Available online at www.sciencedirect.com

SCIENCE DIRECT®

ISPRS Journal of Photogrammetry & Remote Sensing 58 (2003) 31–42

PHOTOGRAMMETRY
& REMOTE SENSING

www.elsevier.com/locate/isprsjprs

A multi-sensor system for positioning in urban environments

Norbert Haala*, Jan Böhm

Institute for Photogrammetry, University of Stuttgart, Geschwister-Scholl-Strasse 24, D-70174 Stuttgart, Germany

Received 8 July 2002; accepted 31 January 2003

Abstract

Within the article a low-cost system for the provision of georeferenced terrestrial images in urban environments is presented. Based on an image with approximate exterior orientation from a low-cost GPS and a digital compass and a 3D CAD model of a visible model as provided from a 3D virtual city model, the exact location of the building in the image is detected automatically and used for a refined orientation of the image. The work is part of a project aiming on the development of a mobile device, which enables access to location-based services in a complex urban environment. The intuitive access to object-related information is realized by so-called telepointing. For this purpose, a spatial model of the user's environment is mapped to an oriented image, allowing for the access to object-related information by pointing to the respective image sections. Since the provision of location-based services currently is one of the most promising markets for the use of spatial data in urban areas, these applications will also be discussed.

Keywords: navigation; silhouettes; building detection; image registration; image orientation; augmented reality; 3D city models; mobile systems; location-based services; multi-sensor systems

1. Introduction

The development of tools for the efficient collection of 3D city models, which are one of the most important databases in urban areas, has been a topic of intense research over the past years. A good overview on the current state-of-the-art of experimental systems and commercial software packages is given in Baltsavias et al. (2001). In addition to Digital Height Models and 2D GIS data representing streets and urban vegetation, 3D building models are the most important part thereof. Meanwhile, a number of algorithms based on 3D measurement from aerial stereo imagery or airborne laser scanner data are available for automatic and semi-automatic collection of 3D building models. Thus, an area covering provision of the required data is feasible. Originally, simulations for the propagation of electromagnetic waves used for the planning of antenna locations were the major application areas for 3D building models. Meanwhile, visualization in the context of urban planning has become the key market for this type of data. An increasing need for an area covering collection of 3D representations of the environment can also be expected due to emerging systems allowing for the 3D navigation of the user. As an example, such a navigation system, which is based on a very detailed city

---

* Corresponding author. Tel.: +49-711-121-3383; fax: +49-711-121-3297.

E-mail address: norbert.haala@ifp.uni-stuttgart.de (N. Haala).

0924-2716/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0924-2716(03)00015-7

model, is described by Rakkolainen and Vainio (2001). These systems substitute standard 2D map-like representations by a 3D visualization of the environment. By these means, especially for complex urban environments, a more realistic and intuitive presentation of the path to be followed by the user can be achieved. A further spread of this type of application can also be anticipated due to the emerging availability of small and increasingly powerful devices like Personal Digital Assistants (PDA). These devices will be able to transmit information at sufficient bandwidth, they will integrate additional sensors like digital cameras of GPS for self-localization and they will provide services for location aware applications. Thus, the need for the required urban data as well as the demand for suitable interfaces will be further stimulated.

The work presented within this article is part of a research project aiming on the development of a generic platform that supports location aware applications for mobile users. One of the key features for the provision of location aware services is the intuitive access to the requested information. Frequently, this type of information is related to objects that are visible within the user's environment. For this reason, the access to this localized information is realized by terrestrial images, which are collected by a camera integrated to the mobile device. Based on the exterior orientation of the captured images, the spatial model of the user's environment can be directly mapped to the corresponding sections of the image. After this step, access to object-related information is feasible by pointing to respective regions of interest directly on the image display.

The overlay of computer graphics representing object-related information to the user's current field of view is a standard feature within augmented reality (AR) applications. Since an emerging spread of these concepts can be expected, they will have a considerable impact on future standards also for spatial data in urban areas. For this reason, potential applications based on these concepts will be discussed briefly within the first part of the article.

In order to enable an access to object-related information by pointing to respective image sections (telepointing), the terrestrial image, as it is captured by the mobile device, has to be co-registered to the 3D model of the environment. Thus, in addition to the collection of the required spatial data, the accurate tracking of the actual position and orientation of the user to enable a precise co-registration of model and image is one of the most demanding tasks in this context. In our approach, the alignment is based on the direct measurement of exterior orientation by low-cost components. For refinement of this coarse measurement, the visible silhouettes of the depicted buildings are localized automatically in the image based on a Generalized Hough Transform (GHT). After this matching, checkpoints can be generated automatically based on the 3D coordinates of the visible building model and subsequently be used for the improvement of the exterior orientation by a spatial resection.

2. Image-based access to object-related information

For the provision of location-dependent information, a representation of the user's environment by spatial models is required. Furthermore, this real world model is supplemented by virtual objects. The resulting augmented world model is a common data model that includes the basic semantic for the provision of location-based information. The basic idea is that the user moves through a world of real objects like buildings, streets, cars, which is augmented by virtual objects. In this context, these virtual objects can be interpreted as metaphors for the availability of localized information. This access to object-related information based on the mapping of a virtual 3D representation of the user's environment to his real surrounding is a common concept used in augmented reality applications.

2.1. Augmented reality

Generally speaking, augmented reality (AR) applications can be characterized as an overlay of computer graphics to the user's actual field of view. Usually, in this manner order object-related information is presented by see-through head-mounted displays. In order to generate the computer graphics—for example, wireframe versions of visible objects enriched by supplementary information—a spatial model of the user's environment is required. In urban areas, a significant amount of the data is already contained in a 3D city model. Experimental AR prototypes have already been demonstrated for applications like manufacturing, city planning or image-guided surgery. A good overview on the current state-of-the-art is given in Azuma et al. (2001). 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 built-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. In this context, 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.

Other promising application arise from the integration of augmented reality techniques into tourist information systems. As an example, for the city of Heidelberg a mobile tourist information system was developed within the project Deep Map (Malaka and Zipf, 2000). Based on a 3D city model, a potential tourist can virtually walk through the old town of Heidelberg in order to plan 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 A similar system helping a user to navigate through a built-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.

2.2. Telepointing using georeferenced terrestrial images

In general, the main advantage of applying head-mounted displays for the presentation and access to object-related information is their capability to present the requested data in a hands-off manner without blocking the user's view of the real world. Nevertheless, intuitive access to localized information is still feasible, if for simplification of the overall system the head-mounted display is replaced by an image of the user's environment. This image can, for example, be captured by a camera integrated into a small handheld display. Telepointing, i.e. the access to object-related information by using an image of the environment, is also aspired within our research project NEXUS (Fritsch et al., 2000).

In order to allow an access to object-related information by pointing to respective image sections, the link between the augmented world model and the observed environment has to be generated. Thus, simultaneously to the capture of the image, the position and orientation of the camera is determined by direct georeferencing. Based on this information, a spatial model of the environment as represented by a 3D city model can be co-registered to the image. For our test area in the city of Stuttgart, two different 3D city model data sets are available, each covering several thousands of buildings The first data set was derived automatically at Stuttgart University by a combination of a dense laser DSM and ground-plan information (Haala and Brenner, 1999). 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 data sets, the outline of each building is defined by the public Automated Real Estate Map (ALK), which provides accuracies in the centimetre level. Even though the availability of a 3D city model at sufficient quality is an important prerequisite for our application, the description of available approaches for the 3D building reconstruction is beyond the scope of this paper.

An exemplary application of our current prototype is depicted in Fig. 1. Since the position and orientation of the user is available, the visible building is selected from the database and corresponding object-related information as it is, for example, provided by a website

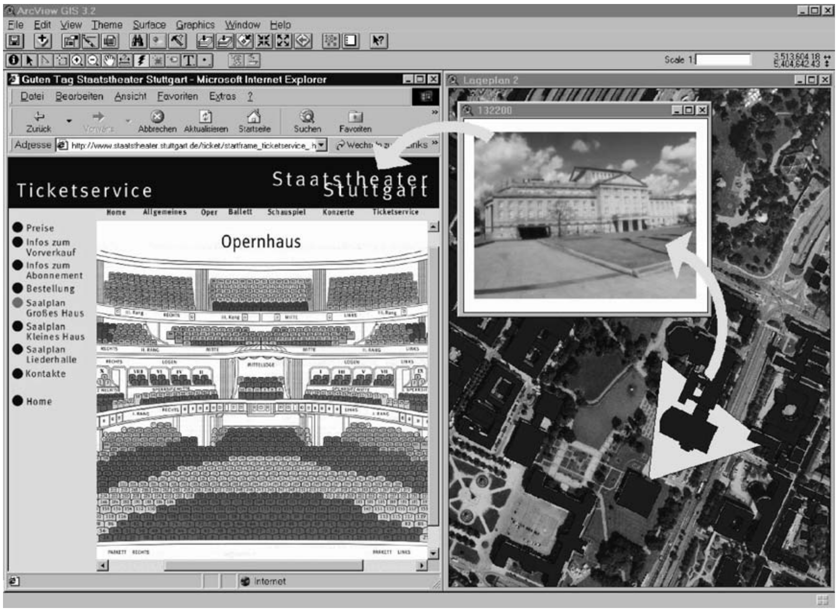

Fig. 1. Exemplary application demonstrating the access to object-related information.

is presented by the graphical user interface. These websites then give access to services like ticket sales if, for example, a theatre is visible. Additionally, the user's location and the selected building can be projected to an ortho-image or a map. For demonstration of the telepointing functionality, this application is currently realized within a standard GIS software package.

The mapping accuracy, which can be achieved using the directly measured parameters of the exterior orientation, is demonstrated in Fig. 2. For the given orientation and calibration of the camera, a rendered view of the depicted building is computed from the geometric representation as it is provided from the available 3D city model. Afterwards the outline of this rendered view, i.e. the visible silhouette of the building, is overlaid to the image. The deviations between model and image as depicted in Fig. 2 clearly demonstrate the limited accuracy of direct georefencing if low-cost components are applied. In this case, only a rough model to image mapping is feasible.

The platform we used for data collection is depicted in Fig. 3. It consists of a standard resolution color video camera with extreme wide-angle lens, a GPS receiver, an electronic compass and a tilt sensor. All the devices are connected to a laptop. While the camera and compass/tilt sensor are handheld, the GPS is attached to a backpack. In the final system, the 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 and information between platform and station will be exchanged by wireless communication. The camera used for our investigations is a consumer style Sony DFW-500 video camera, which is connected to the system via IEEE 1394 also known as FireWire. The camera was pre-calibrated in order to avoid problems due to lens distortions. The GPS receiver is a Garmin LP-25, which can be operated both in normal and differential mode. In our application, the ALF service (Accurate Positioning by Low Frequency) for differential mode providing a correction signal every 3 s was used. While the theoretical accuracy of differential GPS, as it is used in the prototype, is very high, there are a number of practical limitations when applying GPS in built-up areas. Shadowing from high buildings can result in poor satellite configurations, in the worst case the signal is lost completely. Additionally, signal reflections from buildings nearby can give rise to so called multipath effects, which are further reducing the accuracy of GPS measurement. Our experience shows that the system allows for a determination of the exterior

Fig. 2. Silhouette of the building as projected to the image based on the exterior orientation from GPS and digital compass.

Fig. 3. Prototype of the telepointing device.

orientation of the camera to a precision of 7–10 m in planar coordinates. In our system, the vertical component of the GPS measurement was discarded and substituted by height values from a Digital Terrain Model due to the higher accuracy of that data source. The zenith angle provided by the tilt sensor has an error of approximately 1–2°. The applied digital compass is specified to provide the azimuth with a standard deviation of 0.6–1.5°. However, compasses are vulnerable to distortion, because especially in built-up areas the Earth's magnetic field can be influenced by cars or electrical installations. These disturbances can reduce the accuracy of digital compasses to approximately 6° (Hoff and Azuma, 2000).

In addition to the precision of the available exterior orientation, the fit between model and image after co-registration is of course influenced by the geometric accuracy, level of detail and visual quality of the available 3D building models. Fig. 4 shows a rendered 3D view of the building model that was used to generate the silhouette overlaid to the image in Fig. 2. Even though the overall shape of the building is represented reliably, the amount of detail

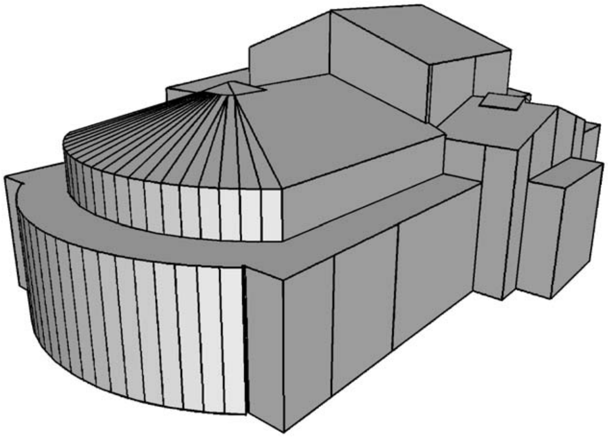

Fig. 4. 3D model used for generation of building silhouette in Fig. 2.

is limited especially for the facades of the building. This situation is typical for almost all available 3D city model data sets, since currently an efficient, area covering collection of these models is only feasible from airborne data. Thus, due to the restricted visibility of the building facades, the amount of detail is very limited for these parts of the building.

2.3. Accuracy demands

Of course, the mapping accuracy required for co-registration of real and augmented world is tightly coupled to the type of supplementary data to be presented during the aspired application. The accuracy requirements are fairly low, if like in the example depicted in Fig. 1, visible buildings have to be selected from an existing database. For this purpose a rough approximation of the actual viewing frustrum is sufficient. Based on a similar level of accuracy also annotations can be overlaid to the objects depicted in the image. By these means, for example, the name or the known function of a building can be presented. Of course all this type of information is coupled to the user's current position and viewpoint. Nevertheless, since the generated graphics have to be aligned only roughly to features of a relatively large scale, a relatively coarse mapping is sufficient for these types of application. In contrast to this, applications like the presentation of routing information for mobile users in complex urban environments will already require an increasing accuracy of co-registration. In this case, the path to be followed to reach a particular destination like the entrance of a building has to be overlaid to the visible environment as it is depicted by the image. For presentation of even more localized information like the wireframe versions of buildings, the accuracy demands are still higher. In such scenarios a user cannot afford if the graphics are not exactly aligned to the real edges of the observed objects. Otherwise, the result will be annoying or possibly even misleading for this type of application. Similar problems will occur during the presentation of otherwise hidden features of buildings such as the location of power lines, water supplies or other infrastructure and utility information. Generally speaking, if the reconstructed virtual environment is directly overlaid to the observer's view, even small deviations between model and real environment in the order of tens of centimetres can lead to significant errors, undermining the effectiveness of the system.

Generally speaking, the augmented world model can only be accurately registered to the corresponding observed object primitives, if both a detailed geometric model of the environment and an accurate tracking system is available. Thus, in addition to the collection of the required spatial data, the co-registration of image and model by an accurate image georeferencing is the most important task in this context. For this reason, also for our system aiming on the precise access object-related information by telepointing the model to image mapping as it is provided by direct georeferencing has to be refined.

2.4. Image georeferencing in urban environments

Spatial resection or bundle block adjustment is a well-known standard approach for georeferencing of images. Especially for the automatic orientation of airborne imagery, commercial software tools are available, which allow for an efficient and almost autonomous data processing. In contrast to this, the required procedures for highly automated tie and control point measurement do not exist for terrestrial images of natural outdoor scenes. 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 this type of application.

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 urban area based on the collected images, which are linked to a digital 2D map. Thus, the quality of the system's 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 sensors are considerable high and, thus, would also meet the specifications for precise object to image mapping. Still, the high costs of the hardware prevent the application for standard users. Alternatively, image sequences can be applied for retrieval of structure of the scene as well as reconstruction of the motion of the camera (Pollefeys et al., 2000). Nevertheless, in our application, only single images will be available from different viewpoints.

Alternatively to the application 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 this case, the terrestrial imagery can be aligned to the reconstructed buildings by an automatic matching of corresponding primitives between object and image space. A system for the alignment of terrestrial images to a 3D city model based on an approximate exterior orientation from GPS and digital compass has been discussed by Jaynes and Partington (1999). A similar approach, which is also based on the matching of linear features of the building facades, is described by Coors et al. (2000). In their approach a very detailed model of the building facade, including features like windows, doors or arches, is required. As already discussed, this amount of detail cannot be presumed for standard data sets. Similarly, the lack of sufficient facade texture information in standard 3D city models usually prevents the use of image correlation techniques for model to image mapping. For this reason, our approach aims on the detection of the overall shape of the building within the image.

3. Localization of building shapes

In order to locate a 3D object in an image, two general strategies can be applied. Firstly, the inherent 3D representation of the object can be matched to the image, which leads to a 3D to 2D matching problem. Secondly, a 2D representation of the object can be applied, which leads to a 2D to 2D mapping problem. While the first approach is the more general and theoretically more appealing, practical problems can occur due to the complexity of the matching task. To allow for a simplified 2D to 2D matching, a suitable representation has to be provided for the 3D shape of the object. One option is to decompose the object shape into several possible views and store the resulting 2D representations in a so-called aspect graph. Since in our application the exterior orientation of the imaging device is available approximately, the generation of the whole aspect graph is not required. Instead, a single 2D view is created on-the-fly for each image in correspondence to the respective orientation. Thus, based on the a priori knowledge, the model to image matching is simplified to a 2D to 2D problem.

Additionally, object recognition requires the extraction of suitable features, which have to be available both in the image and the model data. In our data set, the buildings are modelled as polyhedrons, no in-plane facade detail or texture information is available. This strong discrepancy in feature detail between model and image data prevented us from using edge or corner detection. As already demonstrated in Fig. 2, the overall shape of the building represented by the silhouette can be calculated based on the approximate exterior orientation and the 3D model of the building. For a refined mapping, this representation can be detected and exactly localized within the corresponding image. For this purpose a Generalized Hough Transform (GHT) is applied.

3.1. Generalized Hough transform

The Hough transform is a well-known technique for isolating and locating features of a particular shape within an image. Since the desired features have to be specified in some parametric form, the *classical* Hough transform is most commonly used for the detection of regular curves such as lines, circles, ellipses, etc. (Hough, 1962). The GHT (Ballard and Brown, 1982) is a generalization of this concept for the employment in applications, where a simple analytic description of a feature is not possible. In this case, instead of using a parametric equation of the curve, a look-up table is applied to define the relationship between the boundary positions and orientations and the Hough parameters.

During the off-line phase, which has to be computed once for each 2D prototype shape, a look-up table, the so-called *R*-table is generated. In our application this object shape is defined by the silhouette of the building as it is calculated from the approximate orientation. First, an arbitrary reference point $x_{\mathrm{ref}},y_{\mathrm{ref}}$ is defined. Usually, the centroid of the shape is selected for this purpose. As it is depicted in Fig. 5, the shape of the feature can then be defined with respect to this point of reference by the distance r and angle β of normal lines drawn from the boundary. In order to generate the *R*-table, first, for every point of the prototype shape, the edge orientation ω is calculated. The *R*-table then consists of distance and direction pairs r and β, which are indexed by the orientation ω of all points along the boundary of the given shape. The Hough transform space is now defined in terms of the possible locations of the shape in the image, i.e. the possible range of the centroid's position $x_{\mathrm{ref}},y_{\mathrm{ref}}$.

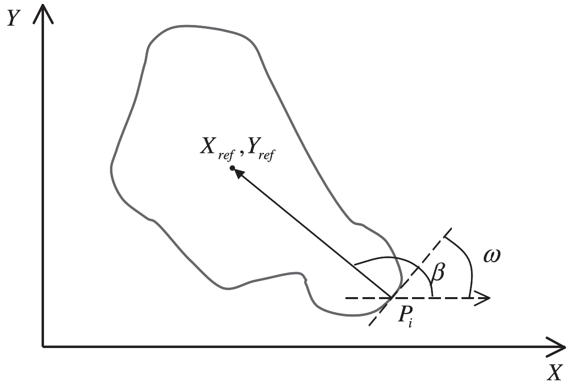

Fig. 5. Representation of prototype shape as recorded in the framework of the Generalized Hough Transform.

In the on-line phase, for localizing of the shape in the collected image, gradients are computed by an arbitrary edge operator within the search image. After this step for each edge pixel i at position x_i,y_i, the orientation ω_i is available. Based on the available orientation ω_i now for each edge pixel $p_i(x_i,y_i)$, the corresponding values for r_i and β_i can be selected from the generated *R*-table. These values are then used to calculate the position $x_{\mathrm{ref}},y_{\mathrm{ref}}$ of the prototype shape in the image by

$$x_{\mathrm{ref}} = x_i + r\cos\beta_i$$

$$y_{\mathrm{ref}} = y_i + r_i\sin\beta_i. \tag{1}$$

During processing now each edge pixel votes for a certain position $x_{\mathrm{ref}},y_{\mathrm{ref}}$ of the given shape. Each vote is represented by an update of the accumulator array at this calculated position. The position in the image receiving the most votes at the end is selected as the position of the shape in the image. If the orientation of the object is allowed to vary, as is the case in our application, a separate *R*-table has to be computed for each discrete rotation angle. The same is true for scaling, which has also to be determined in our system. Thus, the formation of the *R*-tables in this case is quite complex and computationally expensive.

3.2. Matching of building silhouettes

For our implementation the HALCON image processing environment was used, which provides a shape detection mechanism based on the GHT as it is discussed in the previous section (Ulrich et al., 2001). In order to compensate for the computational

Fig. 6. Detected silhouette of the building.

costs of large *R*-tables, this operator includes several modifications to the original GHT. As an example, it uses a hierarchical strategy generating image pyramids to reduce the size of the tables. By transferring approximation values to the next pyramid level, the search space is drastically reduced. Additionally, the expected accuracy of the shape's location can be applied for a further reduction of the search space.

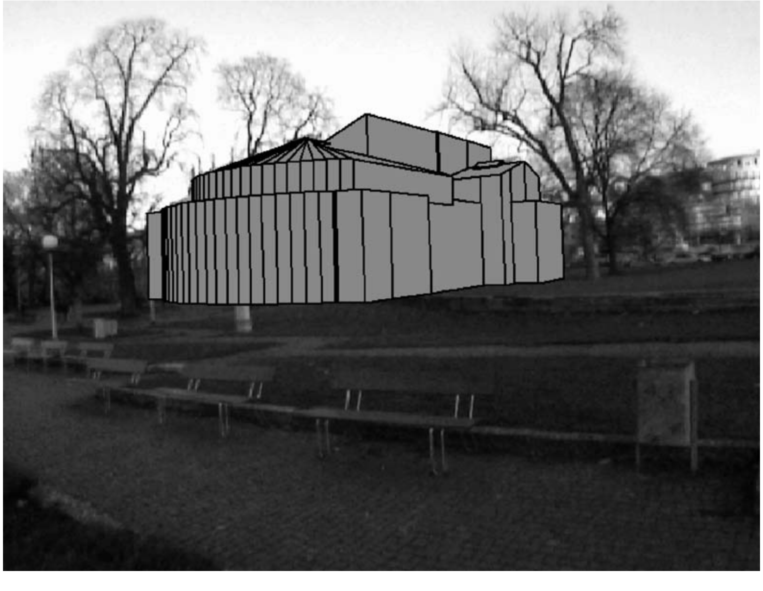

Fig. 7. Building model projected to the image after refinement of orientation.

Fig. 6 shows the silhouette of a building as automatically detected by the GHT within the captured image. Based on the estimated parameters for shift, rotation and scale within this process, the approximate image coordinates of the visible object model already shown in Fig. 2 can now be improved. After the refinement of the shape's position in the image, corresponding points in object and image space are available. These control points can now be applied in order to improve the original exterior orientation by a spatial resection. In principle, the complete process, extraction of building silhouette, improvement of image coordinates by GHT and spatial resection then has to be iteratively repeated in order to avoid errors resulting from the simplification of the original 3D to 2D matching to a 2D to 2D

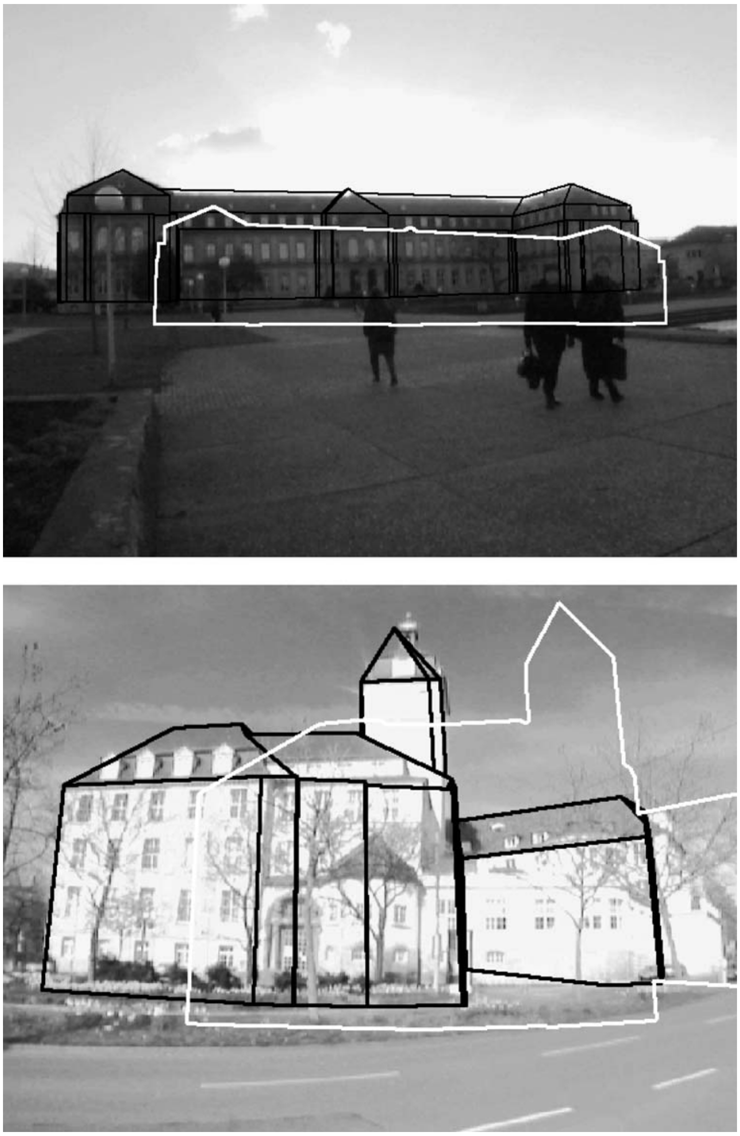

Fig. 8. Additional examples of shape matching. The building silhouette from direct georeferencing is given by the white polygons, the black lines represent the wireframes after refined model to image mapping.

problem. Nevertheless, for our application the differences between the projected wireframe and the image were mainly caused by errors of the available 3D building model. Thus, this iteration was not applied. In addition to measurement errors during the collection of the 3D city model, these errors mainly result from generalization effects. Fig. 7 shows the original 3D building model projected to the image after orientation refinement.

Two other examples of the algorithm for different buildings are given in Fig. 8. In these examples, the silhouettes of the building models as they are provided from GPS and digital compass measurement are overlaid to the images as a white polygons. The results of the GHT for refined mapping are represented by an overlay of the building's wireframes which are represented by the black lines.

4. Discussion

Within the paper the use of oriented images as an interface to provide location-based services has been discussed. By co-registration of a spatial model to an image of the environment, an intuitive access to object-related information is feasible for a user by pointing to the respective image sections. The exterior orientation of the camera is measured directly by standard low-cost hardware, allowing for mapping accuracies of a few meters. Even though his accuracy can be sufficient for some applications, it has to be improved if highly localized information has to be presented. For this purpose, the successful implementation of a fully automated process for the detection of buildings in terrestrial images has been demonstrated. Based on the GHT, the shape of the depicted building can be detected no matter whether it is shifted, rotated or even scaled in relation to the image. The GHT also allows for a certain tolerance in shape deviation. This is essential, since available CAD models usually provide a generalization of the building's actual shape as it is appears in the image. However, the accuracy of the position and orientation measured by GPS and digital compass, respectively, must be sufficient to guarantee that the shape of the predicted outline does not differ significantly from the actual building image.

The detection of the overall shape of buildings of course requires a sufficient distance of the user to the depicted building. Nevertheless, this can be presumed for applications like telepointing or navigation. Still, it is to be expected that an area covering provision of position information can only provided by hybrid systems. In this connection, tagging techniques provide an option already used for indoor applications. By this approach, precise object identification and location is realized by a tag fixed to an object, which sends a unique ID, i.e. via infrared signals.

Another important application requiring a precise georeferencing of terrestrial images in urban environments is the refinement of available 3D city models. Whereas building geometry can usually be provided effectively from airborne data, the collection of facade geometry including doors, windows and other unmodelled structure as well as the extraction of facade texture currently is a current bottleneck during data collection. Thus, the alignment of terrestrial images is an important prerequisite in order to facilitate an efficient enhancement of existing urban databases.

References

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B., 2001. Recent advances in augmented reality. IEEE Computer Graphics and Applications 21 (6), 34–47.

Ballard, D.H., Brown, C.M., 1982. Computer Vision. Prentice-Hall, Englewood Cliffs, NJ.

Baltsavias, E., Grün, A., van Gool, L., 2001. Automatic Extraction of Man-Made Objects From Aerial and Space Images (III). A.A. Balkema, Lisse.

Bosse, M., De Couto, D., Teller, S., 2000. Eyes of Argus: georeferenced imagery in urban environments. GPS World 4, 20–30.

Coors, V., Huch, T., Kretschmer, U., 2000. Matching buildings: pose estimation in an urban environment. IEEE and ACM International Symposium on Augmented Reality. IEEE CS Press, pp. 89–92.

Fritsch, D., Klinec, D., Volz, S., 2001. NEXUS—positioning and data management concepts for location aware applications. Computers, Environment and Urban Systems 25, 279–291.

Haala, N., Brenner, C., 1999. Virtual city models from laser altimeter and 2D map data. Photogrammetric Engineering and Remote Sensing 65 (7), 787–795.

Hoff, B., Azuma, R., 2000. Autocalibration of an electronic compass in an outdoor augmented reality system. Proc. of International Symposium on Augmented Reality. Technische Universität München, pp. 159–164.

Höllerer, T., Feiner, S., Terauchi, T., Rashid, G., Hallaway, D., 1999. Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system. Computers & Graphics 23 (6), 779–785.

Hough, P.V.C., 1962. Method and means for recognizing complex patterns. U.S. Patent 3.069.654.

Jaynes, C., Partington, M., 1999. Pose calibration using approximately planar urban structure. Proc. of Asian Conference on Computer Vision. Available at http://www.metaverselab.org/pub/paper/accv-jaynes-00.pdf (accessed 21 January, 2003).

Malaka, R., Zipf, A., 2000. DEEP MAP—challenging IT research in the framework of a tourist information system. Proc. of 7th International Congress on Tourism and Communications Technologies in Tourism, Springer Computer Science, pp. 15–27.

Pollefeys, M., Koch, R., Vergauwen, M., van Gool, L., 2000. Automated reconstruction of 3D scenes from sequences of images. ISPRS Journal of Photogrammetry and Remote Sensing 55 (4), 251–267.

Rakkolainen, I., Vainio, T., 2001. A 3D city info for mobile users. Computers & Graphics, Special Issue on Multimedia Appliances 25 (4), 619–625.

Sood, R.-A., Fahrenhorst, C., 1999. Geocoded image sequences. In: Fritsch, D., Hobbie, D. (Eds.), Photogrammetric Week '99. Wichmann, Heidelberg, pp. 313–316.

Ulrich, M., Steger, C., Baumgartner, A., Ebner, H., 2001. Real-time object recognition in digital images for industrial applications. In: Gruen, A., Kahmen, H. (Eds.), Optical 3D Measurement Techniques V. Wichmann, Heidelberg, pp. 308–318.

Wolf, M., 1999. Photogrammetric data capture and calculation for 3D city models. In: Fritsch, D., Hobbie, D. (Eds.), Photogrammetric Week '99. Wichmann, Heidelberg, pp. 305–312.