

Towards pedestrian navigation and orientation

¹Darko KLINEC, ²Dieter FRITSCH

Institute for Photogrammetry
University of Stuttgart

¹Phone : ++49 711 1214116, Fax : ++49 711 1213297
Email : darko.klinec@ifp.uni-stuttgart.de

²Phone : ++49 711 1213386, Fax : ++49 711 1213297
Email : dieter.fritsch@ifp.uni-stuttgart.de

ABSTRACT

At the Institute for Photogrammetry (ifp), Stuttgart University, current research aims at the development of tools and applications for Location Based Services (LBS). As location of information is a key parameter of these applications, the current user's position and orientation is essential for information access. Usually, human beings navigate and locate themselves by integrating viewpoint information (landmarks) into a coherent mental representation of the traversed environment. But in urban environments with difficult and complex building structures natural orientation fails very often, because the navigator's viewpoint does not encompass the environment in its totality. Thus, intuitive automatic positioning and navigation becomes necessary.

In this paper our approach for positioning, integrating multi sensors and images will be described in more detail pointing out the process and aspects for automation. Results will be presented and a discussion for integrating this approach for pedestrian navigation in LBS will be given.

KEYWORDS

NEXUS, Location-based, image processing, spatial resection

INTRODUCTION

In the last years mobile systems and services become more and more public to support mobile users. With an emerging availability of small and powerful devices new types of applications becomes feasible. In the meantime simple LBS are offered from various sides for users of cell phones. They enable us to access on data stored on remote sys-

tems or to stay in contact with other persons. People can get guiding support to the next restaurant, and even friend finder functionality is provided. For these applications the mobile devices (cell phones) can be localized by use of the cellular phone network or GPS (Global Positioning System). To support users by overlaying the reality with additional information or to interact with surrounding objects, these techniques are not sufficient. There are several techniques in Augmented Reality (AR) to solve these problems but often they are limited for use in indoor, as the environment has to be equipped with appropriate sensor technology. In outdoor we can't assume the same prepared environment like indoors. The approach being developed at the Institute for Photogrammetry therefore takes into account a model of the environment to support the process.

LOCATION AWARE APPLICATIONS AND INFORMATION ACCESS

If we have a look on the market we will notice that there are several simple information systems providing information dependent on the actual location. As these are commercial systems all of them are using a proprietary architecture for data storage and data access. Each company stores information in their proprietary data format. Access on data of different information providers and exchange are difficult.

Imagine the possibilities if an open platform for location aware applications would be available, where everybody can contribute information to the model. Therefore one challenge of the research within the project NEXUS is the provision of such an open platform – the NEXUS system. This system relies on a model-based concept, called the NEXUS Augmented World Model (AWM) [Nicklas et al. (2001)]. The AWM is the base for the NEXUS system's extensibility and flexibility and it forms the interface to the applications. As it is an open platform also existing data sources like the WWW shall be integrated. This may lead to a great heterogeneity in the data. In NEXUS a federation approach is used to handle that heterogeneity, see Fig. 1.

A closer view to the object oriented AWM shows its basic idea: federation of information and representation of the real world (buildings, streets, cars,...). As example one representation of the real world could be a detailed 3D city model. For interaction with the AWM and the use of NEXUS services the architecture provides also an interface for sensor integration. Here different positioning sensors can be plugged in to provide the necessary position and orientation information to several applications.

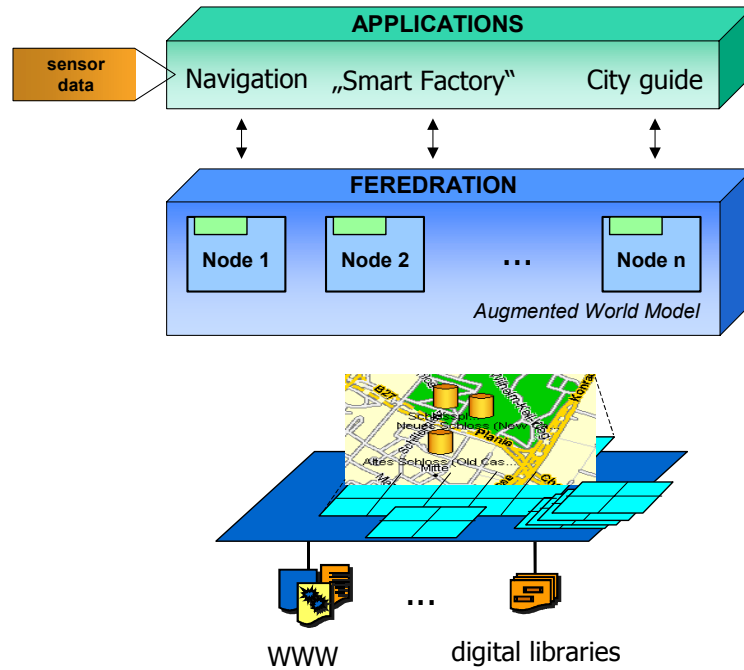


Figure1: Architecture of the NEXUS platform

LOCATION SENSING TECHNIQUES

As location aware applications need position information different techniques can be used to provide location information. In principle there are three methods for automatic location sensing: triangulation, proximity and scene analysis [Hightower & Borriello (2001)]. Triangulation techniques use geometric properties of triangles to compute object locations. This method is divisible into the subcategories of lateration, using distance measurements, and angulation, using primarily angle measurements. As example systems like GPS or a magnetic compass use triangulation techniques. When an object is “near” a known location this is described with proximity. There are three general approaches to sensing proximity. (a) Detecting physical contact, (b) monitoring wireless cellular access points and (d) observing automatic ID systems. The scene analysis technique uses features of a scene observed to conclude to the location of the observer or of objects within the scene. Usually the observed scenes are simplified to obtain features that are easy to represent and compare.

The position and orientation of a user are basic information to provide high quality location based services. Azuma et al. [Azuma et al. (1999)] evaluated that all tracking systems lack accuracy or robustness. This leads to his conclusion that only a combination of different technologies, which he call a hybrid tracking system, should be used. Several systems are using GPS and devices to track the orientation. But often the existing approaches are not accurate enough for the overlay of a reconstruction of real world objects or they cannot be applied to persons walking in a city without the requirement of

fixed positions where they should stand to receive information. The Touring machine of Columbia University in New York is one of the most well known information systems [Julier et al. (2000)]. It broadcasts information about names of the buildings, so positioning and orientation accuracy are not the most important conditions. To provide information about more specific building features the requirements on position and orientation accuracy are higher. The use of image processing techniques is a method to improve results. Beveridge et al. [Beveridge et al. (1996)] and Behringer [Behringer (1999)] used horizon shapes extracted from a visual scene to look up the observer's location from a prebuilt dataset. In You et al. [You et al. (1999)] an augmented reality system is described which tries to correct drift errors of a gyroscope and errors of the compass using images collected by an additional camera.

Pedestrians usually use viewpoint information and landmarks (e.g. buildings) to locate themselves in a familiar environment. To facilitate the same in an unfamiliar environment an automation of this process should be provided by combining the model of the environment and surrounding objects (landmarks). In the research project NEXUS a model of the real world is existing – the Augmented World Model. This model contains various information as well as a 3D representation of real world objects (e.g. buildings). On that condition we are able to integrate divers sensors and operations on these data for pedestrian orientation and navigation.

SCENE ANALYSIS FOR ORIENTATION AND NAVIGATION

With the increased availability of low-cost and low-power consumption imaging devices it seems worthwhile to investigate their integration potential into mobile devices. The combination of mobile computational capabilities, imaging capabilities, positioning devices and network access opens a door for novel applications. As example Augmented Reality applications in urban regions are useful to assist users and to interact with the model of the environment. Especially in NEXUS where a model of the environment exists information about objects can be provided by identifying them. In order to allow information access or to use objects as landmarks for navigation, the link between the Augmented World Model and the observed environment has to be generated. As we concentrate on image information this leads to an approach which tries to overlay the model and the corresponding primitives in the real world. Similar approaches often use a database containing georeferenced photos. There captured photos and the images stored in the database were matched. We do not use such comparison methods but a 3D city model. The base information for our approach is provided by simultaneously capture of the image and direct georeferencing of position and orientation. Using this information the Augmented World Model can be co-registered to the image. For automation of co-registration, image processing techniques and methods for image orientation

are applicable, which extract and detect prominent features in the model in the one hand and in the image on the other hand [Azuma et al. (1999)]. As example spatial resection or bundle block adjustment are well known approaches for georeferencing of images. In our approach we use a shape matching followed by a modified spatial resection. The workflow in Figure 2 depicts details of the process for automated georeferencing of terrestrial images. Required input data for the process are a 3D city model and collected images with known rough exterior orientation. Based on these data objects of interest are identified and the correspondent silhouettes are created. Afterwards a Generalized Hough Transform (GHT) is utilized for detection of the object's position within the image, by estimating the best match [Haala et al. (2002)]. In other words the result of the GHT is a two-dimensional similarity transformation. It is important to point out that after this step only knowledge about relationship in image space, but not between 2D-image and 3D-object space is available. Therefore, in a second step a spatial resection algorithm (SRA) is utilized for co-registration. As the detailed description of the whole process for co-registration will go beyond the scope, the following section will point out two main aspects. First the preparation process for extraction and correlation of straight lines will be discussed and second the feature based spatial resection.

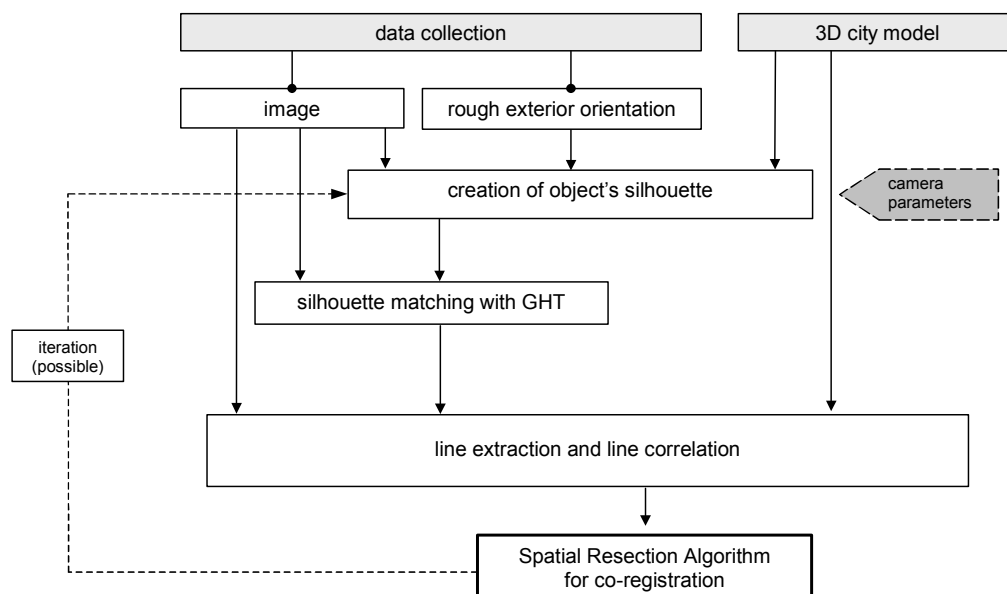


Figure 2: Georeferencing of terrestrial images using object features

Feature extraction

Systems for object recognition require a definition of features used for recognition and algorithms to perform the matching. The decision on the feature type is usually guided by the available model data. In our case we want to recognize buildings, which are modeled as a set of polyhedrons. To achieve a robust detection we decided to detect the object's silhouette, which is a good approximation of its shape. For recognizing the

shape within the image, first it must be rendered from an existing 3D CAD database. For this purpose we assume that the person is standing close to the position estimated by GPS and the user's line of sight approximately corresponds to the direction collected by the orientation module. On that condition an object of interest is selected [Klinec&Fritsch (2001)] and its silhouette is rendered, according to the camera parameters and the CAD model data. An example of a back projected silhouette using rough orientation parameters is shown in Figure 3. After rendering, the silhouette representation is detected in the scene by a silhouette matching based on a generalized Hough Transform. Figure 4 shows the result of a GHT based silhouette matching. More details on how the silhouette matching is working are given in [Haala et al. (2002)]. For further processing the silhouette is an initial feature, as it defines a region for extraction of straight lines needed by the spatial resection algorithm.

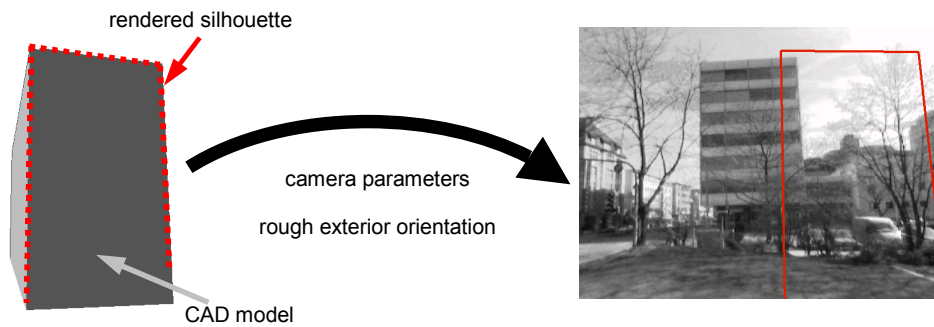


Figure 3: Silhouette back projection using rough orientation parameters

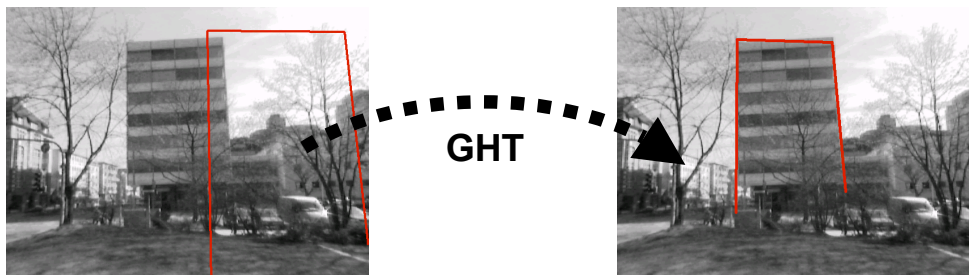


Figure 4: Silhouette matching using GHT

Extraction and correlation of straight lines

We are using straight lines as input data for spatial resection; therefore they have to be extracted within the image. In the literature several approaches for straight-line extraction are given. In our approach we decided to use the conventional Hough Transform [Hough (1962)] to solve this problem. However extraction of straight lines does not solve the whole problem. For each extracted line in the image, the corresponding line in the three dimensional CAD model has to be identified. Usually the detection of three dimensional primitives to corresponding primitives in the two dimensional image space

(image) would lead to a 3D-to-2D matching problem. As approximations of the exterior orientation of the camera and the position of the object's silhouette (determined by GHT) are existing the problem can be reduced to a 2D-to-2D matching. The result of the GHT is a two dimensional similarity transformation consisting of translation, rotation and scale. This offers the possibility to project a selected edge from the CAD model onto the image using the camera's exterior and interior orientation parameters and move the edge to its nearly correct location by similarity transformation estimated by GHT. Now the location of this edge in the image is well known and a buffer in the image around this line can be defined. Searching for the straight line within this extracted image region, using the conventional Hough Transform, gives us as result the correlation to the correspondent edge. In this way a stepwise correlation for each selected edge of the 3D model can be provided. A result of the described process is shown in Figure 5, where the automatically extracted and correlated lines are displayed in red color.

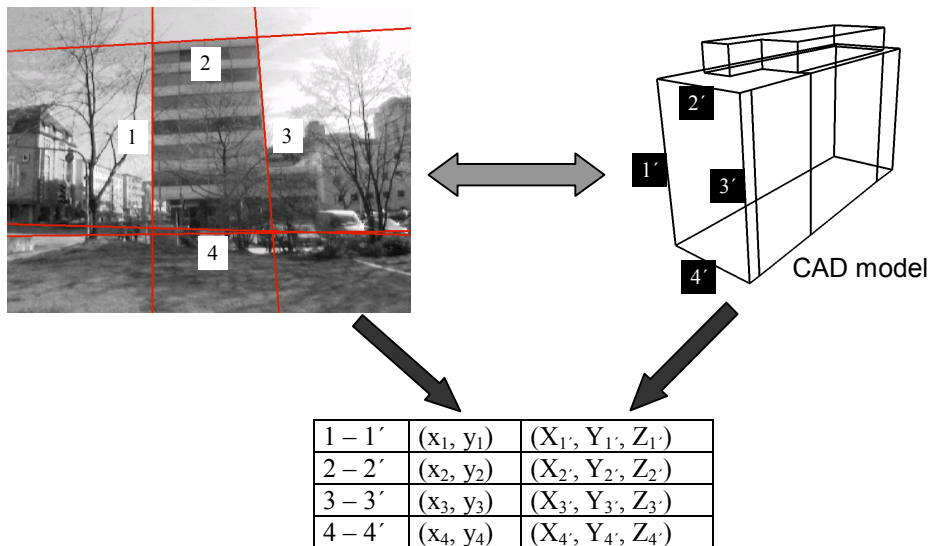


Figure 5: Automatically extracted and correlated lines

Feature based spatial resection

Based on the result of the correlation process the exact co-registration of image and object space can be provided by spatial resection. In principle the algorithm overlays the extracted lines in the image and the corresponding edges of the 3D CAD model using the camera parameters as restrictions. This feature based spatial resection process is illustrated in Figure 6.

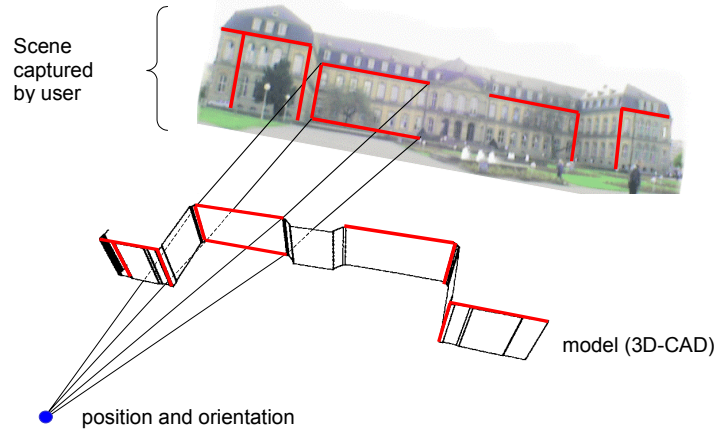


Figure 6: Feature based spatial resection

The implemented algorithm for feature based spatial resection first starts with a transformation of the 3D CAD edges given with initial point and end point, into a parameterized representation using equations (1) and (2). For parameterization of the lines extracted in the image, equations (3) has to be used. To avoid singularities in equations (1), (2) and (3) case differentiations are implemented within the algorithm and dependent on them modifications of the equations are used. For spatial resection the known collinearity equations, e.g. given in [Kraus (1996)] are modified using the parameterizations (1), (2) and (3). This results in equations (4) and (5), which we implemented in our approach. A more detailed mathematical description can be found in [Schwermann (1995)].

$$Y = \alpha_x \cdot X + \gamma_x \quad Z = \beta_x \cdot X + \delta_x \quad (1) / (2)$$

$$y = \alpha_x \cdot x + b_x \quad (3)$$

$$a_x = \frac{r_{31}(Y_0 - \alpha X_0 - \gamma) + r_{21}(\delta - Z_0 + \beta X_0) - r_{11}(\alpha(\delta - Z_0) - \beta(\gamma - Y_0))}{r_{32}(-Y_0 + \alpha X_0 + \gamma) + r_{22}(Z_0 - \delta - \beta X_0) + r_{12}(\alpha(\delta - Z_0) - \beta(\gamma - Y_0))} \quad (4)$$

$$b_x = x_h \frac{r_{31}(-Y_0 + \alpha X_0 + \gamma) + r_{21}(Z_0 - \delta - X_0 \beta) + r_{11}(\alpha(\delta - Z_0) - \beta(\gamma - Y_0))}{r_{32}(-Y_0 + \alpha X_0 + \gamma) + r_{22}(Z_0 - \delta - \beta X_0) + r_{12}(\alpha(\delta - Z_0) - \beta(\gamma - Y_0))} + y_h \quad (5)$$

$$+ c_k \frac{r_{33}(-Y_0 + X_0 \alpha + \gamma) + r_{23}(Z_0 - \delta - X_0 \beta) + r_{13}(\alpha(\delta - Z_0) - \beta(\gamma - Y_0))}{r_{32}(-Y_0 + \alpha X_0 + \gamma) + r_{22}(Z_0 - \delta - \beta X_0) + r_{12}(\alpha(\delta - Z_0) - \beta(\gamma - Y_0))}$$

To solve the spatial resection problem, we implemented a least squares algorithm using the Gauß-Markoff-Modell. Here the unknown parameters (X_0 , Y_0 , Z_0 , ω , φ , κ) are estimated and the co-registration of objects and image is determined.

In principle the complete process, extraction of silhouette, improvement of its position

in the image by GHT, extraction and correlation of straight lines and spatial resection can be repeated iteratively in order to avoid errors resulting from the initial orientation data and to improve the result stepwise.

RESULTS

To investigate the feasibility of our approach several examples were chosen. The investigations of our approach are based on a 3D CAD dataset of the city of Stuttgart, provided by the City Surveying Office of Stuttgart. This 3D CAD city model was created by manual photogrammetric stereo measurements from images at 1:10.000 scale [Wolf (1999)]. For creation of 3D city models also fully automatic methods are existing [Brenner (2000)]. We decided to use the dataset of the City Surveying Office, as a large amount of detail is available as well as the accuracy is high (centimeter level).

Table 1 shows different results. The used dataset is corresponding to the building in Fig. 3 to Fig. 5. Tabular displayed are results of our feature based spatial resection and a spatial resection using manually selected points as input data. To detect the quality of the automatic line extraction also a prepared dataset for optimized line detection was used. As we can see the simulated dataset and the real dataset have nearly the same result when using our feature based spatial resection. Compared to the estimation using points as input data for spatial resection the result is also nearly the same. As the point based estimation represents the optimal result, we can conclude that a position error of $\sim 0,2\text{m}$ occurs by using our line based approach. Normally the results of both algorithms should be the same as they are based on the collinearity equations. In the test case the error using the line based spatial resection occurs dependent on the input data. Based on the resolution and quality of the used images the line detection lack on accuracy and also the result of the line based spatial resection.

Table 1: Positions determined by spatial resection using points vs. straight lines as input data

		<i>Results (spatial resection)</i>		
		Points	Straight lines	
	Rough exterior orientation (collected data)	real dataset (manually selected tie points)	simulated dataset (optimized line detection)	real dataset
X	3512910.62	3512904,45 m	3512904,66 m	3512904,78 m
Y	5405089.17	5405072,13 m	5405072,41 m	5405072,13 m
Z	258.15	252,46 m	252,28 m	253,12 m
ω	10.63	9,9°	10,1°	9,0°
φ	64.7	61,4°	63,0°	61,8°
κ	0.0 (predefined)	-2,56°	-3,4°	-1,91°

DISCUSSION

In this article we have indicated that “simple” systems for location aware applications are commercially available, but they often lack on accuracy if the orientation of users has to be estimated for information access or navigation purposes. A digital compass can help to estimate the orientation, but especially in urban areas it suffers by electro magnetic fields. However for pedestrian navigation we need a system, which is able to detect the orientation and position reliably. The orientation is an important information as a pedestrian can change the orientation more flexible than e.g. a vehicle, which is bound on traffic rules. Also a system for pedestrian navigation should be intuitive to use. These requirements include a different use of sensor systems as they are used e.g. in vehicle navigation. Usually navigation is understood as a service to be guided from point A to point B using the “best” path. Pedestrian navigation requires also to be guided to special object points (e.g. entrance of a building) and more abstract to be “guided” to additional object information. Comparing these requirements and the methods for location sensing (triangulation, proximity and scene analysis) the scene analysis represents a method for intuitive positioning and object detection. This is one reason why we decided to implement the scene analysis for orientation and navigation.

When we look at the research results of our approach we can see that position and orientation can be estimated with sufficient accuracy. That offers the possibility to use these concepts for supporting location aware applications. Also the approach demonstrates a tool that combines 3D representations of the environment to access on information or to overlay information to captured images.

CONCLUSION

The NEXUS concept focuses on the development of a global platform in order to support different kinds of location-based applications (access on information, navigation, ...). As a prerequisite for the realization of location-based applications, the position and orientation of the users has to be known. Within the article we have demonstrated the successful implementation of a fully automated process for estimating the position and orientation of an “observing” camera (mobile user) using images and a 3D city model. Also we have indicated the possibility to use the described approach for information access and navigation purposes. To improve the reliability and the mapping accuracy between spatial models and images further research will aim at the robust integration of 3D city models into the orientation process as well as their preparation.

ACKNOWLEDGEMENTS

The research described in this paper is founded by “Deutsche Forschungsgemeinschaft”

(DFG – German Research Foundation). The research takes place within the Center of Excellence N^o 627 “NEXUS – SPATIAL WORLD MODELS FOR MOBILE CONTEXT-AWARE APPLICATIONS”, at University of Stuttgart.

REFERENCES

- Azuma R., Hof B., Neely III H., Sarfaty R., Daily M., Bishop G., Chi V., Welch G., Neumann U., You S., Nichols R., and Cannon J. (1999), Making Augmented Reality Work Outdoors Requires Hybrid Tracking , in: *Augmented Reality: placing artificial objects in real scenes: proceedings of IWAR '98*, San Francisco, Calif., 219-224.
- Behringer R. (1999), Registration for Outdoor Augmented Reality Applications Using Computer Vision Techniques and Hybrid Sensors, in: *Proceedings of IEEE Virtual Reality 1999*, Houston, Texas, USA.
- Beveridge J. R., Graves C. R. and Leshner C. E. (1996), Local search as a tool for horizon line matching, in: *Image Understanding Workshop*, Los Altos, CA, ARPA, Morgan Kaufmann, 683-686.
- Brenner C. (2000), Dreidimensionale Gebäuderekonstruktion aus digitalen Oberflächenmodellen und Grundrissen, DGK, Reihe C, Nr. 530, München.
- Haala N., Böhm J. And Klinec D. (2002), An integrated approach for the orientation of terrestrial outdoor scenes, in: *International Archives on Photogrammetry and Remote Sensing IAPRS*, Volume XXXIV, Com. II, ISPRS Commission II Symposium, Xian, 157-162.
- Hightower J. and Borriello G. (2001), A survey and Taxonomy of Location Systems for Ubiquitous Computing, Technical Report UW-CSE 01-08-03, University of Washington.
- Hough P.V. C. (1962), Method and Means for recognizing complex patterns, U.S. Patent 3,069,654.
- Julier S., Lanzagorta M., Baillot Y., Rosenblum L., Feiner S. and Höller T. (2000), Information Filtering for Mobile Augmented Reality, in: *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR)*, Munich, Germany.
- Kraus K. (1996), *Photogrammetrie 1*, Dümmler Verlag, Bonn
- Nicklas D., Großmann M., Schwarz T., Volz S. and Mitschang B. (2001), A model based Open Architecture for Mobile, Spatially Aware Applications, in: Jensen C.S. (ed.), Schneider M. (ed.), Seeger B (ed.), Tsotras V. (ed.): *Proc. of the 7th International Symposium on Spatial and Temporal Databases: SSTD 2001*, Redondo Beach, CA, USA.
- Schwermann R. (1995), *Geradengestützte Bildorientierung in der Nahbereichsphoto-*

togrammetrie, Dissertation, Nr. 52, Geodätisches Institut der Rheinisch-Westfälischen Technischen Hochschule Aachen.

Wolf M. (1999), Photogrammetric Data Capture and Calculation of 3D City Models, in: Photogrammetric Week '99, Stuttgart, 305-313.

You S., Neumann U. and Azuma R. (1999), Orientation Tracking for Outdoor AR Registration, in: Proceedings of IEEE Virtual Reality 1999, Houston, Texas, USA.