SEMI-AUTOMATIC CAMERA CALIBRATION AND IMAGE ORIENTATION USING THE CIPA REFERENCE DATA SET

Frank A. van den Heuvel
Delft University of Technology, Department of Geodesy
Thijsseweg 11, 2629JA Delft, The Netherlands
E-mail: F.A.vandenHeuvel@geo.tudelft.nl

KEY WORDS: Architecture, Line-photogrammetry, Camera Calibration, Image Orientation, Object Reconstruction, Automation

ABSTRACT
The goal of this paper is to demonstrate the applicability of line-photogrammetric methods by applying them to a subset of the images of the CIPA reference data set. This data set is set up by CIPA (The ICOMOS / ISPRS Committee for Documentation of Cultural Heritage) in 1999 and consists of images of a historic building: the old city hall of Zürich. Although object reconstruction is regarded as a final goal of a photogrammetric recording, this paper focuses on two tasks in architectural photogrammetry: camera calibration and image orientation. The line-photogrammetric methods applied for these tasks have been developed in the past years. These methods have in common that they require the extraction of image lines and information on the orientation of the related object lines. The methods are not discussed in detail in the paper.
For camera calibration manually and automatically extracted straight image lines of five images were used. The required parallelism and orthogonality information was obtained automatically by vanishing point detection. Correspondence between line features in different images is not required. Camera parameters and their precision are estimated in a least-squares adjustment. The results correspond well with the camera parameters provided in the CIPA reference data set itself and those provided in the literature.
The semi-automatic method for image orientation relies on line extraction and vanishing point detection as well. Furthermore, coplanarity of object lines that reside in a façade is assumed. Fully automatic relative orientation is not possible for all four image pairs used in the experiments. Depending on the image configuration and settings of a few parameters, one or two image points have to be measured for successful orientation. These points are chosen on the corners of the building and thus in the overlap of consecutive object models. This allows the scale to be transferred from one model to the next. Then this method for image orientation results in a partial reconstruction of the building.

Figure 1: The subset of used images. Images 8, 9, 10 (top row), 11, 6, 3, and 16 (bottom). Images 8, 9, 10, 11, and 6 have been used for calibration. The images in the bottom row are used for the image orientation experiments.

1. INTRODUCTION

1.1. Line-Photogrammetry

In the past years line-photogrammetric methods were developed for the main tasks in architectural photogrammetry, i.e. camera calibration (Heuvel, 1999a), image orientation (Heuvel, 2002), and object reconstruction (Heuvel, 1999b). These methods have in common that they require the extraction of image lines (manually or automatically) in combination with information on the orientation of the related object lines. The goal of this paper is to briefly introduce these methods and to demonstrate their applicability by applying them to images of the CIPA reference data set.

1.2. The images of the CIPA reference data set

The CIPA reference data set consists of two sets of images taken with two different digital cameras. Here, only a subset of the images taken with the Olympus C1400 is used. Due to the manual interaction required for the reconstruction of a structured object model, one of the goals of this research is to use only a minimum number of images in order to make the modelling more efficient. Therefore, only four images taken from the corners of the building are selected for image orientation and object reconstruction. The object reconstruction is not discussed in this paper because it is presented in (Hrabacek and Heuvel, 2000). For camera calibration five images were selected. Figure 1 gives an overview of the images used in the experiments.

1.3. Structure of the paper

The paper is split in two main parts. First the camera calibration is discussed in section 2, and then the relative orientation of the four corner images is presented in section 3. These two sections start with a subsection in which the approach is briefly explained, followed by the results of the experiments using the images of the CIPA reference data set. Conclusions are presented in section 4.

2. CAMERA CALIBRATION

2.1. The method for camera calibration

The method for camera calibration is summarised in the following three steps (Heuvel, 1999a):

1. Extraction of straight image lines
2. Detection of the object orientation of the image lines
3. Estimation of camera parameters from parallelism and perpendicularity constraints on image lines

The first two steps can be performed manual as well as automated. In the latter case, a line-growing algorithm is used for image line extraction (Heuvel, 2001), and a vanishing point detection method is applied for the detection of the three dominant object orientations (Heuvel, 1998). The quality of the estimated parameters is dependent on a correct vanishing point labelling of image lines performed in step 2. With the camera parameters unknown, the automatic detection of the three vanishing points that correspond with edges of the three orthogonal object space orientations is more critical than when using a calibrated camera. Two factors play a role. First, unknown lens distortion hinders the straight line detection and prohibits the detected lines intersecting in one point in image space. Second, unknown focal length and principle point make it impossible to limit the search space after detection of one or two vanishing points. As a result each vanishing point is detected independent of previously detected vanishing points. The procedure below is applied to five images of the CIPA data set. The lens distortion is determined first, followed by the other parameters of the camera model, i.e. focal length and principle point.

1. Start with vanishing point detection for those images that contain only one façade of the building. For these images only two vanishing points are to be detected, one for the vertical object orientation, and one for the horizontal object edges.
2. If only images with two (presumably orthogonal) façades are available, only the vanishing point that corresponds to the vertical object orientation is detected and its lines used for estimation of the lens distortion. This approach assumes limited camera tilt and rotation around the optical axis.
3. Estimate lens distortion using the detected and labelled lines of at least one, but preferably more images. In the next steps the estimated lens distortion is eliminated from the observations.
4. Detection of three vanishing points. Three-point perspective imagery (for example image 10 in Figure 1) is required for the estimation of the focal length and the principle point. When the optical axis is nearly horizontal and thus a one-point or two-point perspective remains (see the images in the bottom row of Figure 1) the principle point location in horizontal direction (camera x-axis) cannot be estimated, or only with very low precision.
5. Estimate focal length and principle point using the detected and labelled lines of at least one but preferably more images.

2.2. Camera calibration using the CIPA data set

Image lines are extracted for all the selected images of the CIPA data set. For the line-growing algorithm used for straight line extraction two parameters were set. First the parameter for the minimum gradient strength was selected in such a way that most of the characteristic features of the building – especially the windows – were extracted. The second parameter is the minimum length in pixels of an extracted image line. This parameter was fixed to 30 pixels for all the images used for camera calibration.

Lens distortion

Detection of two vanishing points was performed on two images (numbers 8 and 9) that show only one face of the building. This is step 1 of the procedure described in the previous section. The result is shown in Figure 2. In this and following figures the image lines are colour-coded using the line labelling results of the vanishing point detection.

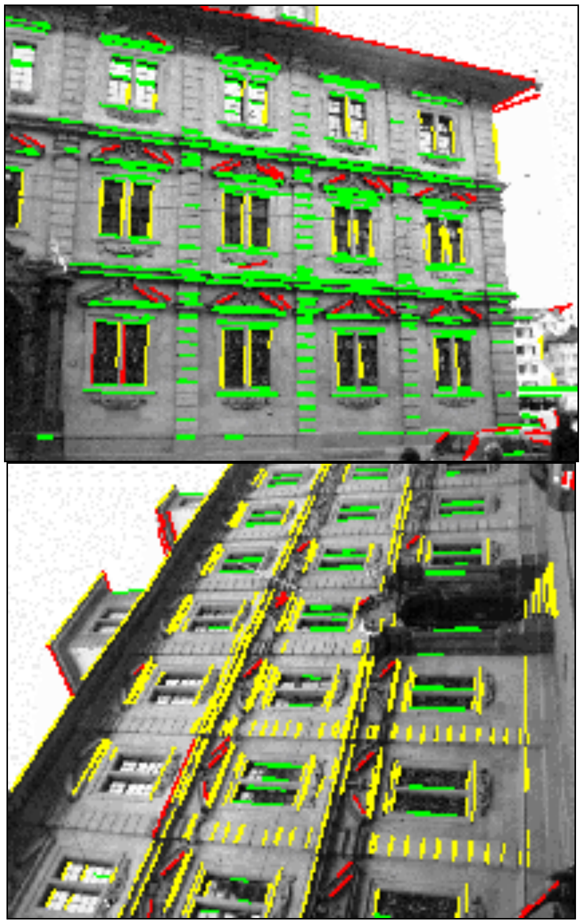

Figure 2 : color-coded image lines of the vanishing point detection for image 8 (top) and 9 (bottom).

For the estimation of the lens distortion (parameter k1) parallelism condition equations for the lines in images 8 and 9 have been used. The estimated value for k1 is -0.570 x10^{-3}. This value is 7.0 times its estimated standard deviation and thus significant. The a priori standard deviation was set to 1 pixel (= 0.006445 mm) for all experiments.

Focal length and principle point

Lines of three vanishing points are required for the focal length and principle point estimation. The lines were extracted and grouped automatically by vanishing point detection, and also measured manually (Figure 3). The estimated parameters using the results of each image separately were not consistent (Image 3 and 6 as an example in Table 1). Especially the difference in the estimated focal length was considerable.

	Image 3	Image 6	Image 3 Manual	Image 6 Manual
Variance fac.	1.027	0.903	1.152	0.132
d.o.f.	487	454	7	7
Max. residual (mm / pixel)	0.018 / 2.8	0.018 / 2.8	0.010 / 1.6	0.003 / 0.4
Focal length in mm (σ)	8.668 (0.052)	9.663 (0.079)	8.514 (0.046)	10.092 (0.058)
Princip pnt. y in mm (σ)	-0.200 (0.057)	-0.053 (0.067)	-0.085 (0.066)	0.376 (0.076)

Table 1 : Results of camera parameters estimation using automatically and manually extracted and labeled lines.

The location of the principle point could not be estimated in x-direction due to the near two-point perspective of the imagery. The precision of the parameters estimated from manual measurements is better, although the number of lines used is only a fraction of the number of automatically extracted lines. The reason is found in the length of the manually extracted lines. Therefore, hereafter only the manually extracted lines are used for the camera calibration. This choice is also to be preferred because it limits the assumptions of parallelism and perpendicularity to the measured edges of the building.

The source of the inconsistency was traced by estimating the parameters using the manual measurements of combinations of two images. The results are shown in Table 2. Only the fit of the combination in which image 3 is missing is of good quality; the overall test is accepted (critical value 1.93).

	Imag 3-6	Imag 3-11	Imag 6-11
Variance factor	31.01	31.63	1.435
d.o.f.	16	16	16
Max. residual (mm / pixel)	0.058 / 8.8	0.052 / 8.1	0.012 / 1.8
Focal length in mm (σ)	9.124 (0.035)	9.234 (0.032)	10.056 (0.037)
Principle point y in mm (σ)	0.158 (0.048)	0.028 (0.046)	0.267 (0.050)

Table 2: Camera parameter estimation using the manual measurement of combinations of images

Image 10 was now added in the estimation in order to verify the results. In Table 3 adjustment results using the manual measurements of three images are presented. As the critical value is 1.8 (5% significance level) both adjustments are accepted. Due to the 90 degree rotation around the optical axis of image 10 the principle point x co-ordinate can also be estimated. Note that the manual line measurements are not suitable for estimation of lens distortion. The results in the last column of Table 3 are the final estimates of the camera parameters used for image orientation described in the next section.

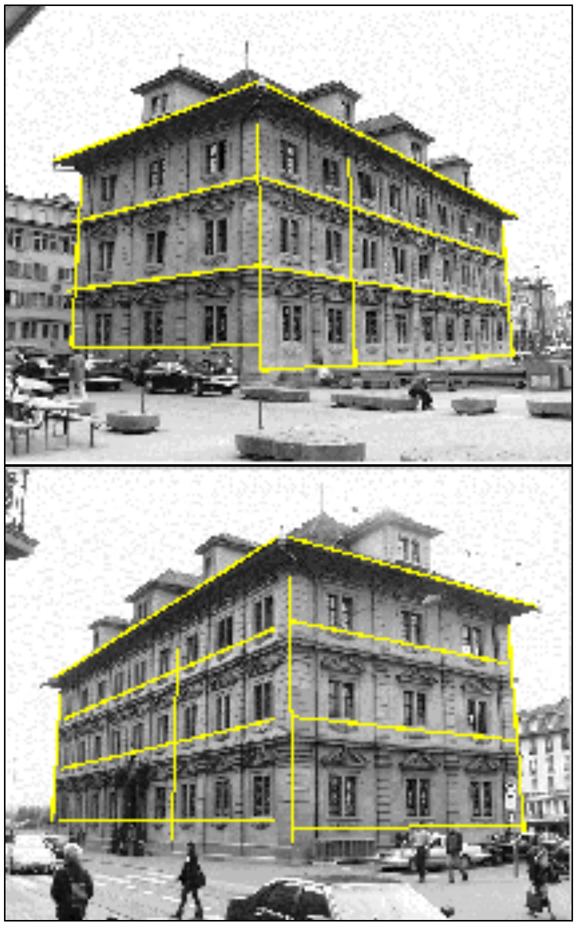

Figure 3: Manual line measurements for camera calibration (images 3 (top) and 6)

	Image 6, 10, 11	Image 6, 10, 11 Final
Variance factor	1.553	1.211
d.o.f.	26	25
Max. residual (mm / pixel)	0.0131 / 2.0	0.0100 / 1.6
Focal length in mm (σ)	10.103 (0.029)	10.116 (0.030)
Principle point x in mm (σ)	-	-0.146 (0.041)
Principle point y in mm (σ)	0.222 (0.047)	0.220 (0.046)

Table 3: Camera parameter estimation using the final combination of three images

The CIPA data set was also analysed by others (Streilein et al., 2000), (Rottensteiner et al., 2001). In the paper by Rottensteiner the images taken with the Olympus camera are split in two groups, one with a focal length of 8.598 mm and the other with a focal length of 10.132 mm (balanced lens distortion). Obviously, the setting of the zoom lens is different for the two sets of images. The camera parameter values from (Rottensteiner et al., 2001) derived with bundle adjustment software Orpheus, are used as a reference. Comparison with camera parameters provided in this set showed a good match. The lens distortion parameter is not compared because the number and definition of distortion parameters differs in the set determined with Orpheus. The difference in the focal length is 3.0 times its standard deviation. Differences in the principle point location are smaller than 2.0 times the standard deviation.

3. IMAGE ORIENTATION

3.1. The method for image orientation

The method for image orientation aims at full automation and is described in (Heuvel, 2002). The camera is assumed to be calibrated, in this case by the method outlined in section 2. The method relies on automated straight line extraction and vanishing point detection, and results in a model coordinate system that is aligned with the building. The building has to fulfil the following requirements for the method to be successful:

- Parallel and perpendicular straight object edges
- Coplanarity of the edges in the façades

Successful orientation can require a few manual measurements to allow for reliably resolving ambiguities inherent in the vanishing point detection and in the repeating and symmetric structures present in most buildings. Furthermore, the manually measured points reduce the computational burden considerably and can be used to guarantee the required overlap of at least one point between consecutive models needed to transfer scale from model to model.

The semi-automatic method for relative orientation outlined above is successfully applied to four images of the CIPA reference data set (Figure 1). In Figure 4 two views on the resulting approximate reconstruction are shown. Relative scale of consecutive models was determined using a manually measured point on each corner of the building. In fact this method results in an approximate and partial reconstruction of the building. The fully automatic relative orientation of two of the four images is described in the next section.

3.2. Image orientation using the CIPA data set

Two characteristics of the CIPA data set images play a major role in prohibiting the method for automated relative orientation to be successful in all cases. The first one is the considerable differences in image scale between images. This is due to the obliqueness of the selected images relative to the façades, as well as to the large differences in the object to image distance (see image 16 in Figure 1). Secondly, the repeating structures in the form of the many identical windows make the detection of the correspondence ambiguous. To some extent it is possible to adapt the parameters to these characteristics. In the example presented here, straight lines are extracted for images 3 and 6 with a minimum line length set to 40 pixels. The maximum distance between two lines to decide for their intersection was set to 10 pixels (Figure 5 on the next page). When these parameters were set to 30 and 5 pixels respectively a correct solution could only be found with two additional manual point measurements. The reason is found in the symmetry of the building; the long façades (on the left in image 6 and on the right in image 3) are erroneously matched when many lines in these façades are available. A longer minimum line length (40 instead of 30 pixels) avoids this.

In Table 4 some statistics of the experiment are presented. The table demonstrates the reduction of the computational burden, inherent in the method, to manageable proportions. The number of possible correspondences is considerably reduced (from 204x214= 43656 to 2378) by checking characteristics of the intersection of the image lines, such as the orientation of the lines in object space that is available from the vanishing point detection. The correspondence hypotheses are being clustered based on a statistical coplanarity test for each combination of two correspondences. Not all combinations are tested; two correspondences with different orientation of the facade are not combined. (# potential tests Table 4). Furthermore, a number of tests can be excluded because of an unlikely relative position of the two images. For instance, the angle between the relative position vector and the vertical is required to be close to 90 degree (threshold set to 10 degree). The clustering results in 3706 clusters of correspondences. For each cluster an overall adjustment is set up. The correspondence with the largest rejected statistical test is removed from its cluster and the adjustment is repeated.

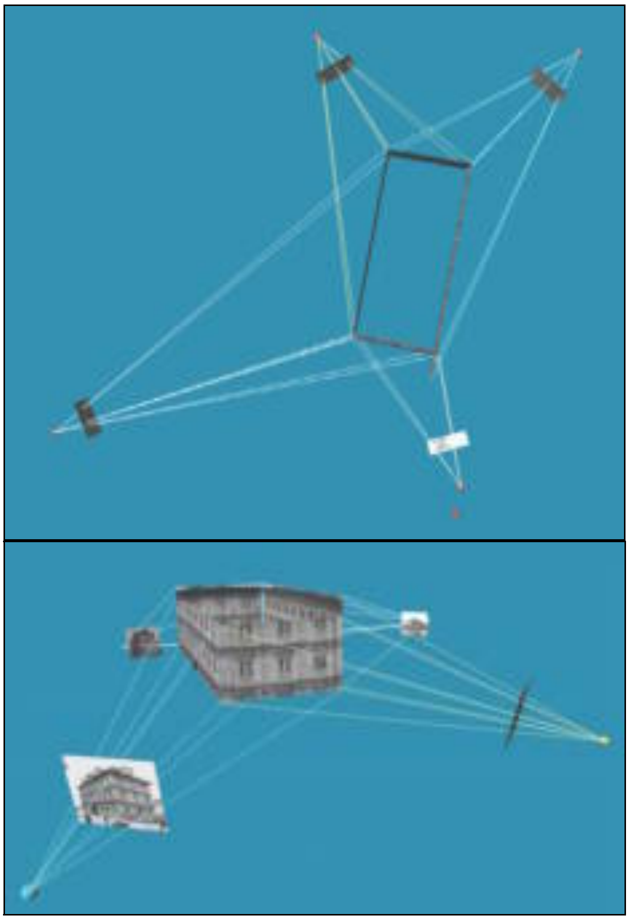

Figure 4: Two views on the approximate reconstruction from the 4 images using the semi-automatic method for orientation

Parameter	Value
Minimum line length	40 pixel
Maximum distance for point creation	10 pixel
# created points image 1 / 2	204 / 214
# correspondence hypotheses	2378
# potential tests	1,182,214
# computed tests	213,991 (18.1%)
# accepted tests	26,063 (3.2%)
# clusters	3706
Maximum # correspondences	27
# clusters after testing	97
Maximum # correspondences	22
Test (ratio with critical value)	3.65

Table 4: Statistics of correct solution for the automatic relative orientation of image 3 and image 6.

The cluster that contains most correspondences after this iterative testing procedure is selected. This cluster contains 22 correspondences and is visualized in Figure 5 (bottom) and Figure 6. The overall test is rejected, although individual correspondence tests are all accepted. This is probably mainly due to the presence of some structure in the façade that violates the coplanarity assumption of the object points and lines.

Figure 5: Top: results of line extraction and vanishing point detection of images 6 (left) and 3 (right). Middle: points from line intersections. Bottom: result of correct correspondence detection.

Figure 6 shows the position of the two images relative to the façade and the reconstructed part of the façade. This reconstruction results from creating a bounding box around the detected points of the façade. Their 3D coordinates are computed through forward intersection.

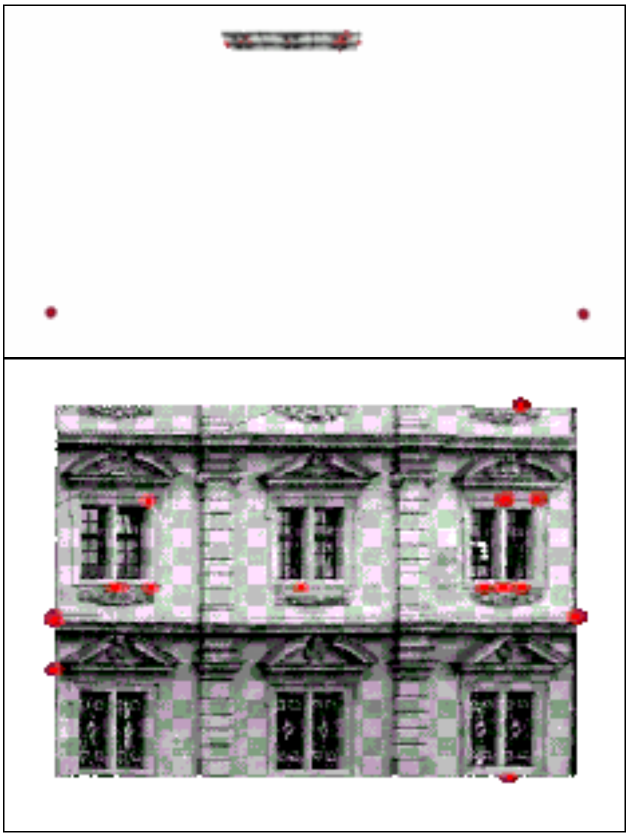

Figure 6: Two views of the rectified part of the façade that contains the 22 corresponding points.

4. CONCLUSIONS

The processing of images of the CIPA reference data set has demonstrated the capabilities and limitations of the developed methods for camera calibration, image orientation, and object reconstruction. Object reconstruction using the same four images as used for the orientation experiments is presented in (Hrabacek and Heuvel, 2000). Figure 7 shows a view of the reconstructed model that is represented in VRML format. All methods used have the following characteristics in common:

- The incorporation of knowledge on the construction of the building
- The use of image line features, manually or automatically extracted
- The use of a limited number of images

Calibration of the camera was performed in a semi-automatic way with five images of the reference data set. Estimation of interior orientation parameters from single images showed that two zoom settings were present in the set of selected images. The quality of the results of single image camera calibration depends greatly on the image characteristics. For a full and accurate camera calibration, several images with different orientation relative the building are needed. Correspondence between the images is not required.

The method for image orientation aims at automation. Tests with images from the CIPA data set show that correct relative orientation can be obtained, but as in the method for automated camera calibration, the chance of success greatly depends on the image configuration and characteristics. Especially the repeating structures (windows) in the façades lead to an ambiguous problem. Between two and three manually measured corresponding points were needed to obtain a correct and reliable solution, and at the same time allowed to determine the relative scale of consecutive models. The paper shows that automatic relative orientation of two images is possible. However, success depends on the settings of a few parameters. A partial object model results as a by-product of the methods for image orientation.

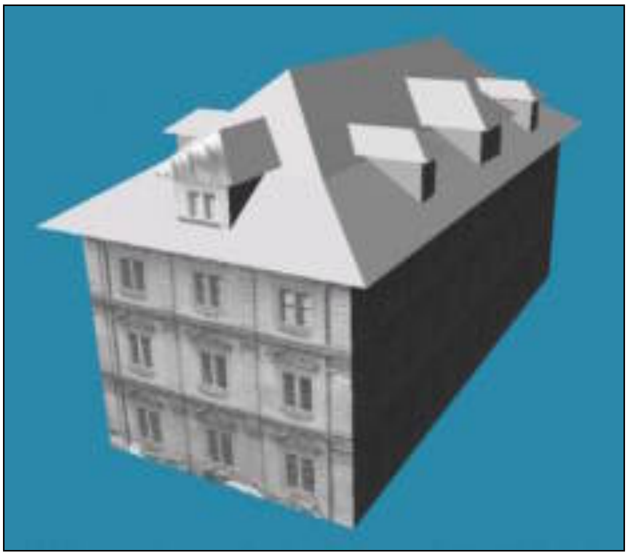

Figure 7: Reconstructed model, partially texture-mapped.

REFERENCES

Heuvel, F.A. van den, 1998. Vanishing point detection for architectural photogrammetry. H. Chikatsu and E. Shimizu (Editors). International Archives of Photogrammetry and Remote Sensing, Vol. 32 part 5, Hakodate, Japan, pp. 652-659.

Heuvel, F.A. van den, 1999a. Estimation of interior orientation parameters from constraints on line measurements in a single image. P. Patias (Editor). International Archives of Photogrammetry and Remote Sensing, Vol. 32 part 5W11, Thessaloniki, Greece, pp. 81-88.

Heuvel, F.A. van den, 1999b. A Line-photogrammetric mathematical model for the reconstruction of polyhedral objects. S.F. El-Hakim (Editor), Videometrics VI. proceedings of SPIE. SPIE, Vol. 3641, San Jose, USA, pp. 60-71.

Heuvel, F.A. van den, 2001. Object reconstruction from a single architectural image taken with an uncalibrated camera. Photogrammetrie Fernerkundung Geoinformation, 2001(4): 247-260.

Heuvel, F.A. van den, 2002. Towards automatic relative orientation for architectural photogrammetry. P. Patias (Editor). International Archives of Photogrammetry and Remote Sensing, Vol. 34 part 5, Corfu, Greece, pp. 227-232.

Hrabacek, J. and Heuvel, F.A. van den, 2000. Weighted geometric object constraints integrated in a line-photogrammetric bundle adjustment, ISPRS congress. International Archives of Photogrammetry and Remote Sensing, Vol. 33, part B5, Amsterdam, pp. 380-387.

Rottensteiner, F., Grussenmeyer, P. and Geneva, M., 2001. Experiences with the digital photogrammetric program package Orpheus based on CIPA's "Zurich city hall" dataset for architectural photogrammetry, XVIII CIPA Int. symposium. proc, Potsdam, Germany, pp. 639-646.

Streilein, A., Hanke, K. and Grussenmeyer, P., 2000. First experiences with the "Zurich city hall" data set for architectural photogrammetry, ISPRS congress. International Archives of Photogrammetry and Remote Sensing, Vol. 33 part B5, Amsterdam, pp. 772-779.