'Photogrammetric Week 01' D. Fritsch & R. Spiller, Eds. Wichmann Verlag, Heidelberg 2001.

Automated image orientation in a location aware environment

NORBERT HAALA, Stuttgart

ABSTRACT

Within the paper the use of directly georeferenced terrestrial images for the realization of spatial queries by so-called telepointing is described. In order to enable an intuitive access to object related information by pointing to the respective image sections, a spatial model has to be mapped to an image of the user's environment. The camera used for image collection is integrated to a mobile device, which has to be capable for data transmission in order to provide the respective location based services. Since for the mapping of the spatial data the exterior orientation of the image has to be available, the device additionally integrates sensor for direct georeferencing. The provision of supplementary object related information by linking a spatial model to the perceived environment is a technique usually applied within augmented reality applications, which will also be discussed briefly. This type of application currently is one of the most promising markets for the use of spatial data. The work presented in this paper is part of the NEXUS project at Stuttgart University, which aims on the development of a platform for the access to location based services.

1. INTRODUCTION

Applications aiming on the provision of location-dependent access to spatial data are currently emerging to the consumer market. In order to provide information in the framework of a location aware environment, the actual position and orientation of the user has to be available. Hence, to reach the goal of an area-covering provision of location aware applications like personal navigation to any user anywhere, these systems have to be based on small, yet powerful, mobile devices, which integrate features like the accurate localization of the user as well as a location-dependent processing and transmission of spatial data via wireless networks.

Up to now photogrammetry has mainly benefited from the dynamic development triggered by these application scenarios due to the resulting demand for the collection and update of geographic data bases. Great efforts have been made on the development of automatic or at least partially automatic methods to enable an efficient and economical geographic data collection. One example is the large amount of standard GIS information on road network. This data is mainly used for navigation purposes, which is one of the most important applications in the context of location aware applications. Additionally, the emerging spread of location based services also stimulated the development of new products currently not provided by standard 2D GIS environments. One of the most obvious examples is the transition from 2D to 3D representations of the environment within future navigation systems. As it is depicted in Figure 1, current developments in car navigation systems aim on the superseding of 2D map-like representations by a 3D visualization of the environment. In this context a more realistic and intuitive presentation of the path to be followed by the user can be achieved. Since this type of visualization of course presumes the availability of a complete 3D model of the environment, this will result in an increasing demand for area covering 3D data collection mainly in urban environments.

Visualization of the 3D environment as well as the development of tools for the required data collection is of course not the only goal of current developments. Whilst a user moves through the real world, supplementary information related to the visible objects has to be provided. Hence, the intuitive access to spatial information is another key feature of location based services. In order to enable a very realistic presentation of object related information as well as a intuitive realization of spatial queries, so-called augmented reality techniques can be applied. By this technique computer graphics presenting object related information are projected to the user's current field of view.

Figure 1: Car navigation system BIRDVIEW based on 3D representation of the environment.

Since these systems have a considerable impact to the realization of location aware services, their principle as well as potential applications will be described briefly within the following section. Within our work, which is embedded in the research project NEXUS at Stuttgart University, terrestrial images provide the link between the visible environment and a so-called augmented world model. This augmented world model consists of a spatial model representing the real world, which is enriched by additional elements. These additional elements for example provide additional services or information. In order to link the spatial model to the image, the 3D world coordinates of the respective objects have to be mapped to the image. This process requires information on the exterior orientation of the collected image data. After the description of this process, the use of the oriented imagery for telepointing, i.e. for the realization of spatial queries by pointing to objects of interest at the displayed image will be presented in the final part of the paper.

2. AUGMENTED REALITY

Augmented reality (AR) is becoming one of the driving forces for the presentation of location dependent information. Generally speaking, it is based on the overlay of computer graphics to the user’s actual field of view by a see-through head mounted display. The computer graphics - e.g. the wire frame versions of actual objects like buildings enriched by supplementary information - are generated based on a spatial model of the visible environment and have to be directly overlaid to the corresponding primitives as they are observed by the user. The main advantage of applying head mounted displays is their capability to present the requested information in a hands-of manner without blocking the user’s view of the real world. Of course, the virtual computer graphic objects have to be correctly mapped to the real objects. Thus, in addition to the generation of the computer graphics, the accurate tracking of the actual position and orientation of the user to enable a precise mapping is one of the most demanding tasks in this context. Experimental AR prototypes have already been demonstrated for applications like manufacturing, city planning or image-guided surgery. Recent developments enable the construction of wearable computers based on commercially available hardware and software, therefore mobile augmented reality systems are becoming more and more affordable also for standard users.

One of the most valuable applications of AR, which can also be expected to become of considerable interest for the consumer market, is its capacity to provide situation awareness in build-up areas. This results from the fact that urban environments are usually complicated, dynamic, and inherently three-dimensional. Within an urban environment, AR can for example be applied for the presentation of name labels or additional alphanumeric data appearing to be attached to a side of a building.

In addition to the visualization of these virtual signposts, more specialized applications could aim on the display of information based on “X-ray vision” in order to present features normally not visible for the user. Typical objects of interest are features hidden behind the facades of a building like the location of rooms or information on infrastructure like the position of power-lines.

The integration of augmented reality into a tourist information system is another very promising application for this kind of technique. As an example, for the city of Heidelberg a mobile tourist information system was developed within the project Deep Map. Based on the 3D city model of the old town of Heidelberg a potential tourist can virtually walk through the 3D model in order to allow the planning of real tours preliminary to his actual visit. On-site, the visible environment is enriched by information relevant for each building. By these means queries on thematic information like opening hours of museums or the generation and overlay of historic views can be realized (Coors et al. 2000). A similar system, helping a user to navigate through a build-up area is also described by (Höllerer et al. 1999). Using a head mounted display, the names of buildings are presented to the user depending on his actual field of view. By pointing to the buildings additional information is made accessible via an integrated wireless access to the internet. This so-called telepointing feature is also realized within the project NEXUS (Fritsch et al. 2000), which is the basis of the work presented within this paper. For simplification of the overall system, the head-mounted display is replaced within the NEXUS environment by an image of the user’s environment. This image can for example be captured by a camera integrated into a small hand-held display. Nevertheless, the link between the augmented model and the observed environment still has to be generated.

3. TELEPOINTING BASED ON ORIENTED TERRESTRIAL IMAGES

If the exterior orientation is available, the augmented world data, i.e. the spatial model of the user’s environment enriched by additional objects, can be directly mapped to the captured image. After this step access to object related information can be realized by pointing to respective regions of interest directly on the image display (telepointing). The required mapping between virtual objects and the real world as it is depicted by the image must guarantee a correct overlay of the generated computer graphics to the corresponding objects in the perceived environment. For this reason position and orientation of the camera or the head mounted display, respectively, has to be determined at an sufficient accuracy.

3.1. Accuracy demands for exterior orientation

In general, the mapping accuracy, which is required for the presentation of supplementary information to an AR system is tightly coupled to the type of data to be integrated. The accuracy requirements are fairly low, if annotations like the name of a building and its known function are overlaid to the corresponding object in the real world, or if the translation of road signs are projected to the respective signposts. The same order of accuracy is also sufficient for the presentation of routing information, for example if the path that has to be followed to reach a particular destination is overlaid to the visible street. Although all this type of information is of course coupled to the user’s current position, a relatively coarse mapping is sufficient since the generated graphics have to be aligned only roughly to relatively large scale features.

In contrast, if the AR system shows highly localized information like the wire frame versions of buildings, a user cannot afford to see them far apart from the real edges of the observed objects. If the graphics are not exactly aligned in this type of application, the result will be annoying or possibly even misleading. Similar problems will occur if the presentation of otherwise hidden features such as the location of power lines, water supplies or other infrastructure and utility information is aspired. If the reconstructed virtual environment is directly overlaid to the observers view, even small errors in the model in the order of tens of centimeters can lead to significant errors, undermining the effectiveness of an AR system. For this reason all of this data can only be accurately registered to the corresponding observed object primitives, if both a detailed model of the environment and an accurate tracking system is available. For telepointing, as it is aspired in our application, the wire-frames do not have to be directly visualized. Thus, accuracies in the order of several decimeters for object to image mapping are sufficient. This can be achieved by direct georeferencing based on sensor components for medium accuracy.

3.2. Georeferencing of image data

For airborne imagery the exterior orientation usually is determined indirectly based on a spatial resection or a bundle block adjustment, respectively. If the image data is captured in a standard block configuration, commercial software tools are available, which enable an efficient and almost autonomous processing. In contrast to that, except for indoor applications in an industrial environment, the required procedures for highly automated tie and control point measurement do not exist for terrestrial images. Hence, direct georeferencing, i.e. the direct measurement of camera position and orientation at the time of image capture by a suitable sensor system, usually is the preferable solution for the processing of these outdoor scenes.

Figure 2: NEXUS georeferencing device

A commercial system, which integrates directly geocoded image sequences together with supplementary information into electronic city maps is described by Sood and Fahrenhorst (1999). Their system mainly aims on a coarse inspection of an area of interest based on the collected images, which are linked to a digital 2D map. Thus, the quality of their collected exterior orientation provided by GPS measurement is not sufficient for a task like telepointing. A system for the collection of georeferenced terrestrial images in urban areas at high accuracies using integrated DGPS/INS measurements is presented by Bosse et al. (2000). In their application the collected images are used for the subsequent measurement of building geometry. Therefore the accuracy demands to be met by their system hardware are considerable high and thus also would fulfill the specifications for precise object to image mapping. Alternatively to the application of very precise and expensive sensors for direct georeferencing, the accuracy requirements of the measured camera position and orientation can be reduced if a 3D model of the buildings at the site is already available. In that case the terrestrial imagery can be aligned to the reconstructed buildings by an automatic matching of corresponding primitives between object and image space as it is for example described by Jaynes (1999). For telepointing like it is aspired by our system, an accuracy of several decimeters up to a few meters is sufficient, since only the respective objects of interest have to be selected. Thus, direct georeferencing at medium accuracies is sufficient.

As it is depicted in Figure 2, our current NEXUS prototype consists of a camera mounted to a standard surveying device. The digital camera is equipped with a fire wire interface. This enables a simple synchronization of image capture and measurement of position and orientation. The georeferencing device combines a DGPS receiver, a digital compass and a tilt sensor module. The camera position is determined by a standard Garmin LP25 differential GPS receiver. By the application of the ALF service (**A**ccurate Positioning by **L**ow **F**requency) accuracies of 1-3 meters could be verified in a test area in the city of Stuttgart. Pose information is provided by a digital compass in combination with a tilt sensor module. The digital compass MAPSTAR measures the orientation at a nominal accuracy of 1°. The tilt sensor is integrated in the distance sensor LADIS and provides similar accuracies. The distance to the observed object, which also can be measured by the system is not integrated to the current data processing. Nevertheless, the telescopic sight, which is mounted to the device is used as view-finder during image collection.

3.3. Spatial Model

In order to provide a link between the visible environment and additional object related information as it is required for augmented reality and telepointing applications, a spatial model has to be available. In an urban environment the main component of a spatial model representing the perceived environment is provided by a 3D city model, which essentially contains the 3D wire-frames of the visible buildings. In addition to other applications in the context of visualization, 3D city models are also required for network planning or city climate and environmental research. Also triggered by the requirements of these applications, recent years have shown great effort in the development of tools for the accurate and efficient reconstruction of 3D city models. For this reason a rapidly growing number of data bases is becoming available. Even though the quality of these 3D city models is an important factor for the accuracy of the model to image mapping, the description of available approaches for the required 3D building reconstruction is beyond the scope of this paper.

For our test area in the city of Stuttgart two different 3D city model datasets are available. The first dataset was derived automatically at Stuttgart University based on a combination of laser DSM and groundplan information (Brenner 2000). The second data set has been collected manually by photogrammetric stereo measurement of images at scale 1:10.000 (Wolf 1999). This data has been provided by the City Surveying Office of Stuttgart. For both datasets the outline of each building is defined by the public Automated Real Estate Map (ALK), which provides accuracies in the centimeter level.

3.4. Model to Image Mapping

As it is depicted in Figure 3, the accuracy of camera position and orientation, which is directly measured by the available low-cost components of our georeferencing device, provide a good initial transformation of the building geometry to the collected terrestrial image. This step already enables a selection of the depicted buildings from the complete 3D city model. Telepointing, as it is required within our NEXUS project, is based on the definition of areas of interest within the captured image. For this purpose a coarse mapping of the building silhouettes to the image using the available directly measured exterior orientation is sufficient. Nevertheless, despite the fact that the aspired telepointing can be realized by the current system, the achieved accuracy is far from being optimal for other applications like the extraction of building texture from the captured image.

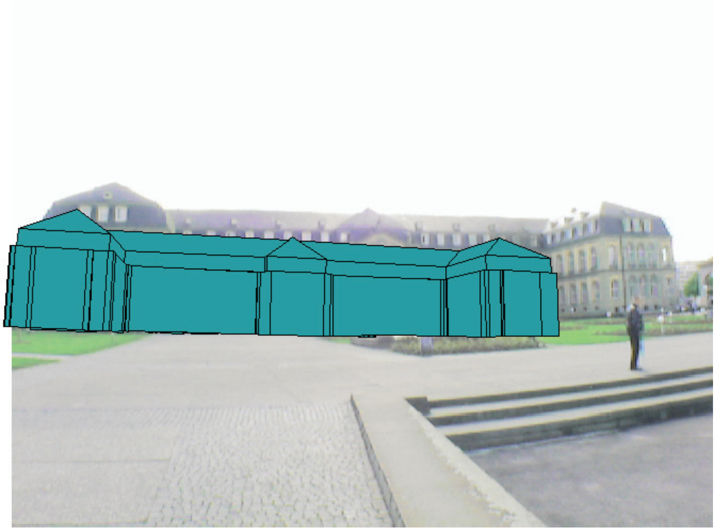

Figure 3: Projection of building based on GPS and digital compass measurement

One option to improve the mapping accuracy is the refinement of the camera pose based on a spatial resection using 3D coordinates of visible building primitives. At the moment, the required correspondence between object and image primitives are provided by manual measurement. Figure 4 gives an example of a refined transformation, which is based on the result of a spatial resection. Currently the refined transformation is used as reference measurement to evaluate the quality of direct georeferencing. By comparison of the results from both approaches an accuracy of the directly measured position of 1 - 2 m as well as an orientation accuracy of 1° - 3° could be verified for our georeferencing system.

Figure 4: Projection of building based on refined orientation from spatial resection.

Since there is a need for a further refinement of the model to image mapping process, current work aims on the automatic matching of corresponding primitives. These corresponding primitives can for example be detected by an automatic matching of GIS data to the collected image as it is described by Hild (2001). In our application the vector data will be provided from the respective 3D building model.

3.5. Examplary application

An exemplary application based on our current prototype is depicted in Figure 5. Simultaneously to the capture of the image the position and orientation of the camera is determined. Afterwards the viewing frustrum is calculated for the captured image and projected to an ortho image or a map. By pointing to a specific object of interest in the image, corresponding object related information as it is for example provided by a website is presented by the graphical user interface. These websites then give access to services like ticket sales if for example a theater is visible.

For demonstration of the telepointing functionality, the system is currently realized within a standard GIS software package. In the final system the NEXUS platform will provide both the management of the positioning components and the provision of the spatial models. A small mobile device (PDA) will be utilized as personal NEXUS station and information between platform and station will be exchanged by wireless communication.

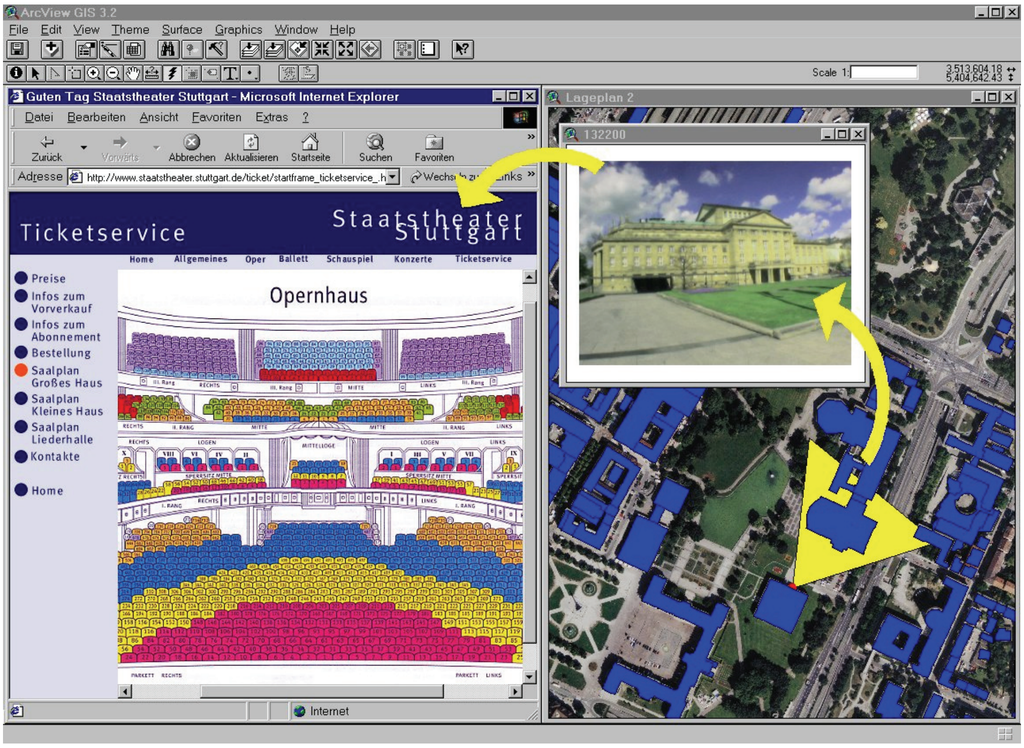

Figure 5: Example application of NEXUS telepointing prototype

4. CONCLUSIONS

Within the paper the use of oriented images as an interface to provide location based services has been demonstrated. Since the spatial model of the environment is mapped to the corresponding image, an intuitive access to additional object related information is feasible. Mapping accuracies of several decimeters to a few meters are sufficient for the described application, hence standard low-cost hardware can be applied for directly measuring position and orientation of the captured imagery. Similar applications of augmented reality like the direct overlay of the wire-frames of visible objects are much more demanding with respect to the required georeferencing precision. For this reason future research will aim on the integration of the available 3D city model into the orientation process to improve the mapping accuracy between spatial model and image. Another area of future interest is the generalization of the spatial model in order to improve its visualization by the relatively small devices, which generally have to be used for mobile applications.

5. REFERENCES

Bosse, M., De_Couto, D., and Teller, S. (2000): Eyes of Argus: Georeferenced Imagery in Urban Environments. GPS World April, 20-30.

Brenner, C. (2000): Dreidimensionale Gebäuderekonstruktion aus digitalen Oberflächenmodellen und Grundrissen. Reihe C, Heft Nr. 530, München.

Coors, V., Huch, T., and Kretschmer, U. (2000): TouriSTAR - ein mobiles AR System in einer urbanen Umgebung. Visualisierung raumbezogener Daten: Methoden und Anwendungen, Bd. II. 191-207. Münster, Institut für Informatik, Uni Münster.

Fritsch, D., Klinec, D., and Volz, S. (2000) NEXUS - Positioning and Data Management Concepts for Location Aware Applications. Proc. of the 2nd International Symposium on Telegeoprocessing. pp. 171-184.

Hild, H. (2001): Automatic image-to-map-registration of opticel remote sensing data. To appear in: Photogrammetric Week ‘01.

Höllerer, T., Feiner, S., Terauchi, T., Rashid, G., and Hallaway, D. (1999): Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system. Computers & Graphics 23, 779-785.

Jaynes, C. (1999): View Alignment of Aerial and Terrestrial Imagery in Urban Environments. ISD'99. 3-19. Springer-Verlag.

Sood, R.-A. and Fahrenhorst, C.(1999): Geocoded Image Sequences. Photogrammetric Week 99. 313-316. Wichmann Verlag.

Wolf, M. (1999): Photogrammetric Data Capture and Calculation for 3D City Models. Photogrammetric Week '99. pp. 305-313, Wichmann Verlag.