

THE PRECISION OF A DIGITAL CAMERA

George Vosselman and Wolfgang Förstner
Institute for Photogrammetry, University of Stuttgart
Keplerstraße 11, D-7000 Stuttgart
West Germany
Commission I

Abstract: A testfield containing a large number of black targets on a white background has been recorded by a digital camera from many points of view. In each digital image, the targets have been located using elementary image processing techniques. Precise coordinates were obtained by matching the targets with artificial masks. The precision of these coordinates was calculated in a bundle block adjustment with self-calibration parameters. The achieved precision amounted to 0.03 pixel, corresponding to 0.8 μm in the image plane.

0. Introduction

The quality of image analysis procedures in image processing and pattern recognition heavily depends on the radiometric and geometric quality of the used sensors. In close range applications mainly video- and - due to their higher geometrical stability - increasingly CCD-cameras are used, as they produce a digital image which can immediately be further processed. The size of contemporary CCD-cameras varies between 128 x 128 and 1320 x 1035 pixels with pixel sizes between 6.8 and 60 μm (cf. Grün 1987).

For a geometric analysis of the image, the calibration of the camera has at least to prove, that

- a. the neighbouring pixels and

- b. the pixel positions show a sufficiently regular and stable geometry.

The ultimate aim could be to calibrate each individual pixel according to its own sensitivity and geometry and have a processor attached to it, to compensate for these effects in the working phase.

The results of this study suggest, that the geometry of the sensors may be fully sufficient for purposes, where the location of small objects (with a diameter larger than, say, 4 pixels), needs to be estimated with a precision of 0.1 pixels or worse, even if no radiometric calibration of the sensor is performed. This corresponds to the accuracies consistently obtained from empirical studies.

The theoretical studies (cf. Förstner 1982, 1984), however, claim, that even much higher accuracies should be obtainable under well controlled conditions, namely standard deviations of 1/20 to 1/100 pixel. It was the primary goal of this study to find out whether these accuracies can really be obtained.

Testfield calibration is the only way to prove the absolute accuracy of a camera with respect to target location. The accuracy requirements are high as for a digital camera of 256 x 256 pixels and a standard deviation of 1/50 pixel the coordinates of the targets have to be measured with a relative accuracy better than 1 : 10 000 in 3D. Up to this point we did not coordinate the targets of our testfield. The precision of the target location process within the images can however reliably be estimated if sufficiently many images are used for simultaneously determining the 3D-coordinates of the targets in a joint estimation process, possibly using additional parameters to compensate for systematic effects, which later on can be used as correction terms.

Originally the testfield was designed with three scopes in mind:

- a. determining the accuracy of the camera,
- b. demonstrating, that an automatic recognition, numbering and location of the targets is feasible and
- c. comparing the empirical results with the theoretical predictions.

This paper describes the testfield, the automatic procedure for deriving labeled image coordinates for the targets and the accuracy results achieved so far.

1. Camera

The used camera is a solid state CCD video camera of type C1000-35M, made by Hamamatsu. It uses the two dimensional Hitachi He 97211 sensor, consisting of 256 rows with 320 photodiodes of a MOS-type. The photodiodes have a light sensitive area of $21.6 \mu\text{m} \times 21.6 \mu\text{m}$. The distance between the sensor elements is $27 \mu\text{m}$ in both row and column direction. Due to the interface, only the central 256 columns can be addressed. We used the central 240×240 patch for the recordings for this investigation. The most important camera data are given in table 1. More details can be found in (Gülch, 1984, 1985).

Imaging device	MOS-type 2-dimensional photo sensor array
Number of pixels	256 rows x 320 columns
Scanning area	8.8 x 6.6 mm
Scanning system	progressive sampling
Resolution	in row direction 240 lines (at center) in column direction 190 lines (at center)
Distortion	less than 1%
Signal-to-noise ratio	46 dB
Spectral response	400 - 1050 nm
Resolution range	256 grey levels

Table 1. Manufacturer specifications of the Hamamatsu C1000-35M video camera

2. Testfield

A reliable calibration demands a high redundancy of observations in the bundle block adjustment. This demand has its impacts on the design of the testfield and on the configuration of the recordings.

2.1. Design

The testfield should contain as many targets as possible. Having many targets, location of these targets can only be done within a reasonable period of time, if the whole process of target recognition and location can be done automatically. For this reason, the target identification numbers have been coded in the testfield. Using four different target shapes (squares, triangles, crosses and circles), two target sizes and straight lines to cluster the targets, a combination of shape, size and topology enables an unambiguous identification of each target. A testfield has been constructed with 64 squares (fig. 1.). Each square contained 16 targets, equally divided among the different shape and size classes. The sizes of the testfield and its targets are specified in table 2., first column. For the investigation presented here, we only used the central 16 squares, together containing 256 targets. The second column of table 2. gives the target sizes in pixels for a recording of exactly 4×4 squares.

2.2. Configuration of the Recordings

The testfield has been recorded from nine different positions, as sketched in fig. 2. At all positions four recordings were made, rotating the camera around its axis by 90 degrees after each recording. These camera rotations

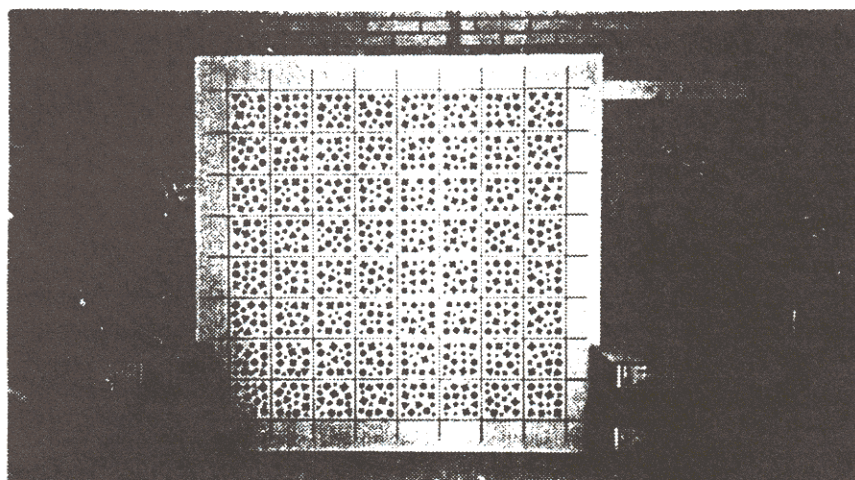


Fig. 1. The testfield

	sizes in mm	sizes in pixels in a recoding of 4x4 squares
size of the small symbols	9 - 15	4 - 7
size of the large symbols	17 - 25	8 - 12
thickness of the straight lines	2	1
distance between the straight lines	125	60
testfield size	1200	

Table 2. Geometric parameters of the testfield

ensure that differences between the pixel size in row direction and the pixel size in column direction can be determined by a self-calibrating bundle adjustment.

The different positions of the camera were simulated by rotations of the testfield (fig. 3.). The three phi-rotations of the testfield were 0, 30 and

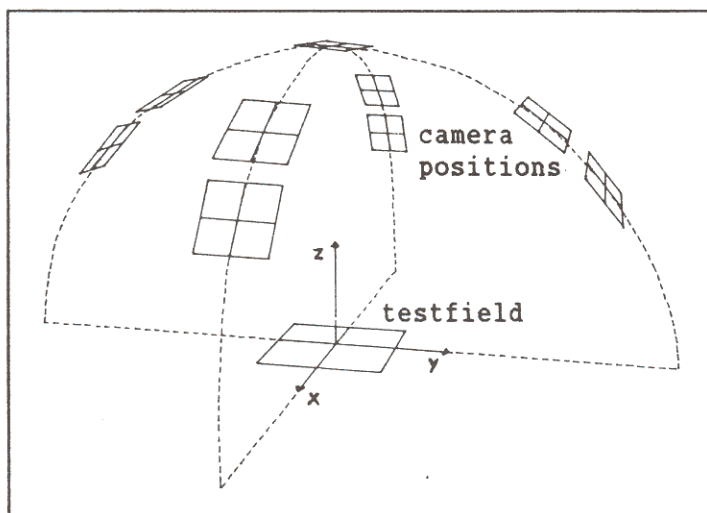


Fig. 2. Configuration of the recordings

45 degrees. The focal length of the camera was 57.5 mm. The distance between the camera and the testfield was about 5 meters.

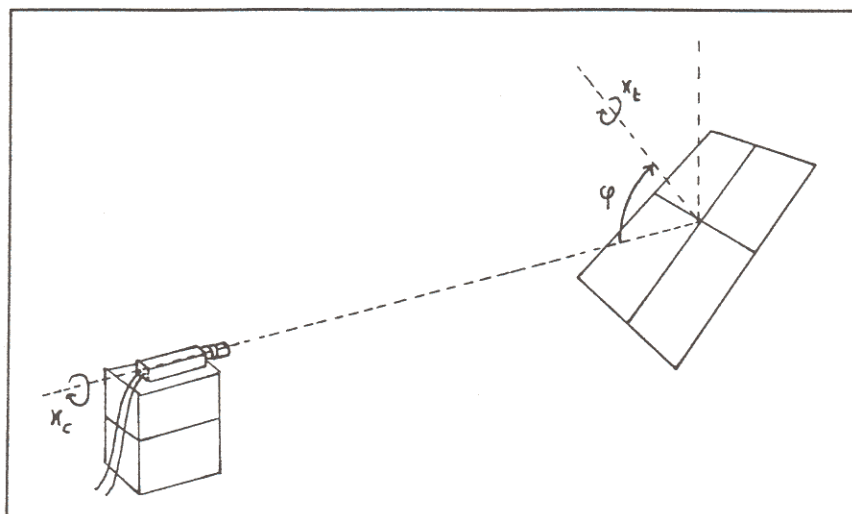


Fig. 3. Simulation of the configuration

3. Automatic Target Recognition and Location

A first step in finding the targets is the recognition and location of the straight lines. Once their position is known, an image can be rectified to a normal image. After thresholding this image, the targets are traced in the binary image and classified. Finally the target numbers are determined.

3.1. Recognition and Location of Straight Lines

Since the lines are about one pixel thick, the grey value curvature is very strong at pixels on a line. Thus, after convolving the original image with the Laplace operator, the lines stand out against the other features. Thresholding this curvature image yields a binary image in which most of the "black" pixels belong to one of the straight lines. In this image the lines can be recognized easily with a Hough transformation. A subsequent least squares line adjustment is used to improve the precision. In this way all straight lines in the recorded images could be located with a precision of 0.1 pixel or better.

3.2. Target Tracing

Knowing the locations of the lines, one now can transform the original image to the normal image, as shown in fig. 4. Processing the normal images has two advantages. First the clusters of targets are contained in squares and, secondly, the characteristics used for the recognition of the targets do not have to be invariant against a projective transformation.

The tracing and recognition of the targets is done in binary images. To increase the amount of information in the binary images, the normal image is enlarged by a factor two (fig. 5a) before the thresholding is applied (fig. 5b). In the binary enlarged image the contours are found using a contour following algorithm (fig. 5c). These contours and the target sizes (the number of pixels within the contour) are used for the recognition.

3.3. Target Recognition

The recognition is done separately for each cluster of sixteen targets. First these targets are divided in eight large and eight small ones. As the sizes of the squares and the triangles were a little smaller than those of the crosses and the circles, we could also divide the two groups of eight targets in four groups of four, using the size as criterion (table 3.).

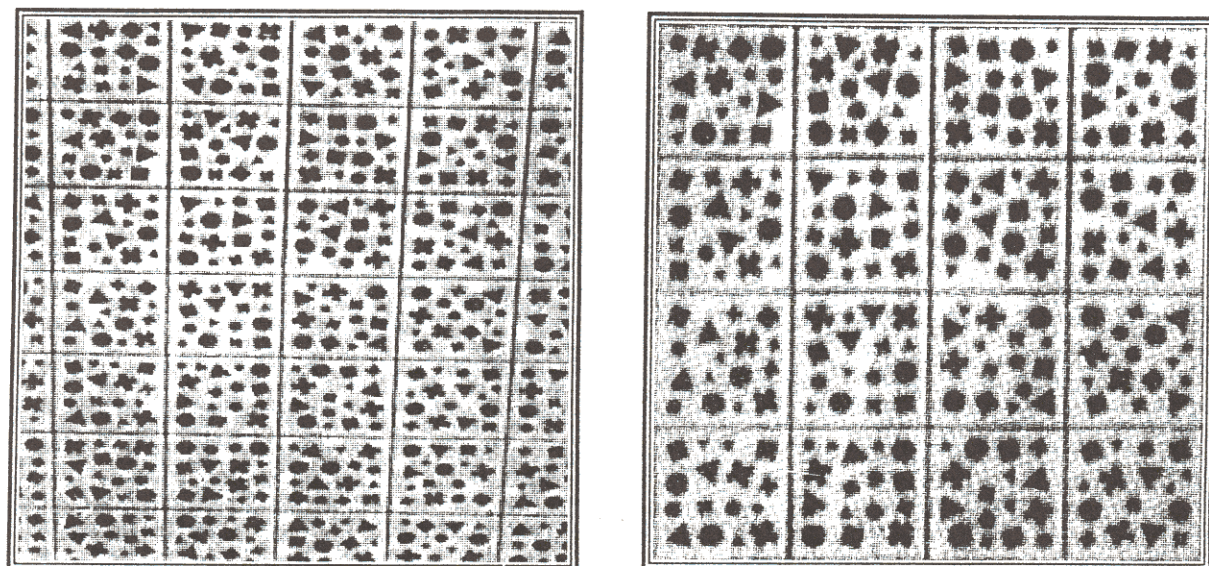


Fig. 4. The original and the normal image

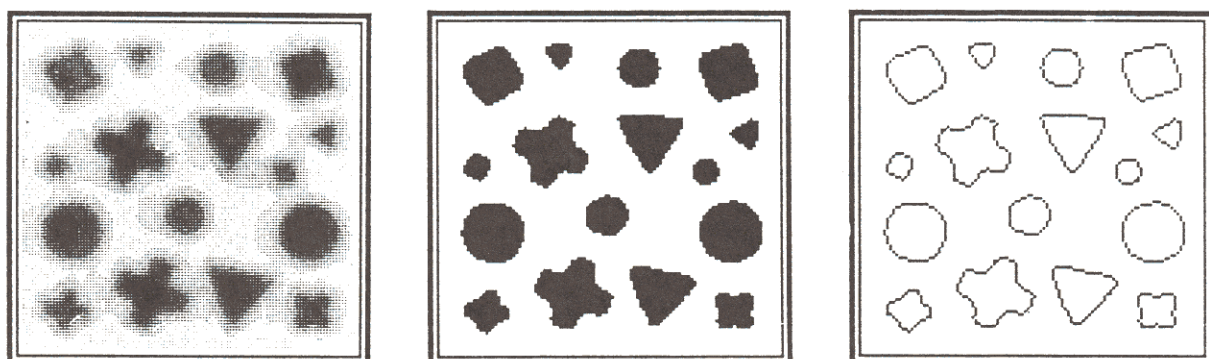


Fig. 5. The enlarged grey value, binary and contour image of one target cluster

size group 1	2 small squares and 2 small triangles
size group 2	2 small crosses and 2 small circles
size group 3	2 large squares and 2 large triangles
size group 4	2 large crosses and 2 large circles

Table 3. Division of the targets in size groups

For each size group, the shapes of the targets were determined by an analysis of the contours. Counting the frequencies of features like "sharp angle" and "concave angle", two likelihoods were derived for each target: one for the target being of type one (e.g. a small square) and one for the target being of type two (e.g. a small triangle). The relations between the frequencies of the features and the likelihoods were taken from a look-up table, which has been made manually after examining a few test images.

Having four targets within one size group, there remain six possible combinations of classifications (knowing that there are two targets of type

one and two of type two). The likelihoods for these combinations of classifications were obtained by multiplication of the likelihoods of the individual targets. The combination with the largest likelihood was assumed to be the correct classification.

With this classification strategy all (4608) large targets were classified right. Only 10 of the 2304 small crosses and circles (0.4%) were classified wrong. The classification of the small squares and triangles gave some more problems. Here 256 of the 2304 targets (11%) were classified wrong. The percentage of failures was heavily depending on the phi-rotation of the testfield. For the three rotations (0, 30 and 45 degrees), the failure rates were resp. 3%, 7% and 17%. Apparently, the small squares and triangles in the images with large phi-rotations became too small to be recognized reliably.

After the classification the targets were numbered using topological relations. The cluster numbers were determined by combinations of intersections between imaginary lines through pairs of targets. (e.g. the line through the large circles does intersect the line through the large squares but not the line through the large triangles). The numbering of targets of identical size and shape was done in a similar way.

3.4. Target Location

The precise coordinates of the targets were determined by a least squares matching of the targets with artificial target masks (Pertl, 1984). This algorithm assumes that the recorded target can be obtained by an affine transformation of the mask matrix and tries to estimate the transformation parameters by minimizing the square sum of the differences between the grey values of the target in the original image and the grey values of the simulated mask matrix. Approximate values for the parameters were calculated from the transformation parameters of the rectification and the centers of gravity of the targets in the binary enlarged normal images. The matching showed very good convergence: only two or three iterations were needed to get the additional corrections to the target coordinates below 0.01 pixel.

4. Self Calibrating Bundle Adjustment

The precision of the obtained target coordinates has been determined by bundle block adjustments using self calibration parameters to model several systematic effects. Unless stated otherwise, the adjustments have been performed with 32 of the 36 images. The four images with a phi-rotation of 0 degrees caused convergence problems and have been left out in the further analysis. All adjustments have been performed with minimal constraints: two horizontal and three vertical control points.

4.1. Testfield Bend

After a first adjustment with one common set of twelve orthogonal additional parameters (Ebner, 1976), the residuals still showed a systematic pattern. The orientation of this pattern was depending on the kappa-rotation of the camera and the residuals in the images of the testfield with a phi-rotation of 45 degrees were larger than those in the images of the testfield with a phi-rotation of 30 degrees. This pattern behaviour could be explained by a bend of the testfield. To check this, an adjustment was performed using all the images in the XOZ-plane of fig. 3. These images have about the same kind of testfield deformation. Only the large circles were used in this adjustment. The calculated object point heights indeed showed a systematic pattern (fig. 6.). The testfield must have had a bend of about 0.8 mm.

To eliminate the effects of the bend, the images were divided into eight groups of four images. All images in one group had the same phi-rotation of the testfield and the same kappa-rotation of the camera, i.e. the same testfield bend and thus the same deformation in the camera coordinate system. For each group the twelve parameter set was added. The determination of the

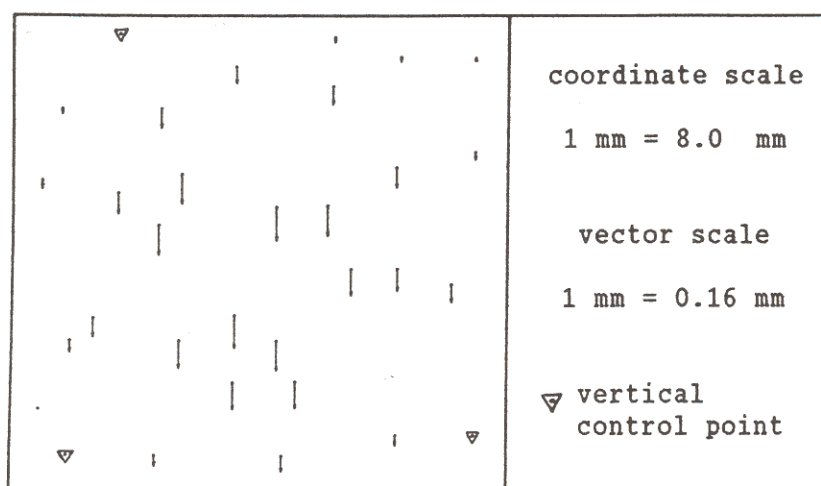


Fig. 6. Object point heights

parameters was done in an adjustment only using the large circles (32 per image). Despite of the many calibration parameters and the usage of only one eighth of the targets, the redundancy is still very large: 1671 of the 2055 observations are redundant.

4.2. Estimated Precision of the Targets from Block Adjustments

All target coordinates were corrected for the systematic effects with the parameters as determined above. For each of the target types, a separate bundle adjustment has been performed. The sigma naughts from these adjustments, representing the standard deviations of the image coordinates, are given in table 4. The large targets are a little preciser than the small ones.

	precision in μm		precision in pixel	
size shape	small	large	small	large
squares	1.09	0.92	0.040	0.034
triangles	1.51	1.34	0.056	0.050
crosses	1.20	0.82	0.044	0.030
circles	1.12	0.90	0.041	0.033

Table 4. Precisions of bundle blocks per target type

Remarkable is the relatively bad precision of the triangles. A possible explanation for the lower accuracy of the positions of the triangles is their type of symmetry: it is different from that of the grid (60° vs. 45° between neighbouring lines of symmetry). This difference may cause small systematic effects depending on the orientation of the triangle with respect to the grid, and, because of the randomness of their orientation is burried in the larger empirical standard deviation.

4.3. Systematic Effects in the Residuals

Fig. 7a shows the residuals of the large squares, crosses and circles in image 1211 ($\phi=30^\circ$, $\kappa_t=0^\circ$ and $\kappa_c=0^\circ$). All systematic parts seem to be eliminated. The root mean square sum of the residuals of the row coordinates is about 30% larger than that of the residuals of the column coordinates. This effect has been observed before (Gülch, 1985). It may be caused

by the scanning and sampling method of the camera.

Although the residual pattern seems to be random, there still is some systematic part in it. If one looks at the residuals of image 1311 (same kappa-rotations, but now a testfield phi-rotation of 45° , fig. 7b), they seem to be random too, but, comparing both patterns, the residuals of corresponding targets clearly show large resemblances in both direction and size. These similar residual patterns appear for all pairs of images, which have the same camera rotation and the same kappa-rotation of the testfield. The reasons for this phenomenon have not been revealed completely and will not be pursued in this paper.

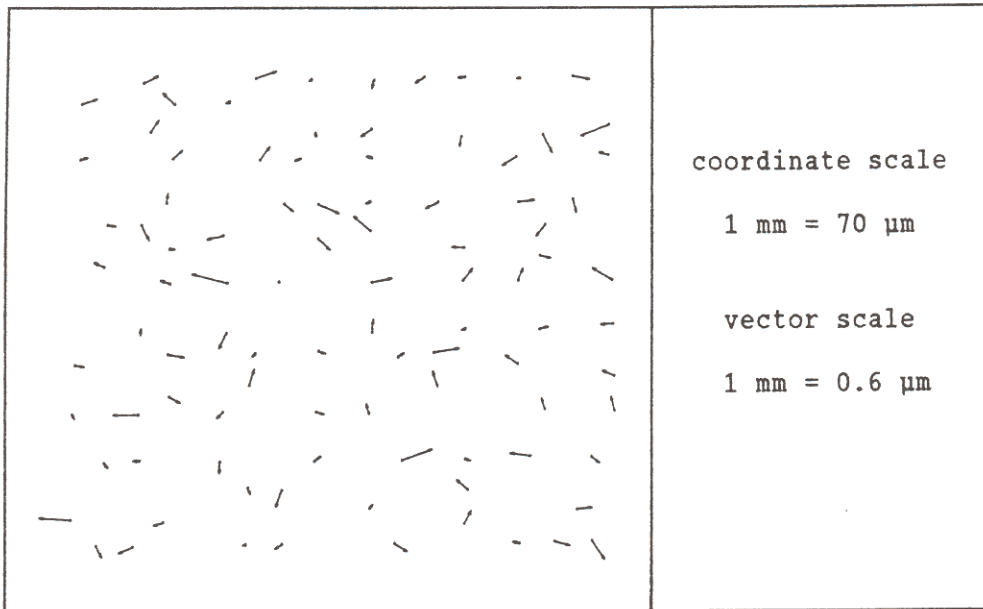


Fig. 7a Residuals of the large squares, crosses and circles in image 1211

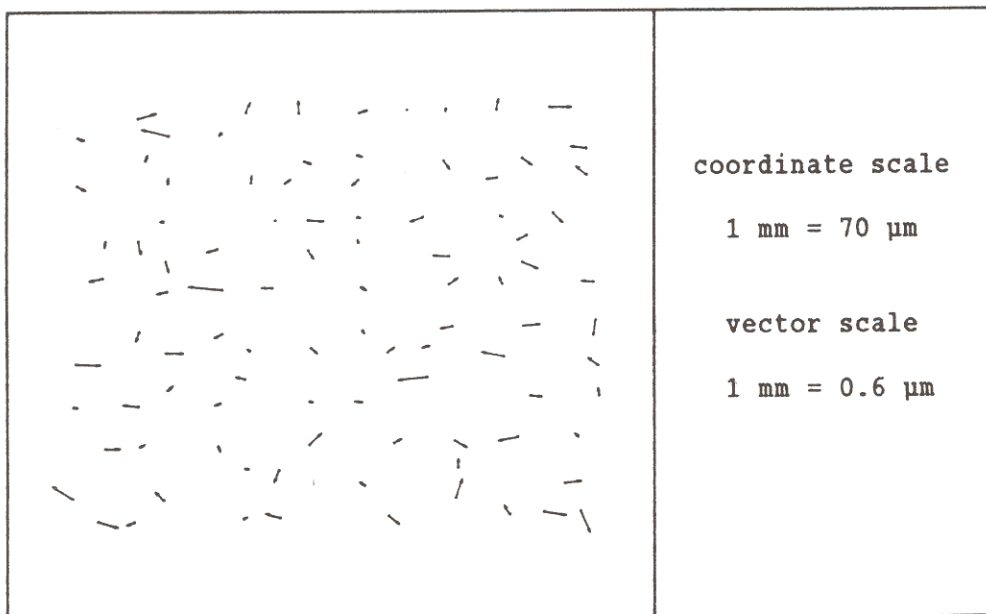


Fig. 7b Residuals of the large squares, crosses and circles in image 1311

4.4. Estimated Precision from Relative Orientations

To determine the impact of these effects, relative orientations have been performed using all targets (240 - 256 per image). The coordinates of these targets were corrected in advance and no additional parameters were used. To stabilize the configuration, all points were added as vertical control point with height zero. The standard deviation of these control points was fixed to 10 mm, so that they did not have any influence on the precision. The sixteen relative orientations of the images with the same kappa rotations had an average sigma naught of 0.48 μm (table 5.). This precision corresponds to 0.018 pixel. About 35% of the observations was redundant.

# image pairs	16		
# observations	1152	- 1232 , average	1199
redundancy	384	- 452 , average	420
σ_0 [μm]	0.39	- 0.56 , average	0.48
σ_0 [pixel]	0.015	- 0.021, average	0.018

Table 5 Performed relative orientations

5. Conclusions

1. The main scope of this project was to estimate the accuracy of target location using a digital camera. The results are very clear: Under well controlled conditions the empirical precision of target location can reach 1/30 pixel for targets with a diameter of appr. 10 pixels. Smaller targets of appr. 6 pixels diameter can be located with a standard deviation of appr. 1/20 pixel. The analysis also showed that there are still local systematic effects. Compensating for these could further reduce the standard deviations to appr. 1/50 pixel. Thus the geometric precision of the used CCD-camera of Hamamatsu is very high.

Due to the extremely high redundancy these results are quite reliable. They were achieved by applying a rigorous least squares matching (LSM) algorithm. However, no radiometric calibration of the camera was performed; only the local variations of brightness and contrast were taken into account in the LSM algorithm.

2. In the course of this study we obtained further results, which are worth mentioning:

- We also determined the centres of the targets solely based on the centres of gravity in the binary image (cf. sect. 3.2) The estimated standard deviations turned out to be worse about a factor 1.5 compared to the LSM approach. Thus the LSM algorithm proved to be really effective. For certain applications, however, the analysis of binary images may be sufficient and under controlled conditions may lead to accuracies of 0.1 pixels.
- The empirical results coincide with the internal standard deviation of the LSM algorithm within a factor 1.5 to 2 in both directions. For the small targets the internal estimates were too pessimistic, for the large targets they were realistic or a bit too optimistic. The main reason for these discrepancies seems to be the different sharpness of the targets in the image and in the used model. If one would introduce a sharpness parameter in to the LSM algorithm, as proposed by Thurgood and Mikhail (1982), the internal estimates could be expected to be too optimistic in all cases, because they only reflect the matching precision and not that of the camera geometry.
- The ratio between the standard deviations for the small and the large targets is very close to the theoretical expectation ((3/2 vs. $\sqrt{2}$).

3. The testfield in its present form, of course, is not suited for a full calibration of a digital camera. Several extensions are necessary:

- The test field has to be extended to 3D, e. g. by controlled movement in the z-direction and the coordinates of the targets have to be determined with a sufficient accuracy.
- In order to come to a fully automatic geometric calibration procedure one needs an automatic recording of the images, which may then be accomplished in much less than ½ hour. The automatic evaluation of the data up to now requires appr. ½ hour per image and some hours for the block adjustment (depending on the number of images and additional parameters) on a Harris H100. With special hardware the total time for the analysis could be reduced to 1 to 4 hours appr., which seems to be reasonable for practical applications.
- On the other hand one needs an analysis of the sensitivity of the sensor elements and of the resolution, in order to be able to compensate for possible defects and to be able to predict the performance of image interpretation tasks.

This study has to be complemented by others using different cameras, hopefully confirming the promising results obtained so far.

Literature:

IAP International Archives of Photogrammetry and Remote Sensing

- Ebner H. (1976): Selfcalibrating Block Adjustment, IAP Vol. 21-III, Helsinki, 1976
- Förstner W. (1982): On the geometric Precision of Digital Correlation, IAP Vol. 24-III, pp. 176-189, Helsinki, 1982
- Förstner W. (1984): Quality Assessment of Object Location and Point Transfer Using Digital Image Correlation Techniques, IAP Vol. 25-A3, Rio de Janeiro, 1984
- Grün A. W. (1987): Towards Real Time Photogrammetry, paper presented at the 41st Photogrammetric Week, Stuttgart, 1987
- Gülch E. (1984): Geometric Calibration of two CCD-Cameras used for Image Correlation on the Planicomp C100, IAP Vol. 25-B3, Rio de Janeiro, 1984
- Gülch E. (1985): Instrumental Realization and Calibration of Digital Correlation with the Planicomp, Proceedings 40th Photogrammetric Week, Stuttgart, 1985
- Pertl A. (1984): Digital Image Correlation with the Analytical Plotter Planicomp C100, IAP Vol. 25-B3, Rio de Janeiro, 1984
- Thurgood J. D. and Mikhail E. M. (1982): Subpixel mensuration of Photogrammetric Targets in Digital Images, Techn report CE-PH-82-2, Purdue University, West Lafayette, 1982
- Vosselman G. (1986) An Investigation into the Precision of a Digital Camera, Eng. Thesis, TU Delft, Dept. of Geodesy, Delft, 1986