PROCESSING OF 3D BUILDING MODELS FOR LOCATION AWARE APPLICATIONS

Norbert Haala*, Jan Böhm, Martin Kada

Institute for Photogrammetry (ifp), University of Stuttgart, Germany
Geschwister-Scholl-Strasse 24D, D-70174 Stuttgart
Norbert.Haala@ifp.uni-stuttgart.de

KEY WORDS: Navigation, Orientation, Augmented Reality, Generalisation, Visualisation

ABSTRACT:

Within the paper the integration of 3D building models in the context of location aware services is presented by two exemplary applications. First the intuitive access to object related information within an urban environment is by telepointing is described. For that purpose terrestrial images are mapped against a 3D model of the environment based on direct georeferencing and a image based localisation of depicted buildings. The oriented image then allows for the access to information related to the depicted objects by pointing to the respective image sections. Secondly, the paper presents an approach for the generalisation of 3D building models with regard to the visualisation of urban landscapes. Especially for applications such as web-based 3D GIS and the presentation of virtual walk-throughs on mobile computing devices, these generated representations of the 3D models are a prerequisite in order to generate level of detail structures for real-time rendering.

1. INTRODUCTION

The development of tools for the efficient collection of 3D city models has been a topic of intense research for the past years. In addition to DTMs and data representing streets and urban vegetation, 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. A good overview on the current state-of-the-art of experimental systems and commercial software packages is for example given in (E. Baltsavias et al. 2001).

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 three-dimensional car navigation systems, virtual tourism information systems or city and building planning has become the key market for that type of data. In our opinion one of the most important development-driving forces for the application of 3D city models is the widespread use of mobile devices for the provision of location based services. For a typical application like personal navigation features such as a realistic visualization of the 3D urban environment in real-time or the accurate localization of the user have to be made available. As it will be discussed in the article, the processing of 3D building models plays an important role while providing these features

Due to the limited amount of storage capacity and computational power of the available hand held devices and their small display sizes on the one hand and the huge amount of data contained within a 3D city model on the other hand, the amount of information to be handled, stored and presented, has to be reduced efficiently. Thus, the generalization of the 3D building models becomes a topic of major interest within that context. Secondly, the provision of location based information like the usage of buildings or the location of rooms, which are hidden behind the faces of a visible building presumes exact knowledge on the actual position and orientation of the user. The work presented within this paper is based on our research project NEXUS, which aims on the development of a generic platform that supports location aware applications with mobile users. For this type of application the world is represented by spatial models, which describe the real world supplemented by virtual objects. This so called augmented world model is a common data model that bears the basic semantic for location based information. The basic idea is that the user lives in a world of real objects (buildings, streets, cars…) which is augmented by virtual objects. These virtual objects are metaphors for external information and act as brokers between the platform and external information sources and services. One of the most important parts of this augmented world model is a 3D city model, which has to provide a highly detailed representation of the world within an urban environment. Our investigations are based on a dataset of the city of Stuttgart provided by the City Surveying Office of Stuttgart. This data was collected by manual photogrammetric stereo measurement from images at 1:10000 scale (Wolf 1999). For data collection the outline of the building from the public Automated Real Estate Map (ALK) was additionally used. Thus, a horizontal accuracy in the centimetre level as well as a large amount of detail could be achieved.

After a short introduction to the requirements of the aspired location aware applications in section 2, the orientation of terrestrial images based on direct georeferencing and the automatic detection of visible building silhouettes will be discussed in section 3. These images are used in order to provide an intuitive interface to object related information. The automatic generalization of 3D building models which is a prerequisite for real-time visualization of urban environments will be presented in section 4.

2. LOCATION AWARE APPLICATIONS

The provision of object related, localized information is a key feature of location aware applications. One option to reach this goal is the application of augmented reality (AR), which 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 -

---

* Corresponding author

are generated based on a spatial model of the visible environment. In our application this data is provided by the 3D city model. Of course the virtual computer graphic objects have to be correctly overlaid to their corresponding primitives in the real world as they are observed by the user. For this reason the accurate tracking of the actual position and orientation of the user in order to enable a precise mapping of the data is required.

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 application for this kind of technique. As an example, for the old town of Heidelberg a mobile tourist information system has been developed (Malaka & Zipf 2000). Within this system preliminary to his actual visit a potential tourist can virtually walk through the 3D city model to allow the planning of real tours. 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 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. Additionally, by pointing to the buildings supplementary information is made accessible via an integrated wireless access to the internet.

This so-called telepointing feature is also realized within our research project NEXUS (Fritsch et al. 2000), which is the basis of the work presented within this paper. For simplification of the overall system, in this application 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 hand-held display.

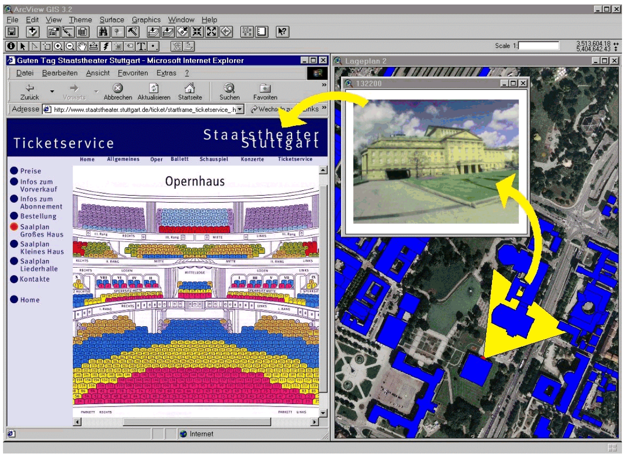

Figure 1: Prototypical telepointing application

An exemplary application based on our current prototype is depicted in Figure 1. Simultaneously to the capture of the image, the position and orientation of the camera is determined by a DGPS receiver and a digital compass. This information is sufficient to calculate the viewing frustrum for the captured image, which then can be 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 theatre is visible. Currently, the system is 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.

3. INTEGRATED ORIENTATION OF TERRESTRIAL IMAGES

In order to enable a precise mapping of the 3D building model to the captured image, an accurate tracking of the imaging device is required. Based on this information the access to the augmented world data, i.e. the spatial model of the user’s environment enriched by additional objects is feasible. This can then be realized by pointing to respective regions of interest directly on the image display. As it it realised in our system the user’s actual position can be provided by the use of a small DGPS receiver and a digital compass in outdoor areas. Nevertheless, in situations where the provided accuracy is not sufficient, the already captured image can additionally be used for a further improvement of image georeferencing.

3.1 Directly Measured Exterior Orientation

The accuracy of the exterior orientation as provided by the available DGPS receiver could be verified to several meters for the positional accuracy, whereas the orientation accuracy as provided by the digital compass and a tilt sensor resulted in an error of 1° – 2°. Figure 2 depicts the application of these low-cost components for the selection of the visible building as well as the initial transformation of the building’s wire-frame to the directly geoereferenced terrestrial image.

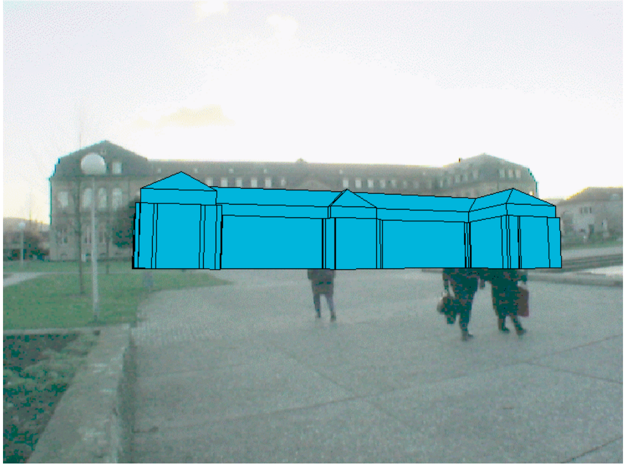

Figure 2: Projection of building based on DGPS and digital compass measurement

Even though this coarse mapping is sufficient for some applications, it has to be refined if highly localized information has to be presented to the user. For this purpose an automated appearance-based detection of buildings in terrestrial images is applied. The problem is stated as follows: From an image with a given approximated exterior orientation and a three-dimensional CAD Model of the building, detect the exact location of the building in the image and use this information in order to improve the directly measured exterior orientation of the camera

3.2 Building Localization

When the task is to detect a three-dimensional shape in an image, two general strategies for object representation are available. One is mapping of the inherent three-dimensional representation of the object, which leads to a 3D to 2D matching problem. Alternatively, two-dimensional representations can be applied, which leads to a 2D to 2D matching problem. While the former is the more general and theoretically more appealing approach, there are several practical problems which often prevent its use. One of the problems is the reliability of feature extraction, the other the exponential complexity of the matching task. For the later approach, in order to have a two-dimensional representation of a three-dimensional shape, one has to decompose the shape into several views and store a two-dimensional representation for each view. This approach is referred to as an aspect-graph. For our system it is not required to build the whole aspect graph, since an approximated exterior orientation of the imaging device is available. In this case, a single view of the shape can be created on-the-fly for each image in correspondence to the respective orientation data.

Additionally, when designing a object recognition system one has to choose the type of features used for recognition. The decision on the feature type is often guided by the available model data. In our case the buildings are modelled as polyhedrons, no in-plane facade detail or texture information is available. This strong discrepancy in feature detail in-between model and sensor data, thwarts the use of edge or corner detection. Since there is no texture information available, also image correlation is not an option.

To achieve a robust detection our system aims on the detection of the overall shape of the building in the image rather than extracting single features. The intent was, that the overall shape is more robust against clutter of the scene, partial occlusion by trees, cars, pedestrians and other negative influences during image capture. A good representation of the overall shape of a building is provided by its silhouette. Thus, based on the existing CAD database, the 3D building model is rendered according to the interior and exterior orientation of the camera used for the collection of the actual image. This 'virtual view' of the building as it is depicted in Figure 2 is used to extract the silhouette of the building. Now, this representation has to be located within the corresponding image. For this purpose a Generalized Hough Transformation (GHT) as described by (Ballard & Brown 1982) was applied.

3.3 Generalized Hough Transformation

Generally speaking, the GHT provides a framework for both the representation and detection of two-dimensional shapes in images. Based on the GHT a shape can be detected no matter whether it is shifted, rotated or optionally even scaled in relation to the image. These degrees of freedom are required since the orientation is only known approximately in our application. Additionally the GHT allows for a certain tolerance in shape deviation. This is also necessary, since the CAD model of the building provides only a coarse generalization of its actual shape as it is appearing in the image.

The Hough transform is a technique which can be used to isolate features of a particular shape within an image. Because it requires that the desired features 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. Compared with this, the generalized Hough transform can be employed in applications, where a simple analytic description of a feature(s) 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. In our application the prototype shape used to compute the look-up table values during a preliminary phase is provided by the silhouette of the building as depicted in Figure 2. First, an arbitrary reference point x_{ref}, y_{ref} is defined within the feature. The shape of the feature can then be defined with respect to this point by the distance r and angle β of normal lines drawn from the boundary to the reference point. The resulting look-up table will consist of these distance and direction pairs r, β, indexed by the orientation ω of the boundary.

The Hough transform space is now defined in terms of the possible positions of the shape in the image, i.e. the possible ranges of x_{ref}, y_{ref}. In other words, the transformation is defined by:

$$\begin{aligned} x_{ref} &= x + r\cos\beta \\ y_{ref} &= y + r\sin\beta \end{aligned} \qquad (1)$$

An arbitrary edge operator provides edges pixels at position x_i, y_i with orientation ω_i for the image. Based on the generated look-up table for the available orientation ω_i the corresponding values for r_i and β_i can be selected. Thus based on equation (1) the accumulator array can now be updated for each edge pixel by the calculated position x_{ref}, y_{ref}. If – as in our case – in addition to the position, the orientation and scale of the feature are also unknown, separate accumulator arrays have to be generated.

For our implementation the HALCON image processing environment was used (Ulrich et al 2001). In order to compensate for the computational 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 an further reduction of the search space.

Figure 3: Detected silhouette of the building

Figure 3 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, the approximate image coordinates of the visible object model as depicted in Figure 2 can now be improved. By application of these corresponding points in object and image space, the original exterior orientation can be improved by spatial resection. Figure 4 shows the 3D building model projected to the image based on the refined orientation from this process.

Figure 4: Improved mapping of the building model

In principle, the complete process - extraction of building silhouette, improvement of image coordinates by GHT and spatial resection – 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 problem. Nevertheless, for our application the differences between the projected wire-frame and the image were mainly caused by errors within the available model due to measurement errors or generalisation effects. Thus this iteration was not applied.

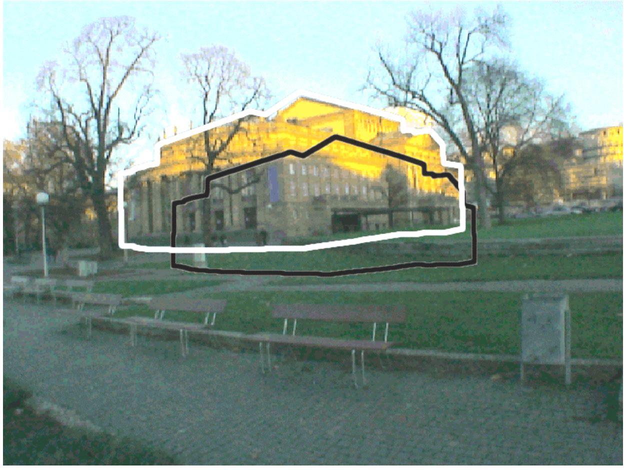

Figure 5: Additional example of shape matching. Silhouette from direct measured exterior orientation is given in black, refined localization is given in white.

Figure 5 gives an other example of the matching process. The silhouette of the visible building as derived from the available 3D model and the directly measured parameters of exterior orientation is depicted in black, whereas the of the image based refinement is outlined in white.

The detection of the overall shape of buildings as it is demonstrated by these examples of course requires a sufficient distance of the user to the depicted building. Since we are aiming on applications like telepointing or navigation, this can be presumed for most cases. Additionally, an area covering provision of position information within a complex urban environment is only feasible based on hybrid systems. Thus in our opinion different sensors and techniques and sensors will be applied for different user scenarios. As an example, tagging techniques can be employed for objects of interest, which are already used for indoor applications. By these systems object identification and location is realised by a tag fixed to an object, which sends an unique ID i.e. via infrared signals.

In addition to the access to object related information, precisely georeferenced terrestrial images in urban environments can also be applied for the refinement of available 3D city models. Whereas building geometry can usually be provided effectively from airborne data, the collection of façade geometry including doors, windows and other unmodeled structure as well as the extraction of facade texture currently is a current bottleneck during data collection. Thus, the automatic alignment of terrestrial images as it is feasible by our approach is an important prerequisite to facilitate an efficient enhancement of existing 3D city models.

4. GENERALISATION OF 3D BUILDING MODELS

During personal navigation, which is one of the main tasks within location based services, the visualization of the environment and the generation of a virtual walk through for planning of actual tours are features of great importance. Due to the small displays of a mobile device, the amount of information to be presented has to be reduced for this purpose. Hence, an automatic generalization of the 3D building models to be presented to the user has to be made available.

Since a building representation by planar faces and straight edges is feasible for most cases, the reconstructed buildings are usually described by general polyhedrons. Hence, for real-time visualization the number of faces to be displayed for each building object has to be reduced considerably. In general, this process presumes the elimination of unnecessary details, whereas features, which are important for the visual impression have to be kept. Especially for man-made-objects like buildings, symmetries are of major importance. For this reason, during the process of generalization the preservation of regular structures and symmetries like parallel edges, perpendicular intersections or planar roof faces has to be guaranteed.

In our approach a simplification of polyhedral building models is achieved by combining techniques both from cartography and computer graphics. In cartography a lot of effort has already been spent on the generalisation of 2D building structures. (Sester 2000) for example uses least squares adjustment for the generalization of building ground plans. Approaches for 3D object generalization have only be proposed recently (Mayer 2000). On the other hand, surface simplification is a widely used technique in the field of computer graphics in order to speed up the visualization of highly complex 3D models (Heckbert & Garland 1997). Usually, surface simplification is applied to general objects, which are either given as polygonal or as triangular surface meshes. Usually the elimination of edges for object simplification is only controlled by geometric properties. Symmetry considerations, which are important for the visual impression of objects like buildings are not taken into account. These symmetries and regularities are stringently preserved during generalization by our approach by integration of a set of surface classification and simplification operations.

The initial step of the generalisation algorithm is to build the so-called constrained building model, which represents the regularization constraints between two or more faces of the polyhedral building model. In the following steps the geometry of the constraint building model is then iteratively simplified by detection and removal of features with low significance to the overall appearance of the building. During this feature removal step the constrained building model is applied in order to preserve the represented building regularities and optimise the position of the remaining vertices.

4.1 Regularization Constraints

In most cases walls are oriented in parallel to the principal axes of the building, which are again often rectangular. It can therefore be assumed that most faces of a building model are coplanar, parallel and rectangular to other faces in the same model. In order to preserve these symmetry relations between faces as well as possible during generalisation, these properties have to be integrated. As this information is usually not explicitly available the constraint building model is constructed. Basically it consists of the polygonal building model enriched by a set of regularization constraints. The lowest element the hierarchy of constraints is the coplanarity constraint, which simply groups a set of faces together, each being coplanar to any other face in the same set. Two faces are assumed to be coplanar if the angle between the normal vectors is close to 0° or 180° and their distance is bellow a given threshold. Sets of coplanar faces are then again grouped together by a parallelism constraint if their faces are parallel to faces of another face set. Finally, two or three sets of coplanar or parallel faces are grouped by a rectangularity constraint if the faces of each set are rectangular to faces in the other two or three sets.

It is our belief that not every topological relation can be detected reliably by an automatic approach. Dependent on the quality of the input model, a number of constraints will almost always be missed due to errors introduced in the generation of the model. Those absent constraints might reduce the quality of the final model if missed in high quantities. An application should therefore offer the possibility to identify and insert more constraints into the constraint building model in a semi-automatic fashion to work around those errors and to improve the overall quality of the final building model. A semi-automatic tool also helps to test the effects of certain constraints on the generalisation process by manually adding or removing those constraints. Thus, in the current algorithm manual improvement of the automatically detected constraints is enabled.

4.2 Model Simplification

As already discussed, purely geometric considerations are not sufficient, if a simplification of objects like buildings is aspired. If the geometry of a building model is simplified by arbitrary removing vertices or edges, the symmetry of the building will irretrievably be disturbed, even if the introduced geometric error is small. In order to preserve the regularity of the model during generalisation our feature detection and removal algorithm allows the use of a manifold set of surface simplification operators, each designed to remove one specific class of feature types. In contrast to rather simple operators used in traditional surface simplification algorithms, our operators remove entire features in one continuous process, while preserving the integrity of remaining parts of the building model.

Figure 6 depicts the three classes of features which are currently distinguished: extrusion, notch and tip. After feature detection small features, which are of low importance to the appearance of the building are removed by a combination of edge collapse and edge foreshortening operations.

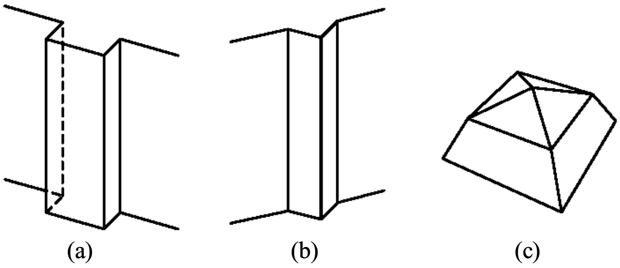

Figure 6: Detected features (a) extrusion, (b) notch and (c) tip.

During feature removal, parts of the building are completely eliminated. The optimal shape of the reduced model, however, should still be determined by all original points, even though if the number of points is reduced in the preceding step. During simplification sets of coplanar faces will for example often be merged. Nevertheless, vertices which are eliminated from the building should still be used to define the overall size of the simplified model. In order to calculate the final shape of the simplified model the available constraints like coplanarity, parallelism or rectangularity between the remaining faces as well as all the points of the original model have to be integrated. In our approach this is realised by application of least squares adjustment.

4.3 Results

The algorithm has been implemented and tested on polygonal building models of a 3D building dataset. The complexity of building models, measured by the number of triangles uses to represent the building could in most cases be reduced by more than 30%.

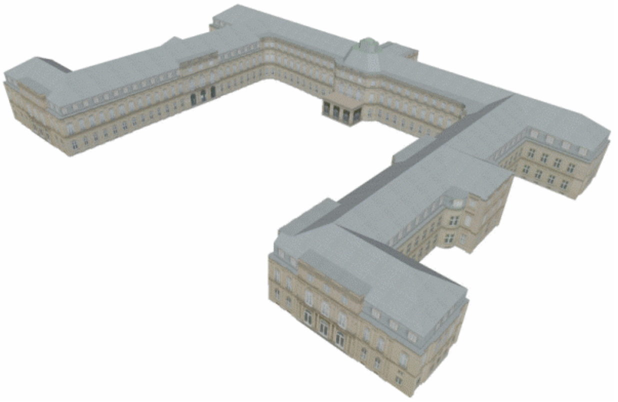

Figure 7: Exemplary object model consisting of 2730 triangles

As an example, the model of the New Palace of Stuttgart depicted in Figure 7 originally consists of 2730 triangles. By removal of extrusions the number could be cut down to 1837 triangles. The result of our algorithm to this model is demonstrated in more detail in Figure 8 to Figure 11. Figure 8 and Figure 10 show a part of the original model as it was captured from stereo imagery and an existing outline from the public Automated Real Estate Map (ALK), respectively. Figure 11 shows the result of the generalisation process. It is clearly visible, that parallelism and rectangularity have been preserved for the remaining faces. Especially if the model is textured again, as it is depicted in Figure 9, this amount of detail is sufficient for realistic visualization even at close distances.

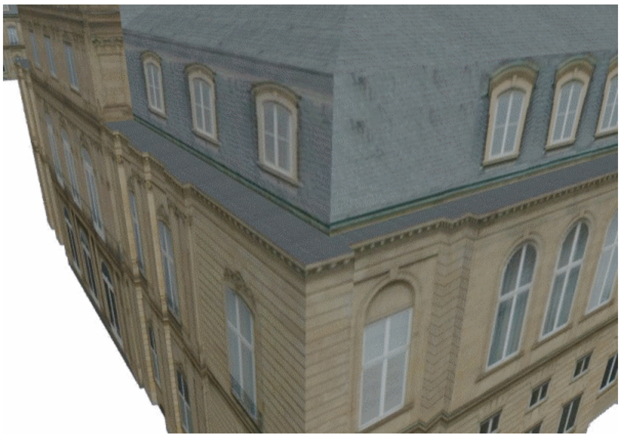

Figure 8: Part of original building model (with texture)

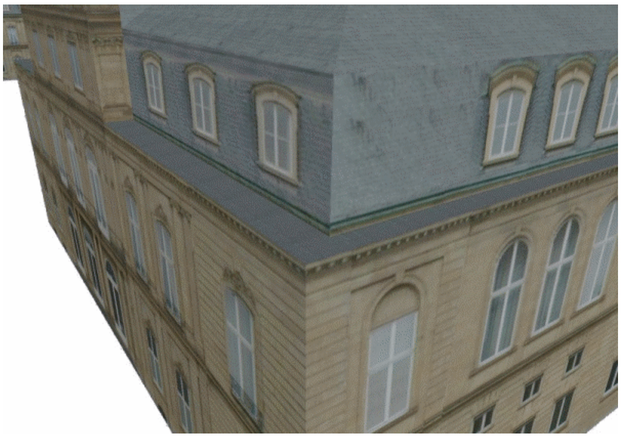

Figure 9: Part of simplified building model (with texture)

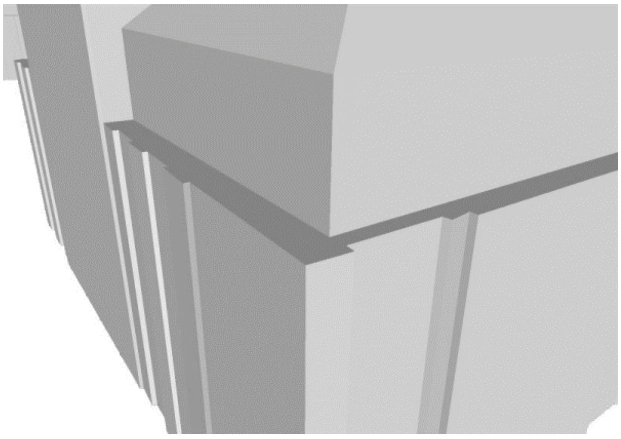

Figure 10: Part of original building model (without texture)

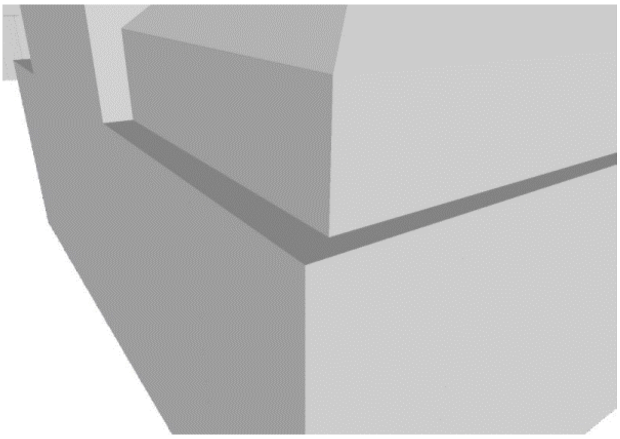

Figure 11: Part of simplified building model (without texture)

5. CONCLUSION

In addition to the further improvement of tools for automatic collection of virtual city models, it is our believe that there is an increasing demand for algorithms allowing for an improved usability of this type of data. One option is the application of 3D building models in the context of location aware applications as described in the paper. Based on the co-registration of the virtual city model to an image of the environment, the access to object related information can be realised. For this application, the exterior orientation of the image was refined by an automatic detection of the building silhouette. Additionally, the current transition from 2D map-like representations to more realistic 3D visualisations will require the availability of tools for generalisation of 3D data. Algorithms like they have been discussed in the second part of the paper are especially necessary, if a good interpretability of the presented 3D data is aspired at very different scales.

6. REFERENCES

Ballard, D.H. & Brown, C.M. (1982): Computer Vision, Prentice-Hall Inc., 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 Publishers

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

Heckbert, P. S. & Garland, M. (1997), Survey of Polygonal Surface Simplification Algorithms. SIGGRAPH '97 Course Notes Multiresolution Surface Modeling , pp. 8-38

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,** pp. 779-785.

Malaka, R. & Zipf, A. (2000), DEEP MAP - Challenging IT research in the framework of a tourist information system. Proceedings of ENTER 2000, pp. 15-27

Mayer, H. (2000): Scale-Space Events for the Generalization of 3D-Building Data. IAPRS Vol. 33 Part B4 , pp. 639-646

Sester, M. (2000): Generalization Based on Least Squares Adjustment. IAPRS Vol. 33 Part B4 , pp. 931-938

Ulrich, M., Steger, C., Baumgartner, A., Ebner, H. (2001): Real-Time Object Recognition in Digital Images for Industrial Applications. 5th Conference on Optical 3D Measurement Techniques, pp. 308-318

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