TEXTURE MAPPING OF 3D BUILDING MODELS WITH OBLIQUE DIRECT GEO-REFERENCED AIRBORNE IR IMAGE SEQUENCES

U. Stilla [a], J. Kolecki [a,b], L. Hoegner [a]

[a] Photogrammetry and Remote Sensing, Technische Universitaet Muenchen (TUM)
(stilla, hoegner)@bv.tum.de
[b] AGH University of Science and Technology, Cracow, Faculty of Mining Surveying and Environmental Engineering,
(DAAD scholarship fellow at TUM),
kolecki@agh.edu.pl

KEY WORDS: Infrared images, image sequences, GPS, inertial navigation, 3D city models, edge matching

ABSTRACT:

Infrared (IR)-radiation of buildings can be captured by thermal image data. Nowadays, thermal images of different parts of buildings acquired in various scales are analysed, however conducted analysis concern mainly the radiometric characteristics, avoiding dealing with the geometry of the images. Aim of our work is the automated texture mapping of existing 3D building models with images recorded by IR cameras. In this way, 2D images are connected with a 3D GIS database and can be further analysed together with the information about the geometry of objects. The extraction of the front facades textures of our test area was already obtained using the terrestrial images acquired from the mobile platform. However those image sequences are restricted to the facades at the street sites. Roof structures or backyards cannot be recorded in this way. Oblique images taken from the helicopter are expected to provide a sufficient complement of the missing data. The flight path was recorded by an inertial measurement unit (IMU) and GPS. In order to utilize the INS/GPS data, first the IMU misalignment with respect to the camera axis has to be corrected. Then the texture mapping is performed without any additional refinement of camera position. The evaluation of obtained textures is performed and some additional methods for their improvement are proposed.

1. INTRODUCTION

1.1 Problem Overview

For texturing the entire hull of buildings, images taken from different viewing direction are required. Airborne imagery provides an appropriate texture for roofs, while terrestrial images are useful for detailed mapping of façades. The manual acquisition of terrestrial imagery for larger parts of a 3d city model is often a time consuming process. For automation of this process IR cameras mounted on a mobile platform, e.g. a vehicle, can be used to capture image sequences of façades (Hoegner and Stilla, 2007a, 2007b). However, the usage of this method is not appropriate for roofs and inner yards. In general, textures from roofs and inner yards can be captured by airborne platforms. Because in nadir views from airborne platforms façades of inner yards do not appear with sufficient resolution, oblique views are preferred.

In this paper an exploitation of direct geo-referenced sequences of IR images is investigated. The evaluation is focused on the accuracy of direct mapping using INS and GPS data for the texture mapping. A precondition for precise mapping is a geometrically calibrated camera with known interior orientation, accurate GPS and INS data, and a geo-referenced 3d building model.

The IR image sequence was taken additionally and simultaneously to a laser scanning during a flight campaign (Hebel & Stilla, 2007). The IMU was mainly required for the laser scanner and a boresight calibration was carried out according to it. The interior orientation of the used IR camera was unknown as well as the shift and misalignment of the camera coordinate system related to the IMU coordinate system.

As a solution to these two problems the bundle block adjustment with a self calibration is proposed. Not only the interior orientation and the distortion parameters could be estimated in this way, but also the sensor shift and misalignment could be estimated. Considering the calculated corrections to the IMU observations a texture mapping was carried out.

1.2 Related Work

In the last decade many papers covering the topic of 3D city models have been published. Combining image data with vector data of 3d city models is of interest for different fields like automatic texturing, navigation, augmented reality etc. Stilla et al. (2000) have shown an approach to match oblique IR image data and 3d city models. Frueh et al. (2004) used high resolution aerial images and presents an approach to automated texture mapping of existing 3D city model. The use of a rough knowledge about images exterior orientation in an edge matching process is presented. Furthermore some aspects concerning optimal image selection are covered. The concept of image and model integration within Virtual Globes is also described by Nebiker et al. (2007), however, rather from an informatics point of view. Eugster and Nebiker (2008) projected simple building models into images taken from an unmanned aerial vehicle (UAV) with using measurements from a low cost and low accuracy IMU and GPS. They observed misalignments and suggested to improve the projection by a matching process.

Eugster and Nebiker (2007) and Cramer (2002) address the problems of systematic errors resulting from the IMU misalignment and shift with respect to the camera. The first one, dealing with a low cost IMUs installed on UAVs, proposes the incorporation of three misalignment angles as the additional parameters of bundle adjustment. The measurement of the coordinates of the shift vector before the flight is assumed to be sufficient. Cramer indicates the need of including shift values to the bundle block adjustment too however his work is mainly related to high quality IMUs used in high-accuracy demanding photogrammetric missions.

Hoegner and Stilla (2008) used terrestrial infrared image sequences for texturing of building models by automatic mosaicing. However, because of the small field of view of the IR camera and lack of edges in the most of the images, the relative orientation of subsequent frames has to be estimated before. Because INS data were not available they first generated a point cloud from the images only and a corresponding camera path. Then a matching of the point cloud and building model was performed using coarse GPS information.

In contrast to our previous work we exploit airborne infrared sequences instead of terrestrial infrared sequences to complement building textures. While terrestrial images show only small building parts airborne images show complete buildings which allows different mapping strategies.

2. DATA OVERVIEW

2.1 The Thermal Images

The thermal images were taken with the IR camera AIM 640 QLW FLIR which was mounted on a platform carried by a helicopter. The helicopter was flying approx. 400 meters over ground. The camera was forward looking and its pitch angle was about 45°. Figure 1 shows the geometry of the acquisition.

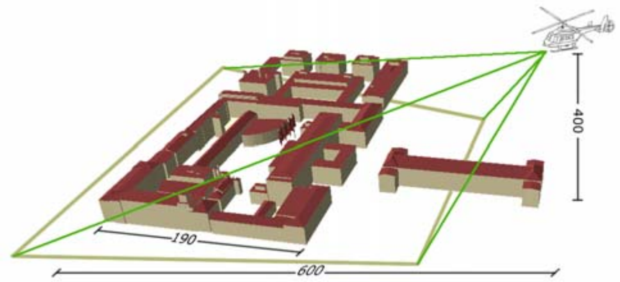


Figure 1. The geometry of the acquisition

For capturing all façades and roof surfaces and to minimize the area of occlusion the same scene was recorded four times (Fig. 2). The helicopter was flying subsequently roughly northward (#1), southward (#2), eastward (#3) and westward (#4). From the entire sequence four short strips showing the test scene were selected. Each consists of 128 frames acquired with a frequency of 25 images per second. The image resolution of the frames is 640 x 512 pixels (see Fig. 3). The interior orientation was given with approximated values.

2.2 GPS and INS Information

The helicopter was equipped with Applanix POS AV 510 which contains an inertial measurement unit (IMU) and a GPS receiver. It should be mentioned that the IMU angles, known as roll, pitch and yaw, are defined in a different way as omega, phi and kappa angels, which are commonly used in photogrammetry. Although the GPS antenna was mounted in the front of the helicopter cockpit, the measured coordinates were referred to the centre of the IMU and integrated with the inertial measurements. As an integration approach, the Kalman filtering was used (Grewal et al, 2007). The inertial data were recorded with a frequency of 200 Hz, while GPS position with frequency of 1 Hz. Unfortunately, there were no DGPS corrections available within the period of flight, so GPS/INS data certainly have a poorer accuracy (Cramer, 2002).

Figure 2. Test area and flight trajectory (Hebel & Stilla, 2007)

Figure 3. An IR image (#13200) from the sequence (#4)

The IMU was mounted close to the infrared camera. However, the exact vector pointing from IMU centre to camera centre is unknown and the IMU coordinate system is not perfectly aligned with the camera coordinate system. So the appropriate corrections to the measured coordinates and measured angles have to be estimated and considered during the further works. Besides angles and coordinates, the Applanix system provides additional parameters like the GPS time, current speed and a total flight distance. During each frame acquisition, the time stamp is written to the image file, which allows the assignment of the measured exterior orientation parameters to an appropriate image.

2.3 3D Building Models

A 3d model in level of detail LOD2 of test area TUM comprising the buildings of Technische Universitaet Muenchen was used. The model was constructed from aerial images using photogrammetric software. Different building shapes were used and fitted to the stereo model using InJECT. The model shows just a generalized view of the scene which results often in small differences between the position of true and modelled building edges.

3. ESTIMATION OF CAMERA ORIENTATION

3.1 Camera Calibration

As accurate values of neither interior orientation parameters, nor distortion were known, a bundle adjustment with self calibration had to be solved. First about 140 control points were measured on an accurate reference stereo model. Control points were uniformly distributed in the test area TUM, as well as in the neighbourhood area. Care had to be taken to measure points that are visible also in the infrared images. The control points were located mostly on the roof corners, as good points located on the ground were difficult to find because of occlusions.

Three frames were selected from each flight stripe. One image was taken from the beginning, one from the middle and one form the end of the stripe. For checking the behaviour of the parameters within a short time period additional subsequent images from a single stripe (4) were included for the bundle adjustment.

The ground control points as well as some tie points were measured on the infrared images and the bundle adjustment was calculated. Besides interior orientation parameters also the distortion confidences were estimated. The estimated lens distortion is significant, and its effect can be easily observed in the images.

3.2 Corrections to the GPS/INS Data

As already mentioned there are few reasons why GPS/INS data cannot be utilized without corrections to perform the direct texturing of the surfaces of the city model. Beside shift and misalignment another worth considering reason of systematic errors in exterior orientations is a possible time shift effect. The time stamp encoded in each image file may not correspond to IMU measurement taken at the moment of an exposure. If the helicopter vibration has a high amplitude and a high frequency, the influence of the time shift becomes meaningful. The influence of the time shift was not considered in this work however it is going to be examined in the further investigations.

Systematic corrections to the measured exterior orientation parameters were determined and the data from stripe 4 were used to verify the received correction parameters. Lines of the 3D model were projected into the images to assess visually the results.

Before applying the corrections there was a significant misalignment of the IMU coordinate system relative to the camera coordinate system. Most significant for the misalignment was the difference in the measured pitch angle. To estimate the misalignment errors, the angular exterior orientation parameters, calculated by the bundle adjustment are utilized. The measured and the calculated angles cannot be compared in a direct way, because inertial measurements are performed with the use of a different rotation system. The Euler rotation matrix is calculated for each photo included from stripe 4 into the bundle adjustment. Afterwards the roll, pitch and yaw angles are calculated from the Euler rotation matrix. Then the comparison between the calculated and measured rotations is performed. Subsequently the differences between the coordinates of the projection centres are also calculated. The results of the comparison are shown in Table 1.

point ID	Δ roll [g]	Δ pitch [g]	Δ yaw [g]	Δ X [m]	Δ Y [m]	Δ Z [m]
13145	1.442	9.786	0.720	0.1	2.8	-3.4
13180	0.774	9.561	0.497	-3.8	-0.7	-1.5
13200	1.523	9.622	0.803	-1.4	3.4	-1.4
13212	1.541	9.789	0.626	0.1	5.8	-2.6
13213	0.478	9.503	0.502	-5.1	0.0	-1.4
13214	0.760	9.535	0.332	-3.7	2.3	-0.7
13215	1.233	9.827	0.527	-1.4	3.6	-2.7
13216	0.844	9.557	0.467	-3.4	2.0	-1.3
13217	1.484	9.650	0.589	-0.1	6.0	-1.6
13218	0.352	9.228	0.496	-6.3	-0.7	0.7
13219	1.298	9.318	0.547	-2.6	4.9	0.3
13220	0.963	9.325	0.640	-4.7	1.9	0.0
13248	1.360	9.217	0.747	-4.4	3.6	0.4
13264	0.791	9.174	0.346	-5.4	4.3	1.5
mean	**1.060**	**9.507**	**0.560**	**-3.0**	**2.8**	**-1.0**
std. dev	0.401	0.222	0.139	2.1	2.2	1.4

Table 1. Differences between measured and calculated exterior orientation elements (sequence #4)

The large differences of the pitch angle can be observed by the mean (9.507) in the Table. The means of differences in roll and pitch angles are smaller but still enough meaningful to affect the results of the texturing. Furthermore, a high standard deviation can be seen for the roll angle. High changes observed between subsequent frames are present in the position differences, too.

The mean values of the calculated differences can be used as corrections to the measured angular exterior orientation parameters of images in stripe #4. The projection of the model lines into an image after applying the corrections is shown in the Fig. 4.

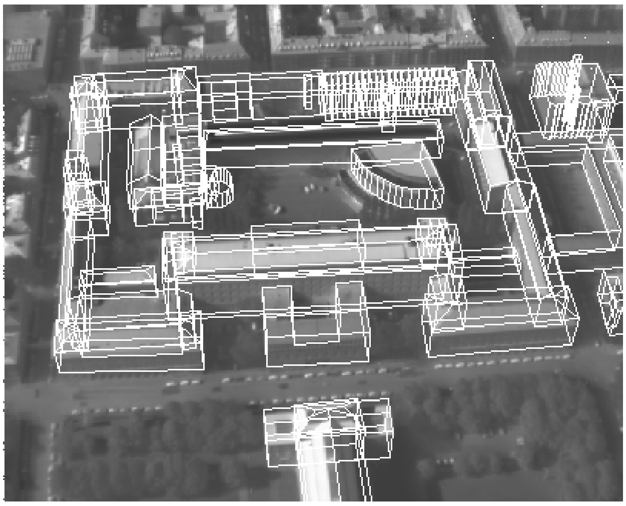


Figure 4. Projections of TUM model lines into IR image #13200 using corrected INS/GPS data.

4. TEXTURE MAPPING

In the texture mapping, 2d texture images are generated for every visible surface of the building model from the input 2d infrared image via the 3d model space. In chapter 3, the projection of the building model polygons into the infrared input image was discussed. Furthermore, a projection of 2d texture space into the 3d model coordinate system has to be calculated setting a certain resolution (pixel/m) of the textures.

4.1 Generating the template textures for the 3d model space

The resulting textures are images given in pixel coordinates. The size of these images is calculated from the size of the corresponding model face and the resolution coefficient that indicates the texture resolution in pixel per meter. The length and the height of the model faces are known and the texture resolution coefficient is set according to the expected resolution of the textures. Next, the pixel coordinates are translated into texture coordinates ranging from 0 to 1 in width and height. This is done by interpolating the texture coordinates for every texture pixel from the known texture coordinates of the vertices of the face. With this texture coordinates and the unit vectors of the face, the texture points are translated into the 3d model space. Two unit vectors define the width and height direction of the face and thus are directly used to interpolate the x,y,z values of the texture points from their texture coordinates. The resulting 3d coordinates and the pixels of the texture image are stored together.

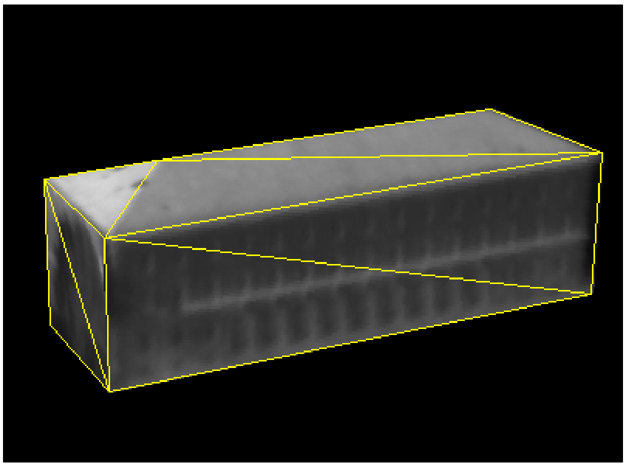
4.2 Extraction of intensity values of the texture pixels

Both the transformation of 2d texture pixel coordinates to 3d model coordinates and the transformation of 3d model coordinates to 2d image coordinates are known. Every 3d texture point is assigned to a pixel of the texture. To avoid holes in the face textures, a backwards projection is used. Every texture point is projected into the image plane of the input IR image and its value is bilinear interpolated from the four neighbouring pixel values. One remaining task is the visibility analysis of these texture points. This is done using the projection of the polygons described in Chapter 3. The projected polygon outlines (see Fig. 4) are now filled in the 2d image space line wise with linear interpolated depth value. Every pixel of every filled polygon is checked against a depth buffer and will only be accepted for the resulting visibility image, if its depth value is smaller than the already stored depth value in the buffer. This visibility check is optimized using computer graphic methods like backface culling and polygon clipping to remove invisible faces in an early process stage. Invisible faces are marked and their texture points excluded from the texture point projection into the image space. The resulting visibility image contains for every image pixel its depth value and the index of the visible face.

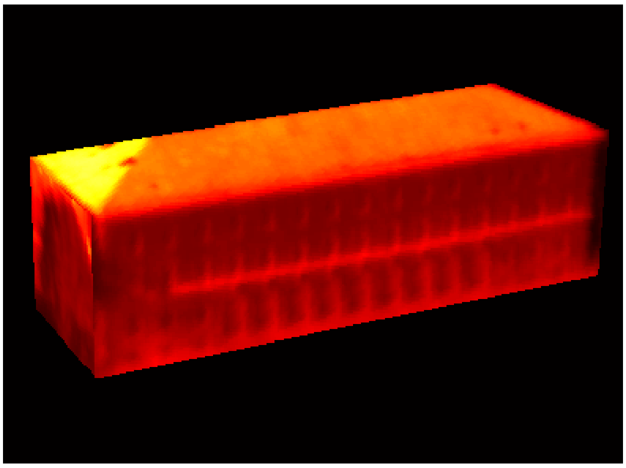
The texture points are now projected into the image space. In this step, every projected texture point is checked against its four neighbours for the index of the visible face. The texture point's value is only interpolated, if its origin face is at least the origin of one of the four neighbouring pixels. Otherwise, the texture point is not visible and its value is set to false. The interpolated values of the texture points are now copied to the texture image by using the stored connection calculated in chapter 4.1.

5. FIRST RESULTS

First results of the texture mapping are shown in Fig. 5. The texture was extracted for a roof and two facades of one of the TUM buildings. The textured building could be seen close to the lower right corner in Fig. 1. The texture of the front facade and the roof was mapped from image #13200 while the texture of the left wall was taken from one of the images from sequence of the first flight path (#1).



a



b

Figure 5. Textured building model. a) gray valued texture with model triangles, b) texture in a colour palette.

The two roof planes visible in the textured model can be easily distinguished. The left part of the roof has a much higher temperature probably because of its southern aspect. Some objects like windows can be seen in the front facade. The left wall is almost fully occluded by a tree growing in front of it. The dark stripe located left from the front facade appeared probably because of a model inaccuracy. The model was acquired by a vectorization of aerial images in which only the very front edge of the roof (eaves) is visible so that the modelled front plane of the building is shifted from its true position. This results also in a brighter stripe located in the bottom of the front facade which actually is the part of the pavement.

The advantage of storing textures assigned to same faces, but coming from different acquisitions (terrestrial or airborne, day or night) or different sensors (visual, midwave or longwave infrared) is the possibility of combined visualisation. Figure 6 shows an example of combining terrestrial and airborne textures. A criterion to select textures can be the texture resolution (see Chapter 4.1)

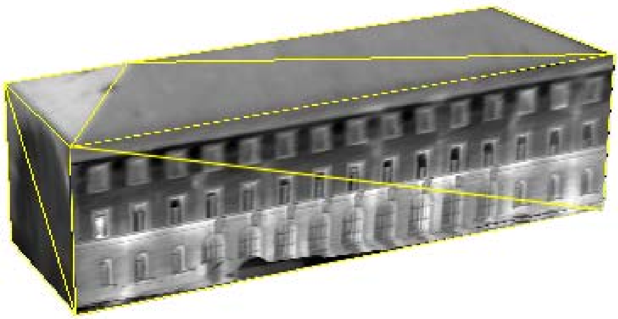


Figure 6. Building with textures from airborne and terrestrial IR images. Front fasade: terrestrial images (Hoegner & Stilla, 2007a), other faces: airborne images

6. DISCUSSION

The first results of our research show that the direct measurement of the exterior orientation elements with the GPS/INS system provides data which are accurate enough to perform the texture mapping of the LOD2 city model. However, it should be mentioned that the corrections resulting from IMU misalignment and shift should be calculated and added to the observations. Very meaningful are the corrections to the angular element of exterior orientation since even a small misalignment could cause a large error. Especially, when a camera is located far away from an object this effect is significant. Results of the projection of the model lines into the image show that in some cases lines do not match the proper image edges. These mismatches results both from the model generalization and from inaccurate exterior orientation elements.

In Table 1 we have seen that the difference between measured values from IMU and values calculated from bundle block adjustment for the roll angle shows a high standard deviation. It can also be seen that subsequent values differ significantly from each other. Figure 7 shows the roll angle over time measured by IMU with a sampling rate of 200 Hz.

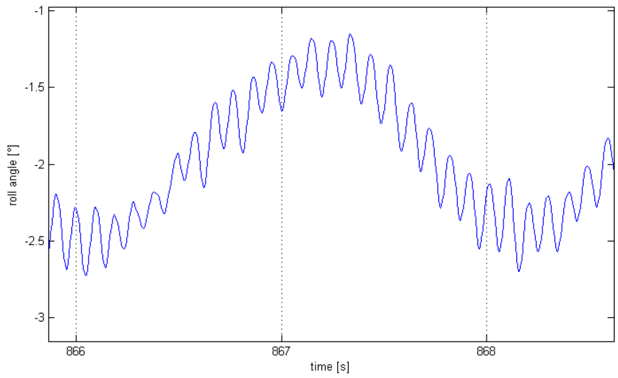


Figure 7. Plot of yaw angle over time

The changes can be described by a slower movement combined with a faster movement. The faster movement has a frequency of about 10 Hz (see dotted lines for a time interval of 1 second). If we consider the frame rate of the camera (25 Hz) it is remarkable that the half frequency is very similar to the frequency of helicopter movements recorded by the IMU. In case that the time stamp for the camera exposure is shifted against the IMU clock, we will receive significant differences in the measured roll angle.

Additionally, further effect can result in texture mapping errors. In case that the IMU has a vibration of 10 Hz we can assume that the camera is also vibrating with a certain frequency which may result in a misalignment. Not investigated and covered in this paper is a rolling-shutter effect. In contrast to a frame shutter the rolling shutter works differently, in that the photodiodes (pixels) do not collect light at the same time. All pixels in one row of the imager collect light during exactly the same period of time, but the time light collection starts and ends is slightly different for each row. The top row of the imager is the first one to start collecting the light and is the first one to finish collecting. The start and end of the light collection for each following row is slightly delayed. The total light collection time for each row is exactly the same, and the delay between rows is constant.

For increasing the accuracy of texture mapping which is required for models with higher level of detail (e.g. LOD3) a more detailed investigation on error analysis is required. A different strategy to increase the mapping accuracy is an automatic matching of building models with image features. This allows a subsequent recalculation of exterior orientation elements of the camera (Stilla at al, 2000, Frueh et al., 2004). A further step will be a matching of parts instead of the entire model of the building or building complex.

ACKNOWLEADGMENTS

The authors would like to thank FGAN-FOM, Ettlingen, for providing image of the flight campaign and K. Eder (TUM-FPF) for providing ground control information.

REFERENCES

Cramer M. (2002) Experiences on Operational GPS/Inertial System Calibration in Airborne Photogrammetry, Journal GIS-Geoinformationssysteme, 6/2002, Wichmann Verlag, Germany, pp. 37-42.

Eugster H., Nebiker S. (2007) Geo-registration of Video Sequences Captured from Mini UAVs – Approaches and Accuracy Assessment, 5th International Symposium on Mobile Mapping Technology - Padua, Italy.

Eugster H., Nebiker S. (2008) UAV-Based Augmented Monitoring – Real-Time Georeferencing and Integration of Video Imagery with Virtual Globes, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 27. Part B1. Beijing.

Frueh C, Sammon R, Zakhor A (2004) Automated Texture Mapping of 3D City Models With Oblique Aerial Imagery, Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'04).

Grewal, MS, Weill, LR, Andrews AP (2007) Global Positioning Systems, Inertial Navigation, and Integration, WILEY-INTERSCIENCE, A John Wiley & Sons, Inc., Publication, pp. 382-424.

Hebel M., Stilla U. (2007) Automatic registration of laser point clouds of urban areas. PIA07-Photogrammetric Image Analysis 2007. International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences, Vol 36 (3/W49A): 13-18

Hoegner L, Stilla U (2007a) Automated generation of 3d points and building textures from infrared image sequences with ray casting. In: Stilla U, Meyer H, Rottensteiner F, Heipke C, Hinz S (eds) PIA07 - Photogrammetric Image Analysis 2007. International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences, Vol 36(3/W49B): 65-70

Hoegner L, Kumke H, Meng L, Stilla U (2007b) Automatic extraction of textures from infrared image sequences and database integration for 3D building models. PFG Photogrammetrie Fernerkundung Geoinformation. Stuttgart: Schweizerbartsche Verlagsbuchhandlung. 2007(6): 459-468

Hoegner L, Stilla U (2008) Case study of the 5-Point algorithm for texturing existing building models from infrared image sequences. XXI ISPRS Congress, Proceedings. International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences, Vol 37(B3B): 479-484

Nebiker S, Christen M, Eugster H, Flueckiger K., Stierli C. (2007) Integrating Mobile Geo Sensors into Collaborative Virtual Globes – Design and Implementation Issue, 5th International Symposium on Mobile Mapping Technology - Padua, Italy.

Stilla U, Sörgel U, Jäger K (2000) Generation of 3D-city models and their utilisation in image sequences. International Archives of Photogrammetry and Remote Sensing. Vol. 33, Part B2, 518-524