

# Multi-Camera Platform Calibration using Multi-linear Constraints

Patrik Nyman, Anders Heyden and Kalle Åström  
Centre for Mathematical Sciences, Lund University, Sweden  
Email: {patnym, heyden, kalle}@maths.lth.se

**Abstract**—We present a novel calibration method for multi-camera platforms, based on multi-linear constraints. The calibration method can recover the relative orientation between the different cameras on the platform, even when there are no feature points that can be seen in more than one camera. It is shown that two translational motions in different directions are sufficient to linearly recover the rotational part of the relative orientation. Then two general motions, including both translation and rotation, are sufficient to linearly recover the translational part of the relative orientation. However, a consequence of the speed-scale ambiguity is that the absolute scale of the translational part is impossible to determine, if no prior information about the motions are known, e.g. from dead reckoning. It is also shown that in case of a planar motion, the vertical component of the translational part can not be determined. However, if at least one feature point can be seen in two different cameras, also this vertical component can be estimated. Finally, the performance of the proposed method is shown in simulated experiments.

## I. INTRODUCTION

Multi-camera platforms have been increasingly popular during the latest years, especially due to the decreasing price of digital cameras and computational power. An important application is to put several cameras on a robot or autonomous vehicle to provide navigational guidance and other visual tasks, see Figure 1.

A multi-camera platform can be regarded as a generalized camera, i.e. a camera with several projection centers, see [1]. There exist general methods for calibrating these types of general cameras, [2], but they are not optimal in the specific case of a multi-camera platform. Another approach, based on factorization can be found in [3].

In order to use a multi-camera platform, it is necessary to calibrate the cameras, i.e. to recover the relative orientation between each pair of cameras. Assuming that the cameras are calibrated, this relative orientation can be described by a rigid transformation, containing a rotation matrix and a translation vector. Assuming also that no information is available on the motion of the platform, the calibration problem becomes non-trivial.

The multi-camera platform calibration problem is similar to the hand-eye calibration problem. If the motion of the platform is known, they are essentially equivalent. However, when the motion of the platform is unknown, it has to be estimated from image information only.

The hand-eye calibration problem was originally formulated as recovering the relative orientation between a robot arm and a camera mounted on the arm. Using known feature points in 3D and known motion of the



Fig. 1. The Care-o-bot, developed at Fraunhofer IPA, [8].

robot arm Lenz and Tsai solved the hand-eye calibration problem in [4] using the hand-eye calibration equation  $AX = XB$ , where  $A$  and  $B$  are known transformations and  $X$  is the unknown relative orientation.

Horaud and Dornaika extended the hand-eye calibration equation to include camera matrices instead of transformations in [5] and also used quaternions to linearize the problem. Later on, Andreff, Horaud and Espiau solved the problem without assuming known 3D-points in [6] and also provided nice linear solutions.

The present paper is partly inspired by Stewenius and Åström, who used multilinear constraints to solve the hand-eye calibration problem in [7]. However, their method can not be applied directly to the multi-camera platform calibration problem, and they only use second order multilinear constraints (epipolar constraints).

We propose a novel method for multi-camera platform calibration based on multilinear constraints, that recovers the relative orientation, without having any common feature points between the different cameras. It is based on the multilinear constraints, thus avoiding unnecessary parameters such as 3D-coordinates of feature points. It is furthermore a linear method, that first estimates the rotational component and then the translational component of the relative orientation.

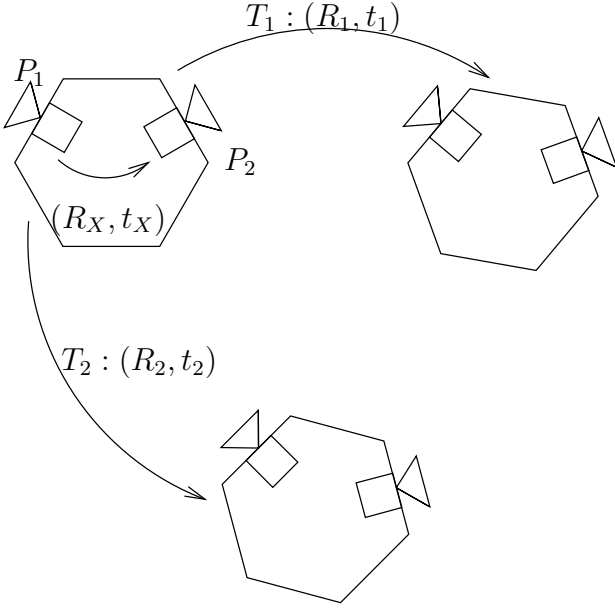


Fig. 2. Illustration of a multi-camera platform making two motions and the corresponding notation.

## II. PROBLEM FORMULATION

Consider a general multi-camera system, see Figure 2, consisting of  $m$  perspective cameras. These can be modeled by the equations

$$\lambda_{ik}\mathbf{x}_{ijk} = P_i T_k \mathbf{X}_j, \quad (1)$$

where  $i = 1, \dots, m$ ,  $j = 1, \dots, n$  and  $k = 1, \dots, p$  denote the camera, point and position, respectively,  $\mathbf{x}$  the image coordinates in homogeneous form,  $P$  the camera matrices,  $T$  transformation matrices encoding the motion of the multi-camera system and  $\mathbf{X}$  the world coordinates in homogeneous form. The multi-camera calibration problem can now be stated as

**Problem 1 (Multi-camera calibration):** Given the image coordinates  $\mathbf{x}_{ijk}$  for  $m$  points, in  $m$  cameras for  $p$  different positions, calculate the camera matrices  $P_i$  (and the transformation matrices  $T_k$  and the object points  $\mathbf{X}_j$ ).

For simplicity, and without lack of generality, we consider the case of two cameras. Furthermore, we assume that there are no points that can be seen in more than one camera and that the cameras intrinsic calibration parameters are known. By a suitable choice of the world coordinate system, we can write (1) as

$$\begin{aligned} \lambda_{i1}\mathbf{x}_{ij1} &= K_1 [I \mid 0] \begin{bmatrix} R_i & t_i \\ 0 & 1 \end{bmatrix} \mathbf{X}_j, \\ \lambda_{i2}\mathbf{x}_{ij2} &= K_2 [R_X \mid t_X] \begin{bmatrix} R_i & t_i \\ 0 & 1 \end{bmatrix} \mathbf{X}_j, \end{aligned} \quad (2)$$

where  $R_X$  and  $t_X$  denotes the calibration parameters for the relative orientation, with  $R_0 = I$  and  $t_0 = 0$ , and  $K_i$  denotes the intrinsic calibration matrices, see [9]. Assuming normalized image coordinates (i.e.  $K_1^{-1}\mathbf{x}_{ij1}$  etc) we can write the camera matrices for the two cameras as

$$P_1 = [R_i \mid t_i], \quad P_2 = [R_X R_i \mid R_X t_i + t_X], \quad (3)$$

see Figure 2. We can now state the multi-camera calibration problem for two calibrated cameras as

**Problem 2:** Given the image coordinates for  $n$  points in 2 cameras, given in (3) for  $p$  different positions, calculate the relative orientation  $R_X$  and  $t_X$ .

## III. RECOVERING THE CALIBRATION PARAMETERS

Assuming that we have a number of feature point correspondences between the image in camera 1, we can write down the multilinear constraints (for three images)

$$\text{rank} \begin{bmatrix} I & 0 & \mathbf{x}_1 & 0 & 0 \\ R_1 & t_1 & 0 & \mathbf{x}'_1 & 0 \\ R_2 & t_2 & 0 & 0 & \mathbf{x}''_1 \end{bmatrix} < 7, \quad (4)$$

where  $\mathbf{x}_1$ ,  $\mathbf{x}'_1$  and  $\mathbf{x}''_1$  denote the image coordinates in image 1, 2 and 3 respectively, see [10]. Similarly for the second camera we obtain

$$\text{rank} \begin{bmatrix} R_X & t_X & \mathbf{x}_2 & 0 & 0 \\ R_X R_1 & R_X t_1 + t_X & 0 & \mathbf{x}'_2 & 0 \\ R_X R_2 & R_X t_2 + t_X & 0 & 0 & \mathbf{x}''_2 \end{bmatrix} < 7. \quad (5)$$

These equations are equivalent to

$$\text{rank} \begin{bmatrix} t_1 & R_1 \mathbf{x}_1 & \mathbf{x}'_1 & 0 \\ t_2 & R_2 \mathbf{x}_1 & 0 & \mathbf{x}''_1 \end{bmatrix} < 4, \quad (6)$$

and

$$\text{rank} \begin{bmatrix} R_X t_1 + R_X (I - R_1) R_X^{-1} t_X & R_X R_1 R_X^{-1} \mathbf{x}_2 & \mathbf{x}'_2 & 0 \\ R_X t_2 + R_X (I - R_2) R_X^{-1} t_X & R_X R_2 R_X^{-1} \mathbf{x}_2 & 0 & \mathbf{x}''_2 \end{bmatrix} < 4. \quad (7)$$

We can now state the calibrated multi-camera calibration problem for two cameras as

**Problem 3:** Given the image coordinates for  $n$  points in 2 cameras, solve (6) and (7) for  $R_X$  and  $t_X$ .

We will now solve the multi-camera calibration problem for two calibrated cameras by first considering two pure translational motions, giving us  $R_X$  and then two general motions, giving us  $t_X$  also.

### A. Recovering the rotational component

Assume that we have a purely translational motion of the multi-camera platform. Then (6) and (7) can be written

$$\text{rank} \begin{bmatrix} t_1 & \mathbf{x}_1 & \mathbf{x}'_1 & 0 \\ t_2 & \mathbf{x}_1 & 0 & \mathbf{x}''_1 \end{bmatrix} < 4, \quad (8)$$

and

$$\text{rank} \begin{bmatrix} R_X t_1 & \mathbf{x}_2 & \mathbf{x}'_2 & 0 \\ R_X t_2 & \mathbf{x}_2 & 0 & \mathbf{x}''_2 \end{bmatrix} < 4, \quad (9)$$

Now, we can use (8) to solve linearly for  $t_1$  and  $t_2$  and then use (9) to solve linearly for  $R_X$  when  $t_1$  and  $t_2$  are known. This can either be done directly by expanding the minors of (9) or by estimating the translational motion of the second camera, using similar equations as in (8). Thus we obtain translations  $s_1$  and  $s_2$  for the second camera, fulfilling

$$s_1 = R_X t_1 \quad \text{and} \quad s_2 = R_X t_2, \quad (10)$$

which can be combined with either  $R_X^T s_1 = t_1$  and  $R_X^T s_2 = t_2$  or the more classical  $s_1 \times s_2 = R_X (t_1 \times t_2)$  to obtain a linear solution for  $R_X$ .

### B. Recovering the translational component

Once the rotational component  $R_X$  is known we can use (6) and (11) for a general motion with known  $R_X$ , to solve for  $t_x$ . First (6) is used to solve for  $R_1, R_2, t_1$  and  $t_2$  based on standard structure-from-motion techniques. Introducing  $\hat{x}_2 = R_X^{-1}x_2$  and similarly for  $x'_2$  and  $x''_2$  and  $\hat{t}_X = R_X^{-1}t_X$ , (7) can be written as

$$\text{rank} \begin{bmatrix} t_1 + (I - R_1)\hat{t}_X & R_1\hat{x}_2 & \hat{x}'_2 & 0 \\ t_2 + (I - R_2)\hat{t}_X & R_2\hat{x}_2 & 0 & \hat{x}''_2 \end{bmatrix} < 4. \quad (11)$$

Since  $R_1, R_2, t_1$  and  $t_2$  are known (11) can be solved linearly for  $\hat{t}_X$ . However, observe that due to the speed-scale ambiguity,  $t_1$  and  $t_2$  can only be recovered up to an unknown global scale, implying that  $\hat{t}_X$  and consequently  $t_X$  can also only be recovered up to an unknown global scale. In order to recover the length of the translational vector  $t_X$  some 3D-information is necessary. One possibility is when the multi-camera platform is mounted on a controlled robot and the actual moved distance is known from dead reckoning. Then the length of  $t_1$  and  $t_2$  are known and consequently the length of  $t_X$  can be calculated. Another option is to use a known distance between two 3D-points.

### C. Planar motions

A common situation when mounting a multi-camera platform on an autonomous vehicle is that of planar motion. When the vehicle is restricted to move on a planar surface, e.g. the floor, and rotation only with a rotational axis perpendicular to this plane, the motion is said to be planar. In this case, it can be seen from (11) that the component in the normal direction to the plane of  $\hat{t}_x$  can not be recovered since

$$\begin{aligned} (I - R_i)(\hat{t}_x + \lambda n) &= (I - R_i)\hat{t}_x + (I - R_i)\lambda n = \\ &= (I - R_i)\hat{t}_x + \lambda n - \lambda n R_i = (I - R_i)\hat{t}_x, \end{aligned}$$

since  $R_i n = n$  and  $n$  denote the normal vector to the plane and  $\lambda$  denotes an arbitrary constant.

One way to resolve this ambiguity is when at least one feature point can be seen in two cameras. Then the epipolar matrix between these two cameras are known, except for the normal component of  $\hat{t}_x$ , which then can be recovered from the epipolar constraint for this point.

## IV. EXPERIMENTS

Some preliminary experiments have been carried out in order to evaluate the ideas. The experiments were done on simulated data. A cloud of world coordinates was randomly generated. The calibration parameters,  $R_X$  and  $t_X$ , were randomly defined. The motions of the platform were defined accordingly. The experiments were first carried out in a noise-free environment, in order to confirm the theory. Secondly, zero-mean noise was added to the image coordinates. In the figure the noise is given in degrees. For a camera with a field of view of 60 and a resolution of 600 pixels, 0.1 degrees correspond to 1 pixel. Several different noise levels were evaluated and these are represented in Figure 3. The error for the rotation matrix,

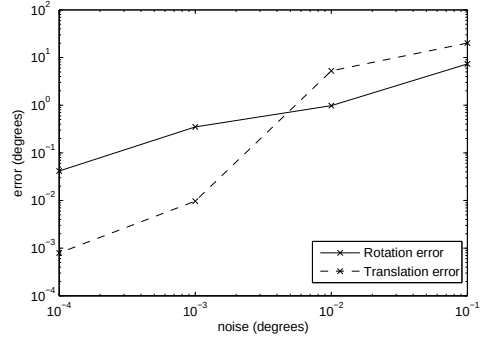


Fig. 3. Error in the determination of the rotational matrix,  $R_X$ , and the translational vector,  $t_X$

$R_X$ , was calculated as the root mean square of difference of the corresponding angles for the estimated  $R_X$  and the correct  $R_X$ . Since the global scale of  $t_X$  could not be determined the error between the estimated  $t_X$  and the actual  $t_X$  was calculated as the angle between the vectors. From Figure 3 one can determine that the method performs well on data with moderate noise and provides sufficient accuracy in order for non-linear optimization methods to be used on higher noise levels.

## V. CONCLUSIONS

We have proposed a novel method for (extrinsic) multiple-camera platform calibration, which is based on multilinear constraints. It is a two-step method that first recovers the rotational component from a purely translational motion and then the translational component from a general motion. We have shown in simulated experiments that it performs well on data corrupted with moderate noise and provide sufficient accuracy for a following non-linear optimization for higher noise levels. Further studies will be made on increasing the robustness and accuracy of the method and showing the applicability on real data.

## REFERENCES

- [1] R. Pless, "Using many cameras as one," in *Proc. International Conference Computer Vision and Pattern Recognition*, 2003.
- [2] M. Grossberg and S. Nayar, "A general imaging model and a method for finding its parameters," in *Proc. International Conference on Computer Vision*, 2001.
- [3] C. Zhang, D.-X. Li, and M. Zhang, "Multi-camera calibration based on iterative factorization of measurement matrix," *International Conference on Multimedia and Ubiquitous Engineering*, pp. 3–8, 2009.
- [4] R. K. Lenz and R. Y. Tsai, "A new technique for fully autonomous and efficient hand/eye calibration," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 3, pp. 345–358, 1989.
- [5] R. Horaud and F. Dornaika, "Hand-eye calibration," *International Journal of Robotics Research*, vol. 14, no. 3, pp. 195–210, 2001.
- [6] R. H. Nicolas Andreff and B. Espiau, "Robot hand-eye calibration using structure-from-motion," *International Journal of Robotics Research*, vol. 20, no. 3, pp. 228–248, 2001.
- [7] H. Stewenius and K. Åström, "Hand-eye calibration using multilinear constraints," in *Asian Conference on Computer Vision*, 2004.
- [8] <http://www.care-obot.de/english/index.php>.
- [9] R. Hartley and A. Zisserman, *Multiple View Geometry*. Cambridge University Press, 2003.
- [10] A. Heyden, "Tensorial properties of multilinear constraints," *Mathematical Methods in the Applied Sciences*, vol. 23, pp. 169–202, 2000.