Image sequence based automatic multi-camera system calibration techniques

Hans-Gerd Maas
Delft University of Technology
Faculty of Civil Engineering and Geo Sciences
Thijsseweg 11, 2629JA Delft, The Netherlands

Commission V, IWG V/III

Keywords: Image sequences, orientation, calibration

Abstract:

A fully automatic and rather flexible procedure for the calibration of stationary multi-camera systems for the 3-D observation of dynamic events is presented. While conventional close-range camera calibration techniques are either based on a stable pointfield with known reference coordinates or on a temporarily stationary pointfield with only approximately known 3-D coordinates, which is imaged from different locations and under different orientations with one single camera, the presented technique is based on stationary cameras and moving targets, making use of the image sequence acquisition nature of most solid state cameras. In the simplest version, only one single easily detectable marker has to be tracked through image sequences of multiple pre-calibrated cameras, thus avoiding the necessity of homologous feature identification for the establishment of multi-view correspondences; 3-D coordinates of the marker position are not required. This single-marker method does not allow for the determination of the interior orientation. In an extended version allowing for full camera orientation and calibration, a reference bar of known length is moved through object space, with the problem of feature identification and establishment of multi-view correspondences being reduced to the tracking of two targets.

The method can only be used with multi-camera systems and is most useful for 3-D motion analysis applications, but may be adapted to a wide range of other applications. The advantages of the method over conventional self calibration techniques are the trivial establishment of multi-view correspondences, the fact that no temporarily stable target field has to be constructed, and the fact that each camera has to be set up only once. After the explanation of the technique, its theoretical performance is examined in detail based on extensive computer simulations, and the practical effectiveness is shown in a pilot study on industrial robot calibration. Based on these studies, recommendations are given concerning the number of reference bar locations, preferable reference bar orientation schemes and the achievable accuracy potential.

1. Introduction

Close range photogrammetry is mostly based on the use of non-metric solid state sensor cameras today. To translate the high accuracy potential of such cameras, which is often in the order of $^1/_{50}$ of a pixel, into object space, these cameras have to be modeled and calibrated thoroughly. Camera models used in digital close range photogrammetry are usually based on the collinearity condition and a number of additional parameters modeling the interior orientation, lens distortion and sometimes further distortions like effects of the image A/D conversion or sensor unflatness. Camera calibration is advantageously performed 'on the job' by self-calibration techniques to represent the instantaneous state of the camera. For highest flexibility and minimum effort, calibration techniques should be based on a minimum of object space information. Photogrammetric self-calibration techniques, which are only based on beam intersections at a number of object points with unknown 3-D coordinates plus one scale information, have been introduced by (Brown, 1971) and (Kenefick et al., 1972). In order to fully reconstruct the interior orientation of one single camera to be calibrated, a temporarily stationary target field has to be imaged by this camera under three different orientations, one of them preferably rotated by 180^o to reduce high correlations between parameters of interior orientation, exterior orientation and 3-D object point coordinates. The simultaneous determination of a horizontal scale factor, necessitated by different clock rates of cameras and framegrabber in systems based on standard videonorm CCD cameras (El-Hakim 1986, Beyer 1987), requires the acquisition of a fourth image, preferably with another camera rotation by 90^o. A practical scheme for self-calibration network geometry, based on the acquisition of seven images, has been presented by (Godding, 1993). The positive effect of additional exposures under the application of rotation strategies has also been shown in (Grün/Beyer, 1992). Due to their high degree of flexibility, these self-calibration techniques have become a standard for the calibration of single-camera systems, which are based on the acquisition of multiple images of a scene for 3-D object reconstruction anyway.

The necessity of the establishment of a temporarily stable target field is part of the actual task in most stationary applications of digital close range photogrammetry, but may be cumbersome in dynamic applications like e.g. human motion analysis (Figure 1). Moreover, photogrammetric systems for 3-D data acquisition in dynamic processes will usually consist of multiple cameras imaging processes simultaneously, with each camera to be calibrated individually. The necessity of the acquisition of multiple images under several camera rotation angles with each individual camera of a multi-camera system causes a considerable effort and does generally reduce the flexibility. Therefore, there is a need for calibration techniques which are adapted to the requirements posed by photogrammetric systems for 3-D measurements in dynamic processes. In the following, the term 'calibration of a multi-camera system' is understood as the orientation and calibration of each individual camera of the system, considering the whole system as a tool for automatic 3-D information collection.

Figure 1: Multi-camera motion analysis system (Motion Analysis Corp.)

The image sequence acquisition nature of CCD cameras does allow to design calibration strategies, which are more appropriate to the calibration of multi-camera systems than conventional self-calibration techniques. In close range applications the datum is often defined as a local coordinate system defined by fixing the minimum of seven datum parameters. The task of relative orientation can then be reduced to the establishment of correspondences at a number of image points, exploiting the redundancy of stereo imaging. The fact that CCD cameras used for the monitoring of dynamic processes do produce image sequences suggests to establish these correspondences sequentially. In many cases this procedure allows to reduce the complexity of the task of relative orientation from the establishment of multi-image correspondences to the detection or tracking a single point in image sequences of multiple synchronized cameras, which can often be performed in realtime. In practice this can be realized by area or feature based tracking in image sequences; as an alternative, active techniques like a projected steerable laser beam, sequentially switched LEDs or a retroreflective marker moved through object space may be employed.

An obvious advantage of these 'moved single-point' methods is their simplicity, as tracking in image sequences with sufficiently high temporal resolution will often be much easier than the establishment of multi-view correspondences; the use of active techniques may even generate highly structured images, in which only one moved feature becomes visible, thus making the determination of homologous points trivial. A disadvantage, however, is the limited suitability of the method for self-calibration: While lens distortion parameters can be determined, the interior orientation of the cameras cannot be reconstructed from the information of a single moved point.

A straightforward extension of the 'moved single-point' calibration is the 'moved reference bar' method, which is based on moving a reference bar of known length through object space, which is imaged by all cameras at a number of locations/orientations over the observation volume. This reference bar can easily be moved through the observation volume by an operator. Compliance with exact locations and orientations is not required, and the calibration data acquisition can be finished within a few seconds. Similar to the 'moved single point' method, the establishment of correspondences is reduced to the detection or tracking of two features or targets (Figure 2). The constant length of the reference bar can be used as additional geodetic information in the self-calibrating bundle adjustment, thus strengthening the solution considerably and allowing for the full calibration of every camera of a multi-camera system, including the parameters of interior orientation.

Figure 2: Moved reference bar

The following considerations will concentrate on the determinability of the interior orientation (camera constant cc, principle point coordinates xp,yp and horizontal scale factor sx) of multiple cameras. These parameters are contained in almost all parameter sets used in digital close range photogrammetry; the determination of lens distortion and further higher order distortion effects turned out to be much less critical and will not be further analyzed.

Several examples for the application of the 'moved reference bar' method can be found in the literature: (Heikkilä, 1990) shows via simulations, that the calibration of a four-camera system based only on intersections at 141 object points is not possible, while after the introduction of 38 distance observations the full interior orientation can be determined. In (Pettersen, 1992) the method is used with pre-calibrated cameras only for the determination of exterior orientation parameters, improving scale control over the network. (Maas, 1997a) shows the application of the technique in industrial robot calibration with a reference bar moved to 27 random positions for the calibration of a three-camera system. The incorporation of stationary distance observations into photogrammetric networks has also been analyzed in (Wester-Ebbinghaus, 1983). So far, there is a lack of detailed analysis of the method; calibration schemes including advice on the ideal number, location and orientation of sequential reference bar observations to be used for system calibration do not exist. Obviously, these parameters are crucial for the accuracy potential of the method. Therefore, the sensitivity of results to the number and distribution of reference bar positions and orientations will be examined in the following, based on extensive numerical simulations and a practical example.

2. Results of simulations

A versatile simulation environment was generated for a thorough examination of the performance of the method. The simulations are based on an observation volume of 3m x 2m x 2m, which is imaged by three standard CCD cameras with 2/3" sensors and 12mm lenses; the cameras are arranged convergent 4m from the observation volume on a base of two times 1.5m (Figure 3). A reference bar with two signalized targets was moved through the observation volume. The length of the reference bar was chosen 1.5m, the standard deviation of the length 0.01mm. The datum was defined as a free network with arbitrary origin and rotation; the scale was defined by the length of the reference bar. Image coordinates of the marker coordinates are assumed to be measured automatically at a precision of 0.25 µm (~ $^1/_{40}$ of a pixel), which can be considered a realistic assumption.

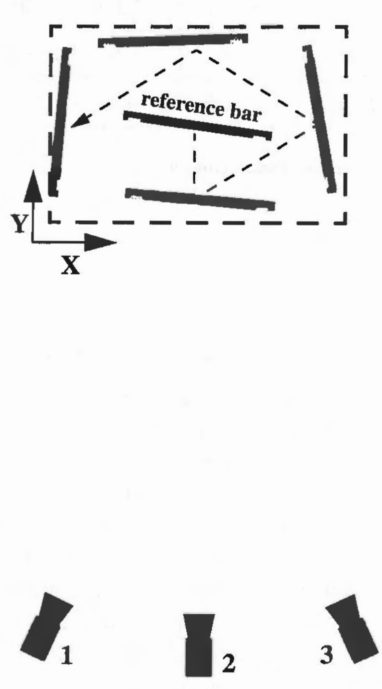

Figure 3: Configuration for simulations

The data was processed by self-calibrating bundle adjustment, using the length of the reference bar at every location as an additional observation. In the following, a 'reference bar observation' means the measurement of the image coordinates of the two markers on the reference bar in the images of all cameras taken synchronously at one reference bar location out of a longer sequence of reference bar locations, plus the known length of the reference bar.

The following simulations were conducted:

- Variation of the number of reference bar observations randomly located and oriented in object space.
- Regular arrangements of reference bar observations.
- Variation of the length of the reference bar.
- Variation of the number of cameras.

The following criteria were used to evaluate the results of the simulations:

- The standard deviations of the parameters of interior orientation and the horizontal scale factor.
- The standard deviations of the object space coordinates of the markers.
- A parameter indicating correlation effects: Self calibrating close range bundle adjustment does often create complex correlation patterns, which cannot easily be evaluated and are difficult to display. Therefore, this analysis is restricted to the obvious criteria of the standard deviation of the camera model parameters and the average standard deviations of object points; high correlations between parameters will usually also increase these standard deviations. In addition, a rudimentary measure for the correlation (c_{90}) is given, provided by the number of values larger than 0.90 in the inverse normal equation matrix.

In the following, the optimum number of reference bar observations and several location/orientation patterns will be evaluated based on the results of simulations.

2.1. Number of randomly distributed reference bar observations

A location/orientation pattern which can be established relatively easily is a random pattern generated by simply 'moving around' a reference bar through object space. The dependency of the standard deviations of the most critical parameters on the number of reference bar observations, based on a large number of simulations, is summarized in Table 1.

$\#_{ref}$	$\hat{\sigma}_{cc}$ [mm]	$\hat{\sigma}_{xp}$ [mm]	$\hat{\sigma}_{yp}$ [mm]	$\hat{\sigma}_{sx}$	$\hat{\sigma}_{k1}$	$\hat{\sigma}_X$ / $\hat{\sigma}_Y$ / $\hat{\sigma}_Z$ [mm]	c_{90}
4	-> singular matrix						
5	0.153	0.153	0.064	0.00414	0.00074	1.219 / 2.315 / 1.639	282
6	0.060	0.060	0.028	0.00140	0.00031	0.533 / 1.224 / 0.630	122
7	0.026	0.035	0.016	0.00051	0.00016	0.263 / 0.558 / 0.255	78
8	0.020	0.020	0.012	0.00048	0.00014	0.232 / 0.521 / 0.250	58
10	0.013	0.017	0.009	0.00022	0.00009	0.151 / 0.332 / 0.141	36
16	0.006	0.010	0.006	0.00014	0.00014	0.097 / 0.264 / 0.099	28
25	0.004	0.007	0.004	0.00011	0.00003	0.075 / 0.232 / 0.082	27
50	0.003	0.005	0.003	0.00006	0.00002	0.067 / 0.210 / 0.069	27
100	0.002	0.003	0.002	0.00004	0.00001	0.064 / 0.212 / 0.066	28
200	0.001	0.002	0.002	0.00003	0.00001	0.062 / 0.213 / 0.065	28
400	0.001	0.002	0.001	0.00002	0.00001	0.061 / 0.204 / 0.064	28

Table 1: Standard deviation of results versus number of reference bar observations

$\#_{ref}$ (in X,Y,Z)	$\hat{\sigma}_{cc}$ [mm]	$\hat{\sigma}_{xp}$ [mm]	$\hat{\sigma}_{yp}$ [mm]	$\hat{\sigma}_{sx}$	$\hat{\sigma}_X$ / $\hat{\sigma}_Y$ / $\hat{\sigma}_Z$ [mm]	c_{90}
4/4/4	0.005	0.007	0.004	0.00009	0.169 / 0.303 / 0.099	32
16/16/16	0.001	0.003	0.002	0.00004	0.099 / 0.237 / 0.068	68
25/0/0	0.086	0.005	0.010	0.00729	0.106 / 0.095 / 4.722	832
0/25/0	0.403	0.068	0.057	0.03523	24.18 / 0.685 / 37.49	2514
0/0/25	-> singular matrix					
25/25/0	0.061	0.005	0.007	0.00511	0.098 / 0.232 / 3.314	1348
25/0/25	0.015	0.006	0.002	0.00010	0.097 / 0.990 / 0.073	678
0/25/25	0.002	0.008	0.002	0.00167	1.540 / 0.238 / 0.080	1080
4/4/4 + 16 diag.	0.002	0.004	0.002	0.00005	0.090 / 0.222 / 0.071	26

Table 2: Results from regular arrangements

Obviously, the quality of the results increases with the number of reference bar observations. The minimum configuration with only five observations does not lead to acceptable results, while the standard deviations of all parameters are only dropping slowly beyond 50 reference bar observations. Balancing accuracy requirements and computational effort, an optimum number of reference bar observations will often be between 25 and 50.

2.2. Regular arrangements of reference bar observations

As an alternative to randomly distributed and oriented reference bar observations, a number of regular schemes were examined. These schemes were generated by arranging varying numbers of reference bar observations parallel to the object space coordinate axes, i.e. by a combination of horizontal, vertical and depth orientations. Table 2 shows the results of these configurations.

The results show that regular arrangements with reference bar orientations parallel to the object space coordinate axes do generally not yield better results than randomly distributed reference bar observations, and that arrangements with orientations parallel to only one or two coordinate axes should certainly be avoided. Solely vertical reference bar orientations in combination with horizonted cameras and the horizontal scale factor in image space do even lead to a singularity in the normal equation system; the latter problem can be solved by rotating one camera by a few degrees, but leads to a very weak solution then, too.

L_{ref} [mm]	$\hat{\sigma}_{cc}$ [mm]	$\hat{\sigma}_{xp}$ [mm]	$\hat{\sigma}_{yp}$ [mm]	$\hat{\sigma}_{sx}$	$\hat{\sigma}_X$ / $\hat{\sigma}_Y$ / $\hat{\sigma}_Z$ [mm]	c_{90}
1500	0.003	0.005	0.003	0.00006	0.067 / 0.210 / 0.069	27
750	0.004	0.006	0.003	0.00009	0.086 / 0.238 / 0.074	28
375	0.007	0.008	0.004	0.00014	0.114 / 0.273 / 0.097	32

Table 3: Dependency on reference bar length

$\#_{cam}$	$\hat{\sigma}_{cc}$ [mm]	$\hat{\sigma}_{xp}$ [mm]	$\hat{\sigma}_{yp}$ [mm]	$\hat{\sigma}_{sx}$	$\hat{\sigma}_X$ / $\hat{\sigma}_Y$ / $\hat{\sigma}_Z$ [mm]	c_{90}
2	0.003	0.005	0.003	0.00008	0.087 / 0.248 / 0.248	18
3	0.003	0.005	0.003	0.00006	0.067 / 0.210 / 0.069	27
5	0.002	0.004	0.002	0.00005	0.057 / 0.197 / 0.056	46

Table 4: Dependency on number of cameras

Only the combination of reference bar orientations parallel to all three object space coordinate axes with the 16 diagonals around and through the observation volume (last row of Table 2) does lead to slightly better results than the same number of randomly distributed reference bar orientations.

2.3. Length of the reference bar

Another parameter to be examined is the length of the reference bar. Table 3 shows the dependency of the standard deviations of the most important parameters on the length of the reference bar, based on 50 randomly distributed reference bar observations.

The analysis shows, that the dependency of the quality of results on the length of the reference bar is relatively uncritical, and that satisfactory results can even be achieved with a reference bar length of only $^1/_8$ of the largest observation volume dimension.

2.4. Number of cameras

In all above simulations the number of cameras was chosen to be three. This assumption is justified inasmuch as modern close range photogrammetric systems for 3-D data capture of moving objects are often based on three synchronized CCD cameras connected to one single RGB-framegrabber, thus allowing to base a system on off-the-shelf hardware components and keeping system cost low. Modern PCI-bus RGB-framegrabbers allow for the realtime transfer of image triplets of three synchronized CCD cameras to host memory at full spatial and temporal resolution (e.g. three simultaneous image sequences of 768x576 pixels at 25Hz in the European CCIR video norm). Nevertheless, the effect of the number of cameras on the results of the simulations is summarized in Table 4.

As expected, the standard deviation of object space coordinates as well as of camera parameters decreases with the number of cameras. However, for many applications this gain in precision will not be large enough to justify the significant extra technical effort of grabbing image sequences from more than three CCD cameras simultaneously. On the other hand, the transition from a two-camera configuration to a three-camera configuration will not require additional technical effort in many cases and will often be self-evident for reasons of reliability and robustness (Maas, 1997b).

3. Practical example

As a practical example, the technique was applied to the calibration of a three-camera system used in a pilot study on industrial robot calibration (Maas, 1997a). In this study, a robot was equipped with a signalizing plate with a number of targets, allowing for the determination of the pose of the end effector at a number of locations in order to improve the knowledge on robot model parameters. Prior to this, the robot was equipped with a 930mm reference bar with two 25mm targets (Figure 4), which was moved to 28 locations/orientations randomly distributed over the robots work range of approximately 1.7 x 1.5 x 1.0 m^3.

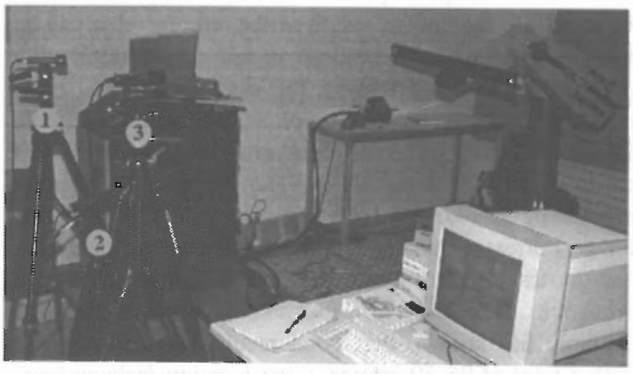

Figure 4: Moved reference bar in a pilot study on industrial robot calibration (Maas, 1997a)

Version	$\hat{\sigma}_{cc}$ [mm]	$\hat{\sigma}_{xp}$ [mm]	$\hat{\sigma}_{yp}$ [mm]	$\hat{\sigma}_{sx}$	$\hat{\sigma}_{k1}$	$\hat{\sigma}_X$ / $\hat{\sigma}_Y$ / $\hat{\sigma}_Z$ [mm]	c_{90}
A	0.38	0.007	0.015	0.008	0.00015	0.148 / 0.376 / 0.092	44
B	0.29	0.004	0.005	0.004	0.00009	0.088 / 0.240 / 0.057	59

Table 5: Standard deviations of parameters in practical test

The given robot model parameters allowed for the prediction of the marker coordinates in image space with an uncertainty of only about one to two pixel, thus making the tasks of marker detection and identification trivial. The results of the bundle adjustment are summarized in Table 5 as version 'A'. In version 'B', 800 positions of markers on the signalizing plate collected for the actual robot calibration (without distance information) were added to the adjustment.

The results are slightly worse than the results of the simulations, but they prove the good determinability of the parameters of the interior orientation. The degradation of the standard deviation of the X-coordinates has to be attributed to problems with line jitter of the analog XC77ce CCD cameras, which were probably caused by effects of electro-magnetic fields of the robot to the camera synchronization. An analysis of computation times (Maas, 1997a) shows that the measurement of the markers in image space can easily be performed in real-time, which is an essential pre-requisite for the efficient use of the method as no image sequences have to be stored.

4. Conclusion

The 'moved reference bar' method can be considered a versatile and reliable method for the calibration of photogrammetric systems consisting of multiple solid state cameras. It avoids the determination of 3-D reference coordinates associated with reference field techniques, and the necessities of establishing a temporarily stable point field and taking multiple exposures under different orientations with each camera to be calibrated, as required by conventional self-calibration techniques. Instead, multi-ocular image sequences of a reference bar have to be acquired; the known length of the reference bar can be used as additional observations in self-calibrating bundle adjustment, thus warranting the determinability of the parameters of interior orientation and the horizontal scale factor caused by different clock rates of an analog CCD camera and a framegrabber. The complexity of data processing is reduced from the establishment of multi-image correspondences at many points to the detection and/or tracking of two discrete features in image sequences. The analysis of simulations and a practical example has shown, that good results can be achieved with a total of 25-50 reference bar locations/orientations, which are preferably randomly distributed over the observation volume.

References:

- Brown, D., 1971: Close-Range Camera Calibration. Photogrammetric Engineering, Vol. 37, No. 8, pp. 855-866
- El-Hakim, S., 1986: A real-time system for object measurement with CCD cameras. IAPRS Vol. 26, Part V
- Godding, R., 1993: Ein photogrammetrisches System zur Überprüfung und Kalibrierung digitaler Bildaufnahmesysteme. ZPF 2/93
- Grün, A., Beyer, H., 1992: System calibration through self-calibration. Workshop 'Calibration and Orientation of Cameras in Computer Vision', Washington D.C., 2. Aug. '92
- Heikkilä, J., 1990: Update calibration of a photogrammetric station. IAPRS Vol. 28, Part 5/2, pp. 1234-1241
- Kenefick, J., Gyer, M., Harp, B., 1972: Analytical self-calibration. Photogrammetric Engineering, Vol. 38, No. 11, pp. 1117-1126
- Maas, H.-G., 1997a: Dynamic photogrammetric calibration of industrial robots. Videometrics V (Ed. S. el Hakim), SPIE Proceedings Series Vol. 3174
- Maas, H.-G., 1997b: Mehrbildtechniken in der digitalen Photogrammetrie. Habilitation thesis at ETH Zurich, Publications of the Institute of Geodesy and Photogrammetry, Vol. 62
- Pettersen, A., 1992: Metrology Norway System - an on-line industrial photogrammetric system. IAPRS, Vol. 29, Part B5, p. 43
- Wester-Ebbinghaus, W., 1983: Ein Beitrag zur Feldkalibrierung von Aufnahmekammern. DGK Reihe C, Nr. 289