External camera calibration for synchronized multi-video systems

Ivo Ihrke Lukas Ahrenberg Marcus Magnor

Max-Planck-Institut für Informatik
D-66123 Saarbrücken

ihrke@mpi-sb.mpg.de ahrenberg@mpi-sb.mpg.de magnor@mpi-sb.mpg.de

ABSTRACT

We present a camera calibration system that is simple to use and offers generality in the positioning of the cameras. This makes it very suitable for the calibration of mobile, synchronized camera setups. We use a *camera graph* to perform global registration which helps lifting restrictions on the camera setup imposed by other calibration methods. A further advantage is, that all information is taken into account at the same time. The method is based on a virtual calibration object which is constructed over time. This implies that no calibration object must be visible simultaneously in all cameras.

Keywords

External Camera Calibration, Registration, Multi-Video System, Graphs

1 Introduction

We have built a mobile camera system to take computer vision research to more general settings than a fixed studio setup. This added flexibility demands a simple to use camera calibration technique that allows for convenient calibration of the camera system.

Since camera calibration is such a crucial task in computer vision, a lot of effort has been spent on this subject. There exist a number of methods, e.g. [Tsai 1987; Hartley and Zisserman 2000] in different flavors. Most of them use carefully prepared calibration objects, while others estimate camera parameters from general images or image sequences. These approaches are called self-calibrating. A relatively new approach is the use of a virtual calibration object. This calibration object does not exist in a physical meaning, but is instead constructed over time (assuming a static camera setup and scene) by tracking an easily identifiable object through an image sequence. The point correspondences hereby obtained are then used for a traditional calibration step [Azarbayejani and Pentland 1995; Chen et al. 2000].

The advantages of using virtual calibration objects are the simple establishment of point correspondences in difficult wide baseline and encircling camera setups as well as the possibility of acquiring as many point measurements as necessary, which is difficult in one image approaches. Furthermore there is no need for a carefully manufactured calibration object, which can break during transportation.

The paper is organized as follows. In Section 2 we discuss some previous work that is important in this context. Section 3 gives an outline of the method, section 4 describes the establishment of 2D correspondences for calibration purposes. In section 5 the computation of pairwise relationships between cameras is discussed. Section 6 describes the main contribution of this paper, namely the usage of a graph structure for external camera calibration. Section 7 presents experiments with synthetic as well as with real data. Finally section 8 concludes the paper and gives some directions for future work.

2 Related Work

Earlier approaches to external calibration with help of virtual calibration objects [Azarbayejani and Pentland 1995; Chen et al. 2000] use light emitting objects (i.e. flashlight, LED) in dark rooms for tracking. In [Azarbayejani and Pentland 1995] the virtual calibration object is introduced to calibrate a pair of stereo cameras. A synchronized pair of cameras is used to track a flashlight, the path of which is used for calibration. Chen et. al. extend this work by applying the virtual calibration object approach to unsynchronized multi-camera setups. The lack of camera synchronization is treated with an Extended Kalman Filter (EKF) which estimates the path of a marker object (LED). Pairwise relationships between cameras are computed employing a structure from motion algorithm [Zhang 1996]. Global registration in a common coordinate system is then done using a triangulation scheme iteratively, to estimate the position of yet unregistered cameras to two already registered cameras. This requires that any unregistered camera is connected to at least two already registered cameras by known pairwise relationships.

Our approach removes this restriction by analyzing a graph consisting of the cameras as vertices and known pairwise position and orientation as the edges. An additional advantage is that all information about the global position of a given camera is imposed simultaneously.

3 Outline

To accommodate general camera setups, we use an approach that does not require a common view of a calibration object for all cameras. Instead a *virtual* calibration object that covers the working volume is constructed over time by tracking an easily identifiable object [Azarbayejani and Pentland 1995; Chen et al. 2000]. We mainly follow the approach of [Chen et al. 2000]. However, we lift the constraints that the cameras must be registered by triangulation from the base camera pair and that the working volume must be dark during calibration.

Intrinsic calibration is done independently for each camera using Tsai's method [Tsai 1987]. This is reasonable since our cameras do not allow for varying internal parameters. This requires us to record an image of a checkerboard for each camera, from which the internal parameters can be computed. However, the checkerboard does not have to be visible from all cameras simultaneously.

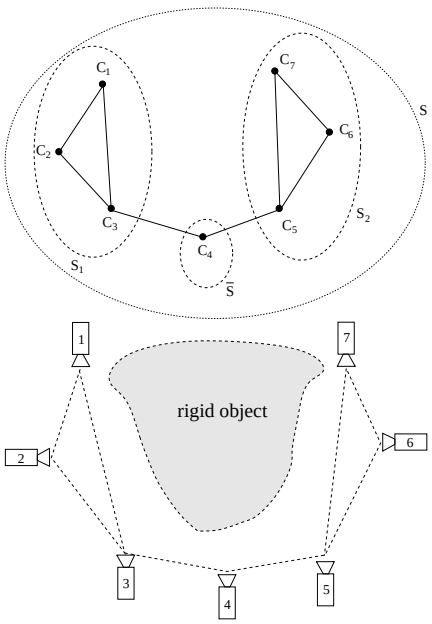

Figure 1: *An example for a graph with nonempty sets $S, S_k, \bar{S}$ (upper image) and a corresponding scene where this situation could occur (lower image).*

At the place of recording, we first obtain a number of 2D correspondences via tracking the marker object depicted in Fig.2. The midpoints of the marker object are determined and the correspondences are used to robustly compute the relative orientation of each camera pair [Zhang 1996; Horn 1990].

Since the pairwise relative position of the cameras can be found up to a scale factor only, a global registration has to be performed to achieve global calibration. This is done by constructing a graph $G = (V, E)$ that represents the relationship between cameras Fig. 1. The cameras $C_i \in V$ are the vertices, and known relative position and orientation between any pair of cameras $(C_i, C_j) \in E$ are represented by the edges of the graph. The graph is undirected, since known relative position and orientation $(\mathbf{R}, \mathbf{t})$ for camera pair (C_i, C_j) also defines the relative position and orientation of (C_j, C_i) as $(\mathbf{R^T}, -\mathbf{R^T t})$.

Once the graph is set up, it is searched for cyclical connected subsets $S_k \subseteq S$ containing all cameras. The set $\bar{S} := S \setminus (\cup_k S_k)$ consists of all cameras not belonging to any cycle. If $S = \bar{S}$, no cycles exist and a special treatment is necessary. Figure 1 illustrates a situation where all sets are nonempty.

The unknown scale factors for the pairwise translations are determined for each S_k independently. This is done by solving an over-determined

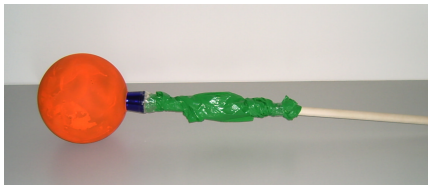

Figure 2: *The marker object that is tracked through the scene to establish 2D correspondences between camera frames.*

linear system of equations. The equations correspond to cycles in the graph and require that the scaled translations along the traversed path sum to zero. That means the cycle represents a closed curve in three-dimensional space. Since the estimation obtained this way is not connected to the image measurements any longer the reprojection error of this solution is usually not satisfactory. Nevertheless the overall shape of the setup is recovered quite well. Therefore this estimate can be used as an initial estimate for bundle adjustment [Triggs et al. 2000; Hartley and Zisserman 2000].

Using these partial registrations, parts of the virtual calibration object (i.e. the path of the tracked marker object) are reconstructed.

The remaining step is to register all subsets S_k and all $C_i \in \bar{S}$ with each other to form a globally consistent calibration.

In this framework the improvement of our method can be stated like this: The restriction of the vertices V being interconnected by three-cycles in the graph G as required by the algorithm of Chen et. al. [Chen et al. 2000] is lifted and arbitrary connectivity of the graph G is allowed for as long as the graph is not disconnected.

4 Obtaining image correspondences

2D point correspondences between images, i.e. projections of the same three-dimensional point onto different camera planes, are the basis for epipolar geometry estimation and therefore the first step in the recovery of the camera structure.

We obtain them by tracking a sphere over time, establishing one correspondence in every frame of the calibration video sequence. The sphere is a suitable object because it has a unique midpoint, the projections of which can be computed from images alone. The sphere projects to a conic section in the image plane 3 which in a real situation will always be an ellipse. If the focal length of the cameras is not too small the ellipse is nearly circular. This observation leads to our sphere detection algorithm.

The sphere detection is run on every frame of the video sequences separately. Our marker object has a color which is not widely present in the scene. Therefore we threshold the image with a color band in Y/Cb/Cr space. We find connected components in the resulting binary image and threshold them according to their size to exclude noise and small artifacts from further processing. A circular Hough transform is performed on every remaining connected component. This yields a best circular fit for each of them.

To decide the best fit for the image, the Hough scores can not be compared directly as they depend on the size of the object. Therefore we calculate the density of object pixels in the circle candidates. A candidate from a round component will typically fill the whole disk and gain a high score. The highest scoring circle is picked as the winner. Finally the score is thresholded. If the score is higher or equal to the threshold we have found the sphere, otherwise we assume that the it is not visible in the image.

If the sphere was found we refine the estimate using orthogonal least squares ellipse fitting. For this purpose we extract the edges of the original image around the position found by the Hough transform. The area for edge detection is limited by the estimated radius plus some safety region. The detected edge pixels are used to robustly fit an ellipse using the RANSAC paradigm [Fischler and Bolles 1981]. We use the squared sum of orthogonal euclidean distances between the data points and the ellipse as an error measure. This gives a sub-pixel estimate of the ellipse midpoint and compensates for small mislocalizations introduced by the Hough transform. We use those midpoints as correspondences.

5 Computation of pairwise position and orientation

Having extracted a set of point correspondences, we can proceed in computing the pairwise relative position and orientation $(\mathbf{R_{ij}}, \mathbf{t_{ij}})$ of any camera pair (C_i, C_j).

This is done by first undoing the effects of the internal camera parameters on the midpoints, followed by the computation of the essential matrix $\mathbf{E_{ij}}$. We use a robust variation of Zhang's approach [Zhang 1996] to essential matrix computation. The difference between Zhang's algorithm and our version is the initial guess of the

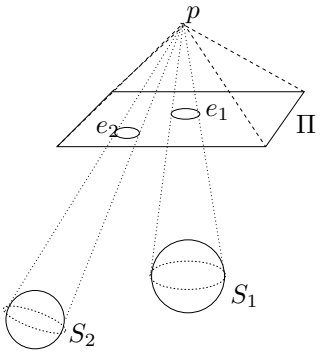

Figure 3: Projection of spheres S_1 and S_2 onto the image plane Π using the focal point p. S_1 is located on a ray through p and the center of Π, thus projection to the middle of the image plane. S_2, on the other hand, lies off from this ray at some angle, and projects to the outer regions of Π. e_1 and e_2 can be viewed of as the image formed by Π cutting the projection cone from the spheres to p. In the case of S_1 the projection, e_1, is a perfect circle as Π is parallel to the base of the projection cone. As for S_2, it deviates from the center of projection, and its projection cone is cut by Π at some angle forming an ellipsoid.

fundamental matrix $\mathbf{F_{ij}}$, which is obtained using RANSAC [Fischler and Bolles 1981] with the seven point algorithm for estimation of the fundamental matrix [Hartley and Zisserman 2000]. This estimate fulfills already the rank two constraint on the fundamental matrix $\mathbf{F_{ij}}$. The following nonlinear refinement step minimizes the symmetric epipolar distance, i.e. the distance of the epipolar line induced by a point to the corresponding point in the other image, using the Nelder-Mead simplex method [Nelder and Mead 1965]. Given the internal parameter matrices $\mathbf{K_i}$ and $\mathbf{K_j}$ an initial guess of the essential matrix is computed $\mathbf{E_{ij}} = \mathbf{K_j}^T \mathbf{F_{ij}} \mathbf{K_i}$. This guess usually does not fulfill the additional constraint that the two nonzero singular values of $\mathbf{E_{ij}}$ are equal. The nearest essential matrix fulfilling the constraints is obtained by setting the two unequal singular values to their mean value [Faugeras and Luong 2001]. However if the two nonzero singular values differ widely, this is not a good guess for the essential matrix minimizing the symmetric epipolar distance. Therefore an additional nonlinear minimization with the simplex-algorithm is performed, this time using only the five parameters of rotation and translation direction to compute the final guess for matrix $\mathbf{E_{ij}}$.

The essential matrix is of the form $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$ and rotation as well as translation can be extracted from it [Horn 1990]. The decomposition of matrix $\mathbf{E}$ yields four solutions for $\mathbf{R}$ and $\mathbf{t}$. The solutions arise because of the unknown scale of $\mathbf{t}$ which can be either positive or negative. The rotations are related by a rotation of 180° around the baseline, connecting the two cameras. In most cases this ambiguity can be resolved by the demand that reconstructed points are in front of both cameras [Hartley and Zisserman 2000; Faugeras and Luong 2001]. Sometimes, in cases of a near 180° angle between the principal axes of the two cameras, it is only possible to resolve a twofold ambiguity, two solutions remain. Since the reprojection error of both solutions is the same there is no measure based on image distances that can be chosen to decide the solution. Therefore we apply a heuristical measure: we choose the solution whose reconstruction has a smaller convex hull. Since we are dealing with euclidean reconstructions this is a reasonable choice.

As a measure of quality we compute a local reconstruction for each camera pair (C_i, C_j) that yielded a solution $(\mathbf{R_{ij}}, \mathbf{t_{ij}})$ and evaluate the reprojection error caused by that reconstruction. We apply a threshold on that error to exclude unstable estimates.

6 Graph building and analysis

From the previously computed pairwise positions and orientations we build a graph that mirrors the availability of rotation/translation information for all camera pairs. The vertices of the graph correspond to the cameras, and edges connecting them indicate a stable solution to the relative position and orientation problem for the cameras connected by that edge. Looking at the problem from a graph theoretical point of view has several advantages. There are standard solutions for problems occurring in general camera setups. It is for example simple to check if a calibration containing all cameras is possible. It is sufficient to check, that the graph is connected. This can already be done during data collection and the user can be guided to create more correspondences for camera pairs that have not sufficient data available yet.

If there are unconnected subsets of vertices in the graph no global calibration can be found, but the subsets of cameras can be calibrated separately. Cameras corresponding to isolated vertices cannot be calibrated.

6.1 Registration of cyclic connected components

We proceed by registering all cameras C_i^k contained in a cyclic connected component G_k of the graph G with each other. This is done for all k separately. For improved readability we skip the subscript k in the further discussion.

Recall that all relative translations $\mathbf{t_{ij}}$ are defined only up to a scale factor and reside in the local coordinate system of camera C_i. The task of the registration procedure is to find consistent scale factors s_{ij} for the $\mathbf{t_{ij}}$ and their transformations to a global coordinate system shared by all cameras.

The registration is based on cycles in the graph $G = (V, E)$. We denote a cycle as $Z = [z_1 \dots z_n], z_i \in \{C_1 \dots C_m\}$ and introduce $\tilde{\mathbf{t}}_{\mathbf{ij}}$ as translation direction $\mathbf{t_{ij}}$ transformed to a common coordinate system. Then a condition on the scale factors s_{ij} for one cycle can be written as:

$$\left(\sum_{i=1}^{n-1} s_{z_i,z_{i+1}} \tilde{\mathbf{t}}_{z_i,z_{i+1}} \right) - s_{z_n,z_1} \tilde{\mathbf{t}}_{z_n,z_1} = 0 \qquad (1)$$

Equation 1 is valid for all cycles $Z^h, h \in \{1 \dots l\}$. It describes the condition that every walk along a cycle in the registered set of cameras returns to its origin, or all translations along a cycle sum up to zero. Every cycle contributes three equations to the system of equations in equation 1, one for each coordinate. Overall there are $3l$ equations that make up a usually overdetermined linear system of equations $\tilde{\mathbf{T}}\mathbf{s} = \mathbf{0}$, which can be solved to obtain the scale factors s_{ij}. Matrix $\tilde{\mathbf{T}}$ and vector $\mathbf{s}$ contain the t_{ij} and s_{ij} in appropriate positions respectively.

To apply this for the registration, a number of preliminary steps are necessary which will be explained in more detail later:

1. find all cycles Z^h in graph $G = (V, E)$,
2. transform all $\mathbf{t_{ij}}$ into a common coordinate system, i.e. compute $\tilde{\mathbf{t}}_{\mathbf{ij}}$,
3. construct and solve the system of linear equations in equation 1,
4. using the computed scale factors s_{ij} compute a consistent calibration

1) An algorithm for computing all cycles in a graph is presented in [Kim et al. 2003].

2) We choose one of the cameras local coordinate systems as the common coordinate system. The corresponding camera is referred to as base camera C_b. We use the camera related to the vertex with the highest degree for that purpose. This choice minimizes the number of edges that have to be traversed to transform the other cameras into the base cameras coordinate system.

For a given camera C_i we transform the translation directions $\mathbf{t_{ij}}, (C_i, C_j) \in E$ originating in C_i into the common coordinate system, yielding $\tilde{\mathbf{t}}_{\mathbf{ij}}$. This is done using the quaternion weighted mean rotation $\overline{\mathbf{R}}_{ib}$ between C_i and C_b: Rotation $\overline{\mathbf{R}}_{ib}$ is computed by finding all paths between cameras C_i and C_b, accumulating rotation matrices along these paths and taking the weighted mean value of their quaternion representations using iterative spherical linear quaternion interpolation[1].

3) The solution $\mathbf{s}$ of the linear system $\tilde{\mathbf{T}}\mathbf{s} = \mathbf{0}$ is defined up to a scale factor only. Therefore matrix $\tilde{\mathbf{T}}$ is singular. This is because the overall scale factor of the camera setup cannot be determined from image measurements alone. To solve the linear system, we compute the basis vector of the right null-space of $\tilde{\mathbf{T}}$. Using the standard approach to this problem, we compute the singular value decomposition (SVD) of matrix $\tilde{\mathbf{T}}$: $\tilde{\mathbf{T}} = \mathbf{U}\mathbf{D}\mathbf{V}^\mathbf{T}$ and extract the column of $\mathbf{V}$ corresponding to the zero singular value.

4) To compute a consistent calibration we multiply the transformed translations $\mathbf{t_{ij}}$ by their corresponding scale factors s_{ij}. If the solution was perfect the calibration would be readily available since every path from C_i to C_b in the graph could be taken to transform camera C_i into the base coordinate system. But because $\mathbf{s}$ is a least squares solution this is generally not true. To get the best estimate of the true translations we compute the weighted mean translation $\overline{\mathbf{t}}_{\mathbf{ib}}$ similarly to the computation of $\overline{\mathbf{R}}_{\mathbf{ib}}$ in step 2). The calibrated cameras $\overline{C}_i$ are now defined as having the projection matrices

$$\mathbf{P_i} = \mathbf{K_i}\overline{\mathbf{R}}_{\mathbf{ib}} \begin{pmatrix} \mathbf{1} & | \overline{\mathbf{t}}_{\mathbf{ib}} \end{pmatrix} \qquad (2)$$

with $\overline{\mathbf{R}}_{\mathbf{bb}}$ being the unit matrix and $\overline{\mathbf{t}}_{\mathbf{bb}}$ the zero translation for the base camera C_b.

6.2 Bundle adjustment

The calibration obtained so far captures the coarse structure of the camera setup quite well, but since it minimizes an algebraic error, namely the least squares error which has no physical meaning, the

[1] The weighted mean value $\overline{x} = \frac{1}{\sum_{i=1}^{n} w_i} \sum_{i=1}^{n} w_i x_i$ can be computed in an iterative manner where a linear interpolation is performed in every iteration. Replacing the x_i by quaternions and using spherical linear interpolation a weighted mean value of rotations can be computed.

standard deviation of uniform Gaussian noise	mean reprojection error left image	mean reprojection error right image
0.1	0.0213	0.0200
0.3	0.0710	0.0677
0.5	0.1174	0.1131
0.7	0.1600	0.1512
0.9	0.2045	0.2151
1.1	0.2548	0.2425
1.3	0.2815	0.2718
1.5	0.3492	0.3262
1.7	0.3612	0.3452
1.9	0.4503	0.4263

Table 1: Mean reprojection error in the left and right images for pairwise relationship computation (inliers only) for different noise levels (experiment 1).

standard deviation of uniform Gaussian noise	mean reprojection error over all camera pairs	mean reprojection error after bundle adjustment
0.1	0.1474	0.1003
0.3	0.7382	0.3285
0.5	1.4748	0.5481
0.7	1.2274	0.7602
0.9	1.5640	1.0145
1.1	1.8092	1.2005
1.3	2.0367	1.4523
1.5	2.7336	1.6161
1.7	8.6253	2.1588
1.9	6.0498	2.1608

Table 2: Mean reprojection error of the reconstruction from the linear solution to the calibration dependent on the noise level and after bundle adjustment (experiment 2).

reprojection error of the points reconstructed with this calibration is quite high. Nevertheless it serves as a good initial guess for bundle adjustment [Triggs et al. 2000]. We implemented a reduced version which optimizes only rotation and translation parameters of the final calibration using a nonlinear optimization method. Usually all parameters (internal ,external and reconstructed point positions) are optimized which is a huge problem with hundreds of variables and can only be handled by exploiting the coarse structure of the problem [Pollefeys 1999].

We parametrize the problem in a similar way as in the essential matrix refinement step, but this time for all cameras simultaneously. The rotation is parametrized as the normalized rotation axis scaled by the rotation angle. These values are obtained from the quaternion representation of the rotation. The translation values are used directly for the parametrization. As the error measure that is to be minimized, we adopt the sum of the mean reprojection errors of the reconstructed points in all cameras. This error measure is optimized using again the simplex algorithm and the previous calibration as an initial guess.

7 Experiments

We performed three experiments to validate our calibration method. The first two experiments use synthetic data with different levels of uniform Gaussian noise to assess the robustness of our method to noise. The first experiment shows the accuracy of the pairwise position and orientation computation (see section 5) under noise. The second experiment evaluates the same for the initial calibration obtained by solving the linear system of scale factors, section 6.1 and the improvements of bundle adjustment. The third experiment uses real data and is the calibration of our in-house multi-video studio [Theobalt et al. 2003], which is obtained using the full method. Errors are mean reprojection errors in pixels where not indicated otherwise.

7.1 Performance of the pairwise relationship estimation

For this experiment we use 100 synthetic 3D data points. The data points are projected onto 5 cameras and disturbed by uniform Gaussian noise with standard deviations ranging from 0.1 to 1.9 pixels. The dependency of the pairwise relationship estimation on the noise level is shown in table 1. The threshold for the RANSAC method was set to 1.0 pixels. The results show, that the method is robust against uniform Gaussian noise and the reprojection error raises slower than the noise level. Nevertheless with higher noise levels fewer points are detected as inliers and the computation time, which raises when the ratio of inliers to total number of points gets smaller, becomes longer.

7.2 Performance of the linear solution to the calibration

The second experiment was performed to test the performance of the linear calibration method of

camera number	median reprojection error
1	1.8665
2	2.8460
3	1.6047
4	2.3499
5	1.7885
6	2.7048
7	5.1755
8	2.4068

Table 3: Median reprojection error of the reconstruction from the linear solution to the calibration dependent on the noise level and after bundle adjustment (experiment 3).

section 6.1. The estimated pairwise positions and orientations from the previous experiment were used to perform the linear calibration. After calibrating the synthetic cameras we reconstructed the 3D points from the noisy image points and reprojected them onto the camera planes. The mean euclidean distance between the image points and their reprojections and the improvements achieved by using bundle adjustment are shown in table 2. It can be noted that the linear method produces high errors quite fast. Nevertheless it gets near the desired minimum of the nonlinear cost function quite well and the bundle adjustment can recover the correct calibration up to the noise level.

7.3 Performance on real data

The third experiment uses real data extracted from a video sequence of 480 frames, recorded at 15 fps. Our sphere detection algorithm is applied to extract the midpoints of the sphere. These midpoints are then used to perform the full calibration. For reconstruction we use all cameras that observed a point. A visualization of the reconstructed virtual calibration object, i.e. the path of the marker object and the reconstructed camera setup are shown in Fig. 4. For the evaluation we use the median reprojection error this time. This is because the data includes outlier which disturb the mean error computation. The results can be seen in table 3. Except for camera 7 which has a very high error all cameras are calibrated reasonably well. The remaining error is most likely due to errors in the internal calibration and due to the fact that we use the midpoints of the detected ellipses as correspondences which do not generally correspond to the projected midpoint of the sphere [Heikkila and Silven 1997]. This introduces some bias which we plan to remove in the future.

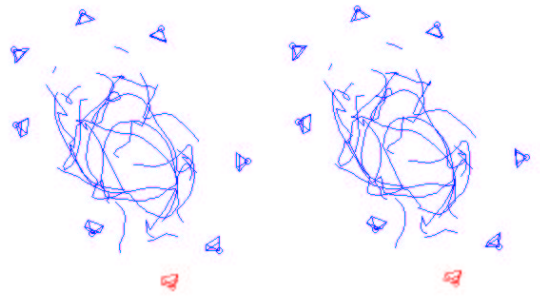

Figure 4: *A top view of the reconstructed virtual calibration object and the resulting camera calibration. The lines show the path of the marker object, the pyramids depict the reconstructed camera positions. This is a cross-eye image that can be fused to give a depth impression.*

8 Conclusions and Future Work

We have presented a flexible camera calibration system that can be used even if some of the cameras have no common field of view. It requires no especially manufactured calibration object and the calibration process can be partly interactive in the way that the user can be guided to create correspondences where it would be helpful. The method is robust against noise and lifts two restrictions imposed by earlier methods. These are:

- the scene does not have to be dark
- the cameras can be calibrated as long as their camera graph is connected

For the future we plan to include the computation of the real projected midpoints of the sphere instead of using the midpoints of the detected ellipses as correspondences. Furthermore is the detection of all cycles in the graph computationally expensive. Since the search for all cycles is a breadth first search this search visits longer cycles later. Therefore the search can be stopped after sufficiently many cycles have been found. The longer cycles are supposedly the less important ones.

References

Azarbayejani, A., and Pentland, A. 1995. Camera self-calibration from one point correspondence. Tech. Rep. 341, MIT Media Lab.

Chen, X., Davis, J., and Slusallek, P. 2000. Wide Area Camera Calibration Using Virtual Calibration Objects. In *Proceedings of CVPR*, vol. 2, 520–527.

Faugeras, O., and Luong, Q.-T. 2001. *The Geometry of Multiple Images*. The MIT Press.

Fischler, M. A., and Bolles, R. C. 1981. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM 24*, 6 (June).

Hartley, R., and Zisserman, A. 2000. *Multiple View Geometry*. Cambridge University Press.

Heikkila, J., and Silven, O. 1997. A four-step camera calibration procedure with implicit image correction. In *CVPR 97*, 1106–1112.

Horn, B. K., 1990. Recovering baseline and orientation from essential matrix, January.

Kim, E. J., Yum, K. H., Das, C. R., Yousif, M., and Duato, J., 2003. Performance Enhancement Techniques for InfiniBand^{TM} Architecture.

Nelder, J. A., and Mead, R. 1965. A simplex method for function minimization. *Computer Journal 7*, 308–313.

Pollefeys, M. 1999. *Self-calibration and metric 3D reconstruction from uncalibrated image sequences*. PhD thesis.

Theobalt, C., Li, M., Magnor, M., and Seidel, H.-P. 2003. A flexible and versatile studio for synchronized multi-view video recording. Research Report MPI-I-2003-4-002, Max-Planck-Institut für Informatik, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany, April.

Triggs, B., McLauchlan, P., Hartley, R., and Fitzgibbon, A. 2000. Bundle adjustment – A modern synthesis. In *Vision Algorithms: Theory and Practice*, W. Triggs, A. Zisserman, and R. Szeliski, Eds., LNCS. Springer Verlag, 298–375.

Tsai, R. Y. 1987. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation RA-3*, 4 (August).

Zhang, Z. 1996. A new multistage approach to motion and structure estimation: From essential parameters to euclidean motion via fundamental matrix. Tech. Rep. RR-2910, INRIA, June.