmanski 2003).

References

Abowd, G., D.; C. G. Atkeson, J. Hong, S. Long, R. Kooper, M. Pinkerton (1997): CYBERGUIDE: A MOBILE CONTEXT-AWARE TOUR GUIDE. ACM Wireless Networks, 3:421-433, 1997

Azuma, R. (1997): A SURVEY OF AUGMENTED REALITY. Presence: Teleoperators and Virtual Environments 6, 4 (August 1997), 355 - 385.

Azuma, R., C. Furmanski (2003): EVALUATING LABEL PLACEMENT FOR AUGMENTED REALITY VIEW MANAGEMENT. Proc. IEEE and ACM Int'l Symp. on Mixed and Augmented Reality (ISMAR 2003) in Tokyo, 2003

Baus, J., A. Krüger, W. Wahlster (2002): A RESOURCE-ADAPTIVE MOBILE NAVIGATION SYSTEM. In: Proceedings of Int. Conference on Intelligent User Interfaces IUI 2002 in San Francisco, California, USA, ACM Press, New York

Chen, S. E. (1995): QUICKTIME VR: AN IMAGE-BASED APPROACH TO VIRTUAL ENVIRONMENT NAVIGATION. In: Proc. of the 22nd Conference on Computer Graphics and Interactive Techniques SIGGRAPH 95, ACM Press

Cheverst,K., N. Davies, K. Mitchell, A. Friday, C. Efstratiou (2000): DEVELOPING A CONTEXT-AWARE ELECTRONIC TOURIST GUIDE: SOME ISSUES AND EXPERIENCES. In: Proceedings of the 2000 Conference on Human Factors in Computing Systems (CHI-00), pages 17–24, New York, ACM Press

Feiner, S., B. MacIntyre, T. Höllerer, A. Webster (1997): A TOURING MACHINE: PROTOTYPING 3D MOBILE AUGMENTED REALITY SYSTEMS FOR EXPLORING THE URBAN ENVIRONMENT. In: Proc. of the IEEE Int. Symp. on Wearable Computing ISWC'97 in Cambridge, MA, USA.

Gartner, G. (2003): PEDESTRIAN NAVIGATION SERVICES: A MULTIMEDIA CARTOGRAPHY APPROACH TO MOBILE INTERNET APPLICATIONS. In: Proc. of the 21th Int. Cartographic Conference ICC 2003, Durban, South Africa

Gartner, G., S. Uhlirz (2001): CARTOGRAPHIC CONCEPTS FOR REALIZING A LOCATION-BASED UMTS SERVICE: VIENNA CITY GUIDE LOL@. In: Proc. of the 20th Int. Cartographic Conference ICC 2001, Beijing, China. Vol. III.

Gibson, S., J. Cook, T. Howard, R. Hubbold, D. Oram (2002): ACCURATE CAMERA CALIBRATION FOR OFF-LINE, VIDEO-BASED AUGMENTED REALITY. In: Proc. of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR'02) in Darmstadt, Germany

Höllerer, T., S. Feiner, D. Hallaway, B. Bell, M. Lanzagorta, D. Brown, S. Julier, Y. Baillot, L. Rosenblum (2001): USER INTERFACE MANAGEMENT TECHNIQUES FOR COLLABORATIVE MOBILE AUGMENTED REALITY. Computers and Graphics 25(5), Elsevier Science

Kennedy, T., M. Slowinski (2001): SMIL – ADDING MULTIMEDIA TO THE WEB. Sams Publishing, Indianapolis, USA

Kimber, D., J. Foote, S. Lertsithichai (2001): FLYABOUT: SPATIALLY INDEXED PANORAMIC VIDEO. In: Proceedings of the 9th ACM Internation Conference on Multimedia 2001, ACM Press, New York, USA

Kolbe, T. H. (2002): FUßGÄNGERNAVIGATION UND ROUTENPLANUNG IN INNENSTÄDTEN UND GEBÄUDEN MIT VIDEOS UND PANORAMEN. In: *Möltgen, Wytzisk* (Eds.): GI-TECHNOLOGIEN FÜR VERKEHR UND LOGISTIK. Tagungsband Münsteraner GI-Tage 2002, IfGI Prints, Nr. 13, Uni Münster, 2002

Kolbe, T. H., A. Middel, L. Plümer (2004): 3D-KARTOGRAPHIE FÜR DIE FUßGÄNGERNAVIGATION: VIRTUELLE WEGWEISER IN PANORAMEN. To appear in: Tagungsband der 40. Sitzung der AG Automation in der Kartographie AgA 2003 in Erfurt. Mitteilungen des Bundesamtes für Kartographie und Geodäsie, Verlag des BKG, Frankfurt am Main, Germany

Lippman, A. (1980): MOVIE-MAPS: AN APPLICATION OF THE OPTICAL VIDEODISC TO COMPUTER GRAPHICS. In: Proc. of the 7th Conference on Computer Graphics and Interactive Techniques SIGGRAPH 80

Malaka, R. and Zipf, A. (2000): DEEP MAP - CHALLENGING IT RESEARCH IN THE FRAMEWORK OF A TOURIST INFORMATION SYSTEM. In: D. Fesenmaier, S. Klein, D. Buhalis (Eds.): Information and Communication Technologies in Tourism 2000. Proc. of ENTER 2000 in Barcelona, Spain. Springer, Wien, New York.

Middel, A. (2003): VIRTUELLE WEGWEISER IN PANORAMEN. MS Thesis, Institute for Cartography and Geoinformation, Univ. of Bonn. http://www.ikg.uni-bonn.de/kolbe/diplomarbeiten/Diplomarbeit_Ariane_Middel.pdf

Milgram, P., F. Kishino (1994): A TAXONOMY OF MIXED REALITY VISUAL DISPLAYS. IEICE Transactions on Information Systems, Vol E77-D, No.12 December 1994.

Pospischil, G, M. Umlauft, E. Michlmayr (2002): DESIGNING LOL@, A MOBILE TOURIST GUIDE FOR UMTS. 4th Int. Symposium on Human Computer Interaction with Mobile Devices (Mobile HCI '02), Pisa, Italy, September 2002.

Rakkolainen, I., T. Vainio (2001): A 3D CITY INFO FOR MOBILE USERS. Computers & Graphics, Vol. 25, No. 4, Elsevier.

Real (2002): REALPLAYER UND REALONEPLAYER. http://www.realnetworks.com

Reitmayer, G., D. Schmalstieg (2003): LOCATION BASED APPLICATIONS FOR MOBILE AUGMENTED REALITY. In: Proceedings of the 4th Australasian User Interface Conference AUIC'03, Adelaide, Australia.

Retscher, G., G. Skolaut (2003): UNTERSUCHUNG VON MESSSENSOREN ZUM EINSATZ IN NAVIGATIONSSYSTEMEN FÜR FUSSGÄNGER. ZfV 2/2003, Zeitschrift für Geodäsie, Geoinformation und Landmanagement, Wißner-Verlag.

Ruddle, R. A.; S. J. Payne, D. M. Jones (1998): NAVIGATING LARGE-SCALE "DESKTOP" VIRTUAL BUILDINGS: EFFECTS OF ORIENTATION AIDS AND FAMILIARITY. MIT Presence, Vol. 7, No. 2, April 1998

SMIL (2001): SYNCHRONIZED MULTIMEDIA INTEGRATION LANGUAGE SMIL 2.0. W3C Recommendation 07 August 2001. http://www.w3.org/TR/smil20/

Tanikawa, T., T. Endo, M. Hirose (1999): BUILDING A PHOTO-REALISTIC VIRTUAL WORLD USING VIEW-DEPENDENT IMAGES AND MODELS. In: Proc. of IEEE Int. Conf. on Systems, Man and Cybernetics SMC'99, Vol. 6

Teller, S., M. Antone, Z. Bodnar, M. Bosse, S. Coorg, M. Jethwa, N. Master (2001): CALIBRATED, REGISTERED IMAGES OF AN EXTENDED URBAN AREA. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition CVPR 2001