FEATURE BASED MODEL VERIFICATION (FBMV): A NEW CONCEPT FOR HYPOTHESIS VALIDATION IN BUILDING RECONSTRUCTION

Babak Ameri
Institute for Photogrammetry (ifp), Stuttgart University, Germany
Babak.Ameri@ifp.uni-stuttgart.de

KEY WORDS: Automation, 3D Buildings Reconstruction, Model-based processing, Validation, Evaluation.

ABSTRACT

Geometric modeling and three dimensional (3D) description of objects, such as buildings, collected through an imaging system has become a topic of increasing importance as they are essential for a variety of applications such as telecommunication, 3D city models, virtual tourist information system, etc. Inevitably, a fully human-based image interpretation system would be a costly and labour intensive operation. Therefore, there is an increasing demand towards fully-, or alternatively (for the time being) semi-automatic machine-based image interpretation systems. This paper introduces a novel method called Feature Based Model Verification (FBMV), for modification and refinement of the reconstructed generic polyhedral-like building objects. The development of FBMV is part of an ongoing research project that aims to develop an automated method for recognition and 3D reconstruction of generic building objects using aerial images. A boundary representation (b-rep) of a coarse building hypothesis is generated in a bottom-up process from simple qualitative geometric primitives in image domain to more complex quantitative primitives in object domain. Subsequently, the reconstructed coarse building undergoes a refinement process based on FBMV concept. The verification process is performed by simultaneously fitting the reconstructed model primitives into the homologous two dimensional (2D) features in images taken from different viewpoints while at the same time the geometrical and topological model information is imposed into the process as external and/or internal constraints.

1 INTRODUCTION

Modeling and 3D description of real world objects collected through an imaging system (passive and/or active sensors) has become a topic of increasing importance as they are essential for a variety of applications. Namely telecommunication for planning of wireless networks in cities (Siebe and Büning, 1997, Leberl et al., 1999), urban environmental planning and design to support the decision making processes for development projects (Danahy, 1999, Lange, 1999), virtual tourist information systems to support the on-line positioning, access and queries on the information of the site of interest (Volz and Klinec, 1999), defense and military organization to support the training operation in virtual environment, architectural design for the realistic visualization of the drawing, to mention only a few.

This paper introduces an automated method called *Feature Based Model Verification (FBMV)*, for modification and verification of the reconstructed generic polyhedral-like building model by back projecting the 3D model into the corresponding images taken from different viewpoints. Treating the hypothesis model as evidence leads to a set of confidence intervals in image space that can be used as a search space to find the corresponding 2D image primitives and performing a consistency verification of the reconstructed coarse model. Theoretically, in stereo image analysis systems it is possible to solve all the unknown parameters of a 3D model from matches to corresponding 2D image primitives. However, in practice, the reliability and accuracy of the parameter determination can be substantially improved by simultaneously fitting the model into the images taken from more than two viewpoints. The methods presented here can be used in either situation.

In fact, the proposed method is a kind of model-based image analysis procedure where a priori knowledge of the shape, and geometric appearance of the object is used during the process of interpretation. In this study, the provided object model is generated in a bottom-up, data-driven approach in contrast to the general approach that a specific user-defined model is introduced into the system. The need and a brief discussion of the general framework of FBMV method is given first. The subsequent sections discuss the fundamental concept of FBMV, its formulation and its robustness. The evaluation of the proposed method, its performance and statistical analysis of the result obtained by some experimental test conclude this paper.

2 MOTIVATION

In (Ameri and Fritsch, 1999) the authors presented a new method for reliable generation of a coarse polyhedral-like building model based on 3D intersection of adjacent polygonal primitives of the roof struture (3D plane-roof polygons)

and stitching them along their line of intersection. The positional accuracy of the reconstructed roof elements such as ridgelines of the roof structure is highly related to the quality of the extracted 3D plane-roof polygons. Failure in correctly estimating the orientation of the 3D plane-roof polygons in object space causes displacement and rotation of the ridgelines with respect to their exact positions during the reconstruction process. In addition, due to the nature of region growing type segmentation algorithm used in the lower level process of this work (Fritsch and Ameri, 1998), the quality of the roof outline is poor. In fact, in real-world images, object boundaries cannot be detected solely on the basis of their photometry because of the presence of noise, occlusion and various photometric anomalies. Therefore, methods for finding boundaries based on purely local statistical criteria are tied to error, finding either too many or too few edges based on arbitrary thresholds (Fua and Leclerc, 1990). To supplement the weak and noisy local information of the images and probable misinterpretation of the orientation of the 3D plane-roof polygon, the geometric and topological information that the coarse object model can provide is incorporated into the chain of the reconstruction process. This information is introduced into the process of the object model verification based on a weighted least squares minimization process. A fine building model is obtained in an iterative, top-down, model-driven estimation process by simultaneously fitting the 3D model into the corresponding images where the geometrical and topological model information are integrated into the process as external and/or internal constraints during the estimation. The ability to apply such constraints is essential for the accurate modeling of complex objects. In particular, when dealing with a generic object model, it is crucial that the model elements are both accurate and consistent with each other. For example, individual components of a building can be modeled independently, but to ensure realism, one must guarantee that they touch each other in an architectural way. The estimation procedure yields a description of the building that simultaneously satisfies all the constraints within all the images. As a result, it allows us to perform a consistency check, and refinement of the model across all the images. Moreover the ability of the estimation method to fuse the information and impose the geometrical and topological constraints over all the images increases the accuracy and reliability of the reconstruction.

In the same line of the major image matching techniques, i.e. feature-based (Förstner, 1986), and area-based least squares matching (Förstner, 1982, Ackermann, 1984), the proposed verification process is called Feature Based Model Verification (FBMV). Similar to feature-based image matching techniques where a set of image-driven geometric features such as points, or edges are utilized in one image to be matched to the homologous features in corresponding images in order to, e.g. describe the surface geometry of the viewed scene. The FBMV uses model-driven geometric primitives to be matched to the respective homologous features in corresponding images taken from different viewpoints in order to verify the geometric description of the object model. In recent years, there has been a considerable increase in the number of publications on parameters solving for model-based vision (Lowe, 1991, Haala, 1995, Fua, 1996, Gülch et al., 1998, Brenner and Haala, 1998). An interesting similar work is reported by (Gruen and Li, 1997). Their method is a semi-automatic approach for 3D extraction of linear features. In fact, this is an extension of a point-wise least squares template matching method (Gruen and Stallmann, 1991, Baltsavias, 1991), where a deformable contour model is used as a template instead of a square or rectangle which is generally used in conventional least squares matching techniques. In our study, their work is categorized as an area-based object extraction or alternatively *Area Based Model Verification* (ABMV).

3 FBMV-MATHEMATICAL FOUNDATION

The objective of this section is to formulate the verification of the hypothesis building model. This is carried out by back projecting the 3D coarse model into the corresponding 2D images. Although this transformation is a non-linear operation it is a smooth and well-behaved transformation, and it is a promising candidate for the application of the well known Gauss-Markov estimation model based upon an iterative least squares minimization error criterion. This method requires the appropriate initial guess for the unknown parameters. These values are provided by the geometric and topological information derived from the reconstructed coarse model itself (Ameri and Fritsch, 1999). In practice, the whole spectrum of the observations derived from the building model's description are divided into three major categories as *1) image based, 2) object based* and *3) image-object based observation equations* which are discussed in the next subsections. At the starting point of the estimation process a dense internal data structure is built from the model description. The structure is used to define identical 3D points, edges, and planar surfaces, as well as their topological relationships. In this manner, the model primitives may move independently while being attached to their adjacent primitives. In this way, an edge element connecting two model points can stretch under the influence of shifting one of its endpoint from its initial location and rotate under the influence of the movement of the another endpoint.

3.1 Image Based Observations

The observations concerned in this class are introduced into the estimation process for solving the unknown parameters of 2D primitives such as the parameters of the 2D image edges or the coordinates of the homologous 2D model points in image space. Two types of observations 1) *linearity* which serves as functional model of the estimation process, and 2) *connectivity* which is applied as topological constraint are integrated into the system as image based observations and are discussed next.

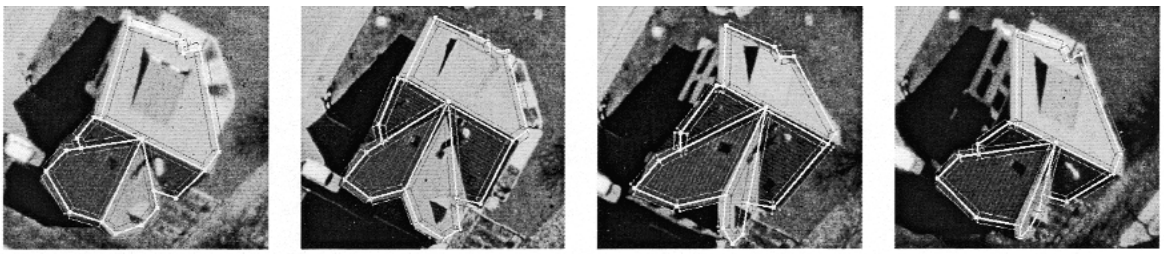

Figure 1: Uncertainty buffer of homologous model edges in corresponding images

3.1.1 Linearity: A Local Internal Geometric Constraint. The estimation process is an orthogonal linear least squares regression problem. It acts as a functional model with the objective to simultaneously minimize the perpendicular sum of the Euclidean distances between the candidate edge-pixels and the projected 2D model edges in all the images. A set of constraints is also integrated into the estimation model to support the minimization process. In fact, it would be sufficient to simply solve a resection equation for modifying the coarse building model if we are able to find the corresponding matches between the model points and their homologous points in the respective images. Thus, in order to overcome the problem of feature corespondence, the match is actually established between the projected model edge and candidate edge-pixels in the image. In other words, since the precise position of the endpoints of image edge is unknown or is missing e.g., due to the occlusion, it is necessary to minimize only the perpendicular distances from the representative points of an image edge to the projected model edge.

In order to measure the perpendicular distance between a representative edge-pixel (x_i, y_i) and the projected 2D model edge e_j, it is useful to express the projected edge in image space in the following form:

$$e_j : \qquad x_i \sin\theta_j - y_i \cos\theta_j - d_j = 0 \tag{1}$$

where d_j is the distance between the origin and the line, and θ_j expresses the angle between the edge and x axis. In practice, the initial parameters of the 2D edge (θ_j^0, d_j^0) are obtained by back projecting the two endpoints of the 3D model edge into the corresponding images using the collinearity equation (8). The computed 2D points are then plugged into the equations (2), (3), in order to derive the initial 2D edge parameters.

$$\theta_j = \arctan \frac{y_{end} - y_{start}}{x_{end} - x_{start}} \tag{2}$$

$$d_j = x_{start} \sin\theta_j - y_{start} \cos\theta_j \tag{3}$$

The projected 2D edge $e_{(r,j)}$ in image I_r serves as initial guess for finding the exact position of the edge model within the corresponding images. An uncertainty buffer with a user specified width is generated around each edge model in image space based on the initial position of the edge and is used as the search space to find the representative edge-pixels. Figure 1 shows the generated buffer around the homologous model edges of a reconstructed coarse building model in four images taken from different view points.

Each pixel within the specified buffer is selected as a representative edge-pixel if it satisfies the following two conditions:

- its gradient direction is approximately perpendicular to the edge model direction, and
- the magnitude of its gray value gradient is more than a data-driven adaptive threshold. This threshold is computed based on a cumulative histogram of the gradient magnitude of all the candidate pixels within the buffer, which satisfy the first criteria. The gradient magnitude associated with each selected pixel is used as its weight in the estimation model.

Applying these two conditions for the selection of representative edge-pixels which should satisfy equation (1) has the following advantages. Firstly, pixels which are laid on the edge image have stronger effect during the fitting procedure because they have more weight in the estimation process. Secondly, the gradients caused by background objects will not interfere with the parameter estimations, as they are not in the approximate direction of the edge model. Figure 2 indicates the selected edge-pixels of the homologous 2D model edges in corresponding images within the generated uncertainty buffer in the first iteration of the estimation process.

At this point, after selecting the edge-pixels representative, we are ready to introduce the linearity constraints into the estimation model. Let us represent the equation of the projected 2D edge in image I_r, passing through the selected edge-pixel (x_i^{img}, y_i^{img}) in the following form:

$$f_{(r,j)}(\theta, d) = x_i^{img} \sin\theta_{(r,j)} - y_i^{img} \cos\theta_{(r,j)} - d_{(r,j)} = e_i(x_i^{img}, y_i^{img}) \tag{4}$$

Figure 2: Selected edge-pixels during the first iteration of the estimation process

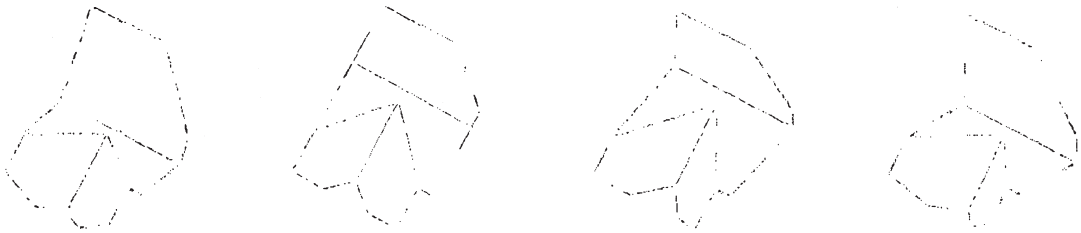

Figure 3: Selected edge-pixels during the last iteration of the estimation process

In this formulation, the orthogonal distance e_i represents an added error parameter, which acts as a cost function and should be minimized during the estimation (see figure 4).

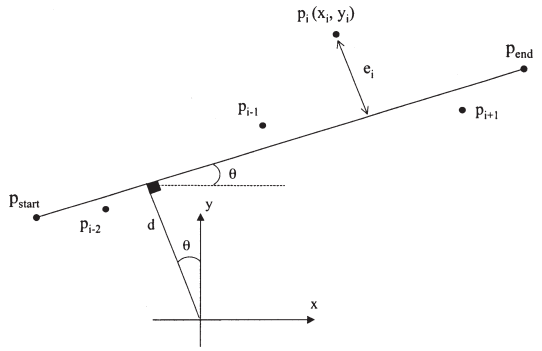

Figure 4: Regression of a 2D image edge to the representative edge-pixels

Linearization of the equation (4) with respect to its parameters $(d_{(r,j)}, \theta_{(r,j)})$ results in the following formulation:

$$\frac{\partial f_{(r,j)}}{\partial \theta_{|\theta=\theta^0_{(r,j)}}} \Delta\theta_{(r,j)} + \frac{\partial f_{(r,j)}}{\partial d_{|d=d^0_{(r,j)}}} \Delta d_{(r,j)} - l_i = e_i(x_i^{img}, y_i^{img}) \tag{5}$$

where

$$l_i \quad = \quad d^0_{(r,j)} - x_i^{img(0)} \sin\theta^0_{(r,j)} + y_i^{img(0)} \cos\theta^0_{(r,j)}.$$

For every selected edge-pixel (x_i^{img}, y_i^{img}) of each 2D edge model $e_{(j,r)}$, within every image I_r, an equation of type (5) is inserted into the system of equations. The total system of equations can be written in matrix form as:

$$A_{linear} \cdot x - l_{linear} = e \quad ; \quad \sigma_0^2 P^{-1}_{linear}. \tag{6}$$

The l_{linear}, is the observation vector containing the orthogonal distance between the candidate pixels and their respective 2D model edge in image space. x is the vector of unknowns consisting of the correction of the edge parameters $(\Delta\theta_{r,j}, \Delta d_{r,j})$, A_{linear} is the associated design matrix including derivatives of the observation equations with respect to the unknowns. The matrix P_{linear}, is the corresponding weight matrix which is introduced as a diagonal matrix and is determined based on the normalized gradient magnitude of each candidate pixel, and e is a error vector with the following statistical assumptions:

$$E(e) = 0 \qquad ; \qquad E(e^T e) = \sigma_0^2 P^{-1}_{linear}.$$

The system of (6) is the well known Gauss-Markov estimation model. The least squares estimation in this model gives a unique and most probable set of estimates for all the parameters of the 2D model edges.

To make the selection process robust and impose a self-diagnosis mechanism, the generated buffer is updated in a regular interval during the iteration process. As the process is iterated, the initial parameters of the 2D edges are updated based on the minimization of the orthogonal distance error between the selected edge-pixels and their respective 2D edges. Consequently the updated parameters define a new orientation for the generated buffer. In addition by introducing a smaller width, the size of the buffer is reduced. As a consequence, as outliers are excluded from the estimation process, this procedure reduces the computational burden and increases the accuracy and speeds up the convergence of the process. Figure 3 indicates the selected edge-pixels of the corresponding 2D edges in the figure 2 for the last iteration of the estimation process.

3.1.2 Connectivity: A Global Internal Topological Constraint. The connectivity constraints are integrated into the estimation model as a topological constraint based on the intersection point between adjacent model edges. Figure 5 depicts a corner of the building model when two edges $edge_1$, and $edge_2$, are connected through the intersection point p_{int}.

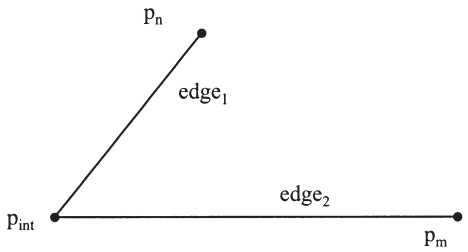

Figure 5: Intersection of two adjacent edges

In general, it is possible to introduce the connectivity constraint into the estimation process both in object space or image space, because it is invariant under the transformation and it is independent of the embedded space. In this model, it is categorized as an image-based observation, in order to overcome the problem of correspondence between the model points within the images. In other words, by setting up the following formulation, the location of the corresponding 2D model points in different images is introduced implicitly based on the topological information, not the geometrical one. That means we do not compute the location of the intersection point explicitly based on the intersection of the two adjacent edges. Therefore, the problem of finding homologous points in respective images is not encountered as it is required in the feature-based matching techniques. In fact, if we had the correspondence relationships between the homologous model points in different images, then the verification of the coarse model would be done simply by obtaining the exact location of the 3D model points based on the simple resection technique such as multiphoto geometrically constrained matching (MPGC) procedure (Baltsavias, 1991, Gruen and Stallmann, 1991).

Similar to the derivation of linearity observation equations (see equation 5), the connectivity constraint between adjacent edges for every associated edge member of a model point is introduced into the total system of equations. Let us consider the equation (4), to represent a 2D edge $e_{(r,j)}$, in image I_r. Linearization of this equation with respect to its parameters, in this case, 2D edge parameters $(d_{(r,j)}, \theta_{(r,j)})$ and 2D coordinates $(x_{int}^{img}, y_{int}^{img})$ of the intersection point in image space results in the following formulation:

$$\begin{aligned} &\frac{\partial f_{(r,j)}}{\partial \theta_{|\theta=\theta^0_{(r,j)}}} \Delta\theta_{(r,j)} \quad + \quad \frac{\partial f_{(r,j)}}{\partial d_{|d=d^0_{(r,j)}}} \Delta d_{(r,j)} + \\ &\frac{\partial f_{(r,j)}}{\partial x_{|x=x_{int}^{img(0)}}} \Delta x_{int}^{img} \quad + \quad \frac{\partial f_{(r,j)}}{\partial y_{|y=y_{int}^{img(0)}}} \Delta y_{int}^{img} - l_{int} = e_i(x_{int}^{img}, y_{int}^{img}) \end{aligned} \tag{7}$$

where

$$l_{int} \quad = \quad d^0_{r,j} - x_{int}^{img(0)} \sin\theta^0_{(r,j)} + y_{int}^{img(0)} \cos\theta^0_{(r,j)}.$$

The arrangment of the above equations for all the connected edges in a matrix form result in the Gauss-Markov model similar to equation (6), except that x is the vector of unknowns consisting of the corrections of the 2D edge parameters $(d_{(r,j)}, \theta_{(r,j)})$ and corrections of the coordinates of the intersection point $(\triangle x_{int}^{img}, \triangle y_{int}^{img})$.

3.2 Image-Object Based Observations

These types of observations are the essential parts of the estimation model. They are integrated into the estimation process in order to establish the required link between the image and object space. They are acting as a bridge to tie the estimated corrections of the unknown parameters obtained in image space to their respective model parameters in object space during each iteration.

3.2.1 Collinearity: A Global External Geometric Constraint. The mapping relation between a point in 3D object space $P_i(X,Y,Z)$, and its perspective projection in 2D image space $p_i^{img}(x,y)$, can be represented by the classical collinearity equations as follows:

$$\begin{aligned} x_i^{cam} &+ F_i^x(X,Y,Z) = 0 \\ y_i^{cam} &+ F_i^y(X,Y,Z) = 0 \end{aligned} \tag{8}$$

Additionally, a mapping relation is needed to relate a point in camera system $p_i^{cam}(x_i,y_i)$ to its corresponding point in image coordinates system $p_i^{img}(x_i,y_i)$. The transformation parameters are expressed in the terms of an affine transformation:

$$\begin{aligned} x_i^{cam} &= c_{11}\cdot x_i^{img} + c_{12}\cdot y_i^{img} + c_{10} \\ y_i^{cam} &= c_{21}\cdot x_i^{img} + c_{22}\cdot y_i^{img} + c_{20} \end{aligned} \tag{9}$$

where $c_{11}, c_{12}, c_{21}, c_{22}$ and c_{10}, c_{20} are the rotation and translation parameters of the affine transformation respectively. Assuming the interior and exterior orientation parameters of each image are given, then the unknowns to be determined are the image coordinates (x_i^{img}, y_i^{img}) of the model point and its corresponding coordinates (X_i, Y_i, Z_i) in the 3D object space. Therefore, if the coordinates of a point is given in object space, the corresponding image coordinates of the point is simply derived from the equations (8) and (9), or alternatively, if an object is imaged from more than one image and the interior and exterior orientation parameters of the images are given, then 3D coordinates of the point can be reconstructed by simultaneous intersection of the above collinearity conditions (resection in space).

Hence, the linearized equations concerning a 2D point in image coordinates $p_i^{img}(x,y)$ with respect to its 3D position $P_i(X,Y,Z)$ can be formulated as:

$$\begin{aligned} \frac{\partial F_i^x}{\partial X_{|X=X_i^0}}\Delta X_i + \frac{\partial F_i^x}{\partial Y_{|Y=Y_i^0}}\Delta Y_i + \frac{\partial F_i^x}{\partial Z_{|Z=Z_i^0}}\Delta Z_i &+ \\ c_{11}\Delta x_i^{img} + c_{12}\Delta y_i^{img} - l_i^x &= e_i^x \\ \frac{\partial F_i^y}{\partial X_{|X=X_i^0}}\Delta X_i + \frac{\partial F_i^y}{\partial Y_{|Y=Y_i^0}}\Delta Y_i + \frac{\partial F_i^y}{\partial Z_{|Z=Z_i^0}}\Delta Z_i &+ \\ c_{21}\Delta x_i^{img} + c_{22}\Delta y_i^{img} - l_i^y &= e_i^y \end{aligned} \tag{10}$$

where

$$\begin{aligned} l_i^x &= -(x_i^{cam(0)} + F_i^x(X^0,Y^0,Z^0)) \\ l_i^y &= -(y_i^{cam(0)} + F_i^y(X^0,Y^0,Z^0)). \end{aligned}$$

The equation (11) for all the model points can also be arranged in the standard Gauss-Markov model (see equation 6), whereas x is the vector of unknowns consisting the correction of the model points in object $(\Delta X_i, \Delta Y_i, \Delta Z_i)$, and image space $(\Delta x_i^{img}, \Delta y_i^{img})$.

3.3 Object Based Observations

As it has been discussed so far, the main objective of the FBMV method is to integrate the model-driven information into the estimation model as supporting constraints. The set of observation equations which is classified in this category are directly obtained based upon the discription of the reconstructed coarse building model, before or during the estimation process. These constraints are introduced between the model primitives in object space as global or local geometric constraints and are linearized and applied as weighted observation equations. In this manner, the integration of the model primitives as unknowns into the total system of equations is completely flexible. Introducing the relationship between the model primitives as a strict condition by increasing its weight or alternatively reducing its influence into the system by decreasing its weight.

The system of (6) is the well known Gauss-Markov estimation model. The least squares estimation in this model gives a unique and most probable set of estimates for all the parameters of the 2D model edges.

To make the selection process robust and impose a self-diagnosis mechanism, the generated buffer is updated in a regular interval during the iteration process. As the process is iterated, the initial parameters of the 2D edges are updated based on the minimization of the orthogonal distance error between the selected edge-pixels and their respective 2D edges. Consequently the updated parameters define a new orientation for the generated buffer. In addition by introducing a smaller width, the size of the buffer is reduced. As a consequence, as outliers are excluded from the estimation process, this procedure reduces the computational burden and increases the accuracy and speeds up the convergence of the process. Figure 3 indicates the selected edge-pixels of the corresponding 2D edges in the figure 2 for the last iteration of the estimation process.

3.1.2 Connectivity: A Global Internal Topological Constraint. The connectivity constraints are integrated into the estimation model as a topological constraint based on the intersection point between adjacent model edges. Figure 5 depicts a corner of the building model when two edges $edge_1$, and $edge_2$, are connected through the intersection point p_{int}.

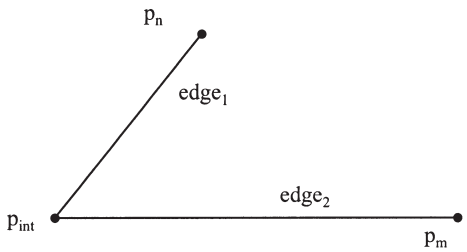

Figure 5: Intersection of two adjacent edges

In general, it is possible to introduce the connectivity constraint into the estimation process both in object space or image space, because it is invariant under the transformation and it is independent of the embedded space. In this model, it is categorized as an image-based observation, in order to overcome the problem of correspondence between the model points within the images. In other words, by setting up the following formulation, the location of the corresponding 2D model points in different images is introduced implicitly based on the topological information, not the geometrical one. That means we do not compute the location of the intersection point explicitly based on the intersection of the two adjacent edges. Therefore, the problem of finding homologous points in respective images is not encountered as it is required in the feature-based matching techniques. In fact, if we had the correspondence relationships between the homologous model points in different images, then the verification of the coarse model would be done simply by obtaining the exact location of the 3D model points based on the simple resection technique such as multiphoto geometrically constrained matching (MPGC) procedure (Baltsavias, 1991, Gruen and Stallmann, 1991).

Similar to the derivation of linearity observation equations (see equation 5), the connectivity constraint between adjacent edges for every associated edge member of a model point is introduced into the total system of equations. Let us consider the equation (4), to represent a 2D edge $e_{(r,j)}$, in image I_r. Linearization of this equation with respect to its parameters, in this case, 2D edge parameters $(d_{(r,j)}, \theta_{(r,j)})$ and 2D coordinates $(x_{int}^{img}, y_{int}^{img})$ of the intersection point in image space results in the following formulation:

$$\begin{aligned} &\frac{\partial f_{(r,j)}}{\partial \theta_{|\theta=\theta^0_{(r,j)}}} \Delta\theta_{(r,j)} \quad + \quad \frac{\partial f_{(r,j)}}{\partial d_{|d=d^0_{(r,j)}}} \Delta d_{(r,j)} + \\ &\frac{\partial f_{(r,j)}}{\partial x_{|x=x_{int}^{img(0)}}} \Delta x_{int}^{img} \quad + \quad \frac{\partial f_{(r,j)}}{\partial y_{|y=y_{int}^{img(0)}}} \Delta y_{int}^{img} - l_{int} = e_i(x_{int}^{img}, y_{int}^{img}) \end{aligned} \tag{7}$$

where

$$l_{int} \quad = \quad d^0_{r,j} - x_{int}^{img(0)} \sin\theta^0_{(r,j)} + y_{int}^{img(0)} \cos\theta^0_{(r,j)}.$$

The arrangment of the above equations for all the connected edges in a matrix form result in the Gauss-Markov model similar to equation (6), except that x is the vector of unknowns consisting of the corrections of the 2D edge parameters $(d_{(r,j)}, \theta_{(r,j)})$ and corrections of the coordinates of the intersection point $(\triangle x_{int}^{img}, \triangle y_{int}^{img})$.

3.3.1 Coplanarity: A Global External Geometric Constraint. Due to the pragmatic assumption that building roof structures are geometrically described by the aggregation of $k \geq 1$ planar surface(s), all the bounding points $P_i(X, Y, Z)$, of the 3D plane-roof polygon F_k, should satisfy the coplanarity condition defined as follows:

$$f_{(i,k)}(\vec{n}, D) = A_k X_i + B_k Y_i + C_k Z_i + D_k = e_i \tag{11}$$

where (A_k, B_k, C_k), are the components of the surface normal $\vec{n}_k$, D_k is the distance from origin to the plane-roof polygon F_k, and e_i is an added error parameter. The partial derivatives of this equation with respect to the unknown parameters, that is the 3D coordinates of the model points, is obtained by:

$$\frac{\partial f_{(i,k)}}{\partial X_{|X=X_i^0}} \Delta X_i + \frac{\partial f_{(i,k)}}{\partial Y_{|Y=Y_i^0}} \Delta Y_i + \frac{\partial f_{(i,k)}}{\partial Z_{|Z=Z_i^0}} \Delta Z_i - l_i = e_i \tag{12}$$

where

$$l_i = -(A_k^0 X_i^0 + B_k^0 Y_i^0 + C_k^0 Z_i^0 + D_k^0).$$

In fact, the corrections $(\Delta X_i, \Delta Y_i, \Delta Z_i)$, in each iteration represent changes to the initial location of the model points, while the best planar fit to the updated model points is obtained. Introducing an equation of the type (12) for every point of the plane-roof polygons F_k, in the estimation model and arranging all the equations in the matrix form result in a Gauss-Markov model similar to equation (6), whereas x is the vector of unknowns consisting the corrections of the coordinates of the model points $(\Delta X_i, \Delta Y_i, \Delta Z_i)$.

3.3.2 Conditional Constraints. The FBMV is an iterative procedure based on Newton-Raphson method, thus it converges to the minimum and becomes closer to the correct solution in every iteration, unless the system is degenerated, the initial values are so far away from the true solution, or the estimation model is incorrectly established. This property enables us to integrate additional constraints between the model primitives during the iteration process, if certain conditions are satisfied. As we have mentioned previously the strength of our method is a data-driven generic data model. That means instead of imposing certain regularities or conditions into the model in the earlier stages of the reconstruction process, these regularities and constraints are introduced into the model in the higher level process of reconstruction. Such constraints are the *orthogonality*, or *parallelity* between the adjacent model edges, *symmetricalness* or *semi-symmetricalness* between the adjacent faces, and so on. The decision to impose these constraints into the estimation model is made during model verification process when the required criteria are met. The triggered constraints are integrated into the model, simply by adding a new row to the total system of equations. For the sake of completeness the orthogonality constraints are elaborated in details next, the other constraints can be dealt with in the same manner.

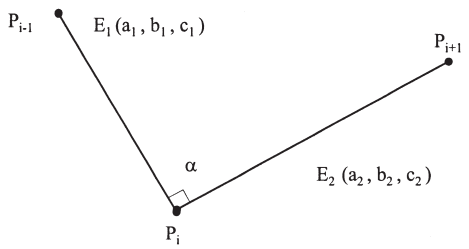

Figure 6: Two orthogonal adjacent edges

Orthogonality: A Local External Geometric Constraint. Figure 6 represent the angle α, between two 3D model edges E_1, and E_2. The conditional geometric constraint of the orthogonality is applied for every model point P_i, when α satisfies the following condition during each iteration:

$$90 - t \leq \alpha \leq 90 + t \tag{13}$$

where t is a threshold (e.g., $5°$) indicating the small deviation of α from its expected value i.e. $90°$. In fact, when two adjacent edges are considered orthogonal then the following constraints should be met:

$$f_i(X, Y, Z) = a_1 a_2 + b_1 b_2 + c_1 c_2 = 0 = e_i \tag{14}$$

where $\vec{E}_1(a_1, b_1, c_1)$, and $\vec{E}_2(a_2, b_2, c_2)$, are the directions of the E_1 and E_2 respectively and e_i is an added noise parameter. Linearization of the equation (14) with respect to the position of the model points $P_i(X, Y, Z)$ result in the form:

$$\frac{\partial f_i}{\partial X_{|X=X_i^0}} \Delta X_i + \frac{\partial f_i}{\partial Y_{|Y=Y_i^0}} \Delta Y_i + \frac{\partial f_i}{\partial Z_{|Z=Z_i^0}} \Delta Z_i - l_i = e_i \tag{15}$$

where

$$l_i = -f_i^0(X, Y, Z) = -(a_1^0 a_2^0 + b_1^0 b_2^0 + c_1^0 c_2^0).$$

Therefore, introducing an equation of the type (15) as a weighted observation equation for every model point which satisfies the orthogonality conditions, and formulating them in a matrix form will result in the well known formulation of a Gauss-Markov model such as equation (6), except that x is the vector of unknowns consisting of the corrections to the coordinates of the model points $(\Delta X_i, \Delta Y_i, \Delta Z_i)$.The other conditional constraints can be implemented in the same manner as discussed for the orthogonality constraint and integrated into the estimation model.

Based on the assumption that all the parameters involved in the estimation model are considered observations, and consequently equations arising with the constraints or conditions are introduced into the total system of equations as weighted observation equations, then we are able to join the different estimation models which are formed by the (5), (7), (11), (12), (15), or any other equations into a unified combined least squares adjustment process where its solution gives the correction to the initial parameters of the model.

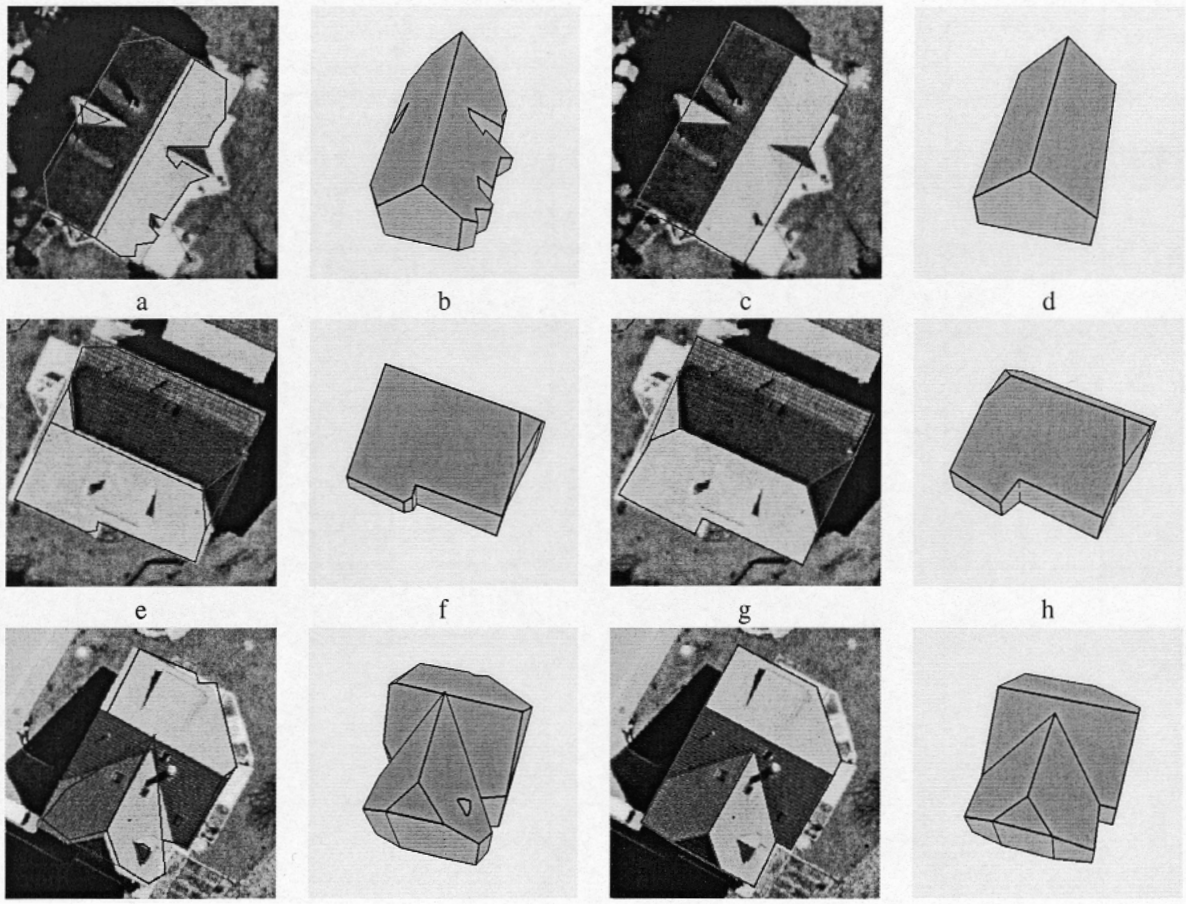

Figure 7: Reconstructed fine building models: a, e, f) reconstructed coarse building model overlaid on aerial image, - b, f, m) perspective view of the reconstructed 3D coarse building model, - c, g, n) reconstructed fine building model overlaid on aerial image, - d, h, q) perspective view of the 3D fine building model.

4 EXPERIMENTS AND RESULT

Three representative buildings of the Avenches data set (Mason et al., 1994), are selected to represent the performance of the proposed method and to visualize the outcome of the FBMV algorithm. It should be noted that the whole process of model verification is only applied on the roof structure and the vertical walls are added to the building model at the final stage. This is performed based on the analysis of the bounding edges and points of the verified fine roof structures. The first building shown in first row of the figure 7 is a simple gable roof structure. The reconstructed coarse model (figure 7-b) indicates that the reconstruction process (Ameri and Fritsch, 1999), recovered the fundamental structure of the building. However, due to the presence of a dormer window on top of the roof and also the low contrast of the roof outline, the bounding edges of the building are broken into the small pieces and are very rugged. The FBMV approach successfully verified the model and removed the redundant edge segments. Furthermore, imposing the orthogonality constraints during

the verification process enables FBMV to accurately recover the building corners (see figure 7-d), which were trimmed off during the segmentation process. Figures (7-a and 7-c) show the initial and modified building model overlaid on the corresponding aerial image, respectively.

The second example deals with reconstruction of a hipped-gable roof structure (see second row of the figure 7). As it is shown in figures (7-e) and (7-f), although the major structures of the building, the bounding edges and even small protruding structure of the roof is reconstructed correctly. Due to the low contrast of the edge segment of the hip tile of the roof, the reconstructed edge is completely shifted away from its correct position. In addition, the position of the corner points is not precisely defined. The modification of the model based on a multi-photo estimation process and imposing the global constraints as discussed previously, enables FBMV to correctly recover this part of the roof and forces the displaced model edge to located in its true position. Furthermore it also defines the positions of the model points more accurately (see figures 7-g and 7-h).

The last example shown in third row of the figure 7 represents a more complex roof structure. The generated hypothesis coarse building model describes the building completely, but there are still displacements in some of the model primitives, specially the intersection point between three adjacent plane-roof polygons is shifted significantly from its real position (see figure 7-l). This is due to the failure in defining the correct orientation (slope) of the respective 3D plane-roof polygons, which is caused by the low quality of the utilized DSM (Ameri and Fritsch, 1999). By applying the FBMV process, the normal vectors $\vec{n}_k$ of every plane-roof polygons are recovered precisely and consequently the intersection point is moved to its real position. In addition, the real bounding edges of the model are also precisely located and the redundant one is eliminated from the final model (figure 7-q). To complete the performance evaluation of the FBMV, the following section is dedicated to numerical analysis and assessment of the quality of the final reconstructed building obtained from the estimation model.

5 QUALITY ASSESSMENT

One of the key issue of the FBMV method is its ability to provide the essential tools for evaluation of the quality of the reconstructed model and its geometric primitives. The combined least squares solution provides an estimate for the variance factor $\hat{\sigma}_0^2$ which can be used for the performance evaluation of the estimation process. In other words, it is used to judge whether or not the estimation model is consistent with the earlier assumption that the noise distribution follows a normal distribution function with a given standard deviation, which was the motivation to apply a least squares minimization of the error criterion. In addition, considering sufficient agreement between the estimation model and our early assumption, the standard and statistically well known covariance matrix Cov_{xx} of the estimated parameters can be obtained as follows:

$$Cov_{xx} = \hat{\sigma}_0^2 (\sum A_c^T P_c A_c)^{-1} \tag{16}$$

The estimated variances of the unknown parameters, specifically in our case the coordinates of the model points in 3D space $(\sigma_X^2, \sigma_Y^2, \sigma_Z^2)$, are the qualitative measures which indicate the accuracy of the model primitives and act as the decision criteria in order to reject or accept the estimated model elements based on the simple thresholding process. The evaluation process can be integrated into the whole chain of reconstruction process as an edition process (traffic light concept (Förstner, 1996)). In a simple manner, these measures give a hint to the end user to perform a visual check on the end product and perform the required modifications on the signalized model primitives if necessary. In the following are the numerical results and the statistical analysis of the complex building (figure 7), where its corner points are numbered as illustrated in figure (8).

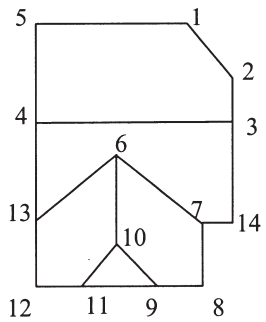

Figure 8: Top view of a single complex building

The test was carried out for the verification of the model based on utilizing two, and four corresponding images taken from different views. In addition, the reconstructed building in every test is compared with a reference model digitized

manually by an operator, in order to show a realistic quality measure of the process as well. The results are tabulated in tables (1), and (2) respectively.

Point-ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	RMSE
ΔX	0.26	0.20	0.09	0.17	0.20	0.10	0.17	0.41	0.16	0.16	0.23	0.12	0.38	0.18	0.22
ΔY	-0.02	-0.11	-0.04	-0.16	-0.02	-0.09	-0.28	0.14	-0.15	-0.08	0.01	0.07	-0.21	0.30	0.15
ΔZ	0.03	0.03	0.12	-0.17	0.05	0.10	-0.18	0.22	0.03	0.28	-0.03	-0.26	-0.35	-0.17	0.18
σ_x^2	0.44	0.44	0.65	0.65	0.30	0.21	0.61	0.07	0.58	0.23	0.37	0.37	0.37	0.37	
σ_y^2	0.71	0.70	0.23	0.23	0.35	0.35	0.37	0.54	0.54	0.64	0.42	0.43	0.45	0.11	
σ_z^2	0.69	0.68	0.16	0.16	0.26	0.12	0.33	0.20	0.62	0.15	0.30	0.31	0.24	0.08	

Table 1: Verified coarse building model based on two corresponding aerial images

A comparison of the estimated variances of the model points coordinates with respect to the absolute values of the differences between the estimated coordinates and the reference coordinates deduce consistency and agreement in both tests. Moving from the estimation model based on two images toward the one utilizing four images indicates a tendency in increasing the accuracy of the estimated model points, as it was expected. In addition imposing more images into the estimation process increase the reliability of the model.

Point-ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	RMSE
ΔX	0.19	0.20	0.19	0.17	0.20	0.17	0.11	0.28	0.03	0.10	0.23	0.12	0.38	0.06	0.19
ΔY	-0.09	-0.16	-0.12	-0.20	-0.10	-0.17	-0.27	-0.17	-0.18	-0.19	-0.02	-0.01	-0.18	-0.29	0.17
ΔZ	-0.15	-0.15	-0.06	-0.23	-0.10	0.08	-0.13	0.09	-0.14	0.04	-0.27	-0.23	-0.19	-0.13	0.16
σ_x^2	0.34	0.34	0.49	0.49	0.20	0.23	0.64	0.08	0.56	0.23	0.22	0.22	0.30	0.26	
σ_y^2	0.30	0.30	0.23	0.23	0.23	0.32	0.36	0.30	0.32	0.46	0.32	0.32	0.39	0.11	
σ_z^2	0.34	0.34	0.21	0.21	0.26	0.16	0.46	0.17	0.64	0.20	0.25	0.25	0.26	0.09	

Table 2: Verified coarse building model based on four corresponding aerial images

The above experimental results show the strength and generality of the proposed FBMV in recovering the reliable and accurately defined geometric primitives of different sorts of building structures, which is an essential part of any automated vision system. It shows that the proposed method is capable of working with any complex polyhedral-like object model, if an appropriate initial hypothesis model is available.

6 CONCLUSION AND FUTURE WORK

The main objective of this paper was to introduce the concept of the FBMV and give some hints of how the information derived from the model itself can support the verification process.

The problem considered in this study was to determine the precise geometric description of a polyhedral-like building model given matches between the model primitives and the image features. However, the proposed framework allows different non-polyhedral object models to be used. A consequent disregarding of this restriction is that the projected model edges are not necessarily straight edges and the model faces are not inevitably planar surfaces, thus the geometric routines should be adopted with different geometry. An important aspect of the FBMV method is the ability to solve the model parameters by simultaneously fitting all the geometric primitives of the 3D model into all the homologous 2D image features, taking into account the external and internal geometric and topologic properties of the model structure and imaging process as constraints during parameter estimation. This is important because it allows the earlier initial matches or the partial matches between the 3D model primitives and 2D image features force the location of other structural elements of the model. Thereby new matches that can be used to verify or reject the initial estimated model parameters are generated.

REFERENCES

Ackermann, F., 1984. Digital image correlation: Performance and potential application in photogrammetry. Photogrammetric Record 11(64), pp. 429–439.

Ameri, B. and Fritsch, D., 1999. 3-D reconstruction of polyhedral-like building models. In: IAPRS, Vol 32, Part 3-2W5, Munich, Germany, pp. 15–20.

Baltsavias, E. P., 1991. Multiphoto Geometrically Constrained Matching. PhD thesis, Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland.

Brenner, C. and Haala, N., 1998. Fast production of virtual reality city models. In: Proc. ISPRS Conference Comm. IV, Stuttgart.

Danahy, J., 1999. Visualization data needs in urban environmental planning and design. In: D. Fritsch and R. Spiller (eds), Photogrammetric Week '99, Herbert Wichmann Verlag, Heidelberg, pp. 351–365.

Förstner, W., 1982. On the geometric precision of digital correlation. In: International Archives of Photogrammetry and Remote Sensing, Vol. 24, 3/3, pp. 176–189.

Förstner, W., 1986. A feature based correspondence algorithm for image matching. In: International Archives of Photogrammetry and Remote Sensing, Vol. 26, 3/3, Rovaniemi.

Förstner, W., 1996. 10 pros and cons against performance characterization of vision algorithms. In: Workshop on Performance Characterization of Vision Algorithms, Cambridge.

Fritsch, D. and Ameri, B., 1998. Geometric characteristics of digital surfaces: A key towards 3D building reconstruction. In: Proc. ISPRS Conference Comm. III, Columbus, Ohio, pp. 119–126.

Fua, P., 1996. Model-based optimization: Accurate and consistent modeling. In: Proc. ISPRS Congress Comm. III, Wien, pp. 222–233.

Fua, P. and Leclerc, Y. G., 1990. Model driven edge detection. Machine Vision and Applications (3), pp. 45–56.

Gruen, A. and Li, H., 1997. Linear feature extraction with 3-D lsb-snakes. In: A. Gruen, O. Kuebler and M. Baltsavias (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), Birkhäuser Verlag, Basel, Boston, Berlin, pp. 287–298.

Gruen, A. and Stallmann, D., 1991. High accuracy edge matching with an extension of the mpgc-matching algorithm. In: Industrial Vision Metrology, SPIE Proceedings Series, Vol. 1526, pp. 42–55.

Gülch, E., Müller, H., Läbe, T. and Ragia, L., 1998. 3-D reconstruction of polyhedral-like building models. In: Proc. ISPRS Conference Comm. III, Columbus, Ohio, pp. 331–338.

Haala, N., 1995. 3D building reconstruction using linear edge segments. In: D. Fritsch and D. Hobbie (eds), Photogrammetric Week '95, Herbert Wichmann Verlag, Heidelberg, pp. 19–28.

Lange, E., 1999. The degree of realism of gis-based virtual landscapes: Implications for spatial planing. In: D. Fritsch and R. Spiller (eds), Photogrammetric Week '99, Herbert Wichmann Verlag, Heidelberg, pp. 367–374.

Leberl, F., Walcher, W., Wilson, R. and Gruber, M., 1999. Models of urban areas for line-of-sight analyses. In: IAPRS, Vol 32, Part 3-2W5, Munich, Germany, pp. 217–226.

Lowe, D. G., 1991. Fitting parameterized three-dimensional models to images. IEEE Transactions on Pattern Analysis and Machine Intelligence 13(5), pp. 441–450.

Mason, S., Baltsavias, M. and Stallmann, D., 1994. High precision photogrammetric data set for building reconstruction and terrain modelling. Data description, Institute for Geodesy and Photogrammetry, ETH Zurich.

Siebe, E. and Büning, U., 1997. Application of digital photogrammetric products for cellular radio network planning. In: D. Fritsch and D. Hobbie (eds), Photogrammetric Week '97, Herbert Wichmann Verlag, Heidelberg, pp. 159–164.

Volz, S. and Klinec, D., 1999. Nexus: The development of a platform for location aware application. In: Proceedings of the third Turkish-German joint Geodetic Days, Vol. 2, Istanbul, Turkey, pp. 599–608.