

# Experiments on mobile robot stereo vision system calibration under hardware imperfection

Ramil Safin<sup>1,\*</sup>, Roman Lavrenov<sup>1</sup>, Subir Kumar Saha<sup>2</sup> and Evgeni Magid<sup>1</sup>

<sup>1</sup> Laboratory of Intelligent Robotic Systems, Higher Institute for Information Technology and Information Systems (ITIS), Kazan Federal University, 35 Kremlyovskaya street, Kazan, 420008, Russian Federation

<sup>2</sup> Mechanical Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi-110 016, India

**Abstract.** Calibration is essential for any robot vision system for achieving high accuracy in deriving objects metric information. One of typical requirements for a stereo vision system in order to obtain better calibration results is to guarantee that both cameras keep the same vertical level. However, cameras may be displaced due to severe conditions of a robot operating or some other circumstances. This paper presents our experimental approach to the problem of a mobile robot stereo vision system calibration under a hardware imperfection. In our experiments, we used crawler-type mobile robot «Servosila Engineer». Stereo system cameras of the robot were displaced relative to each other, causing loss of surrounding environment information. We implemented and verified checkerboard and circle grid based calibration methods. The two methods comparison demonstrated that a circle grid based calibration should be preferred over a classical checkerboard calibration approach.

## 1 Introduction

Stereo vision is widely used in robotics for navigation, obstacle avoidance and path planning [1], object tracking, 3D scene reconstruction, simultaneous localization and mapping [2]. Calibration of a stereoscopic vision system is a necessary step in 3-D computer vision in order to achieve high accuracy in deriving metric information from 2D images such as distance between objects and their dimensions [3]. Calibration is used in order to estimate specific parameters of cameras, which help establishing proper correlations between 2D images and 3D world, and to determine lens distortion models.

Ideally, stereo vision system sensors should be perfectly aligned along the same horizontal line passing through the principal points (image plane center) of both images of the stereo cameras. Otherwise, misalignment could lead to the error growth in measurements [4]. However, in practice, we have to deal with initially presented or appeared later due to multiple reasons hardware imperfection issues, for example, due to severe operation conditions. As in general case it is not possible to solve this issue operatively, the calibration process becomes more complicated and error-prone.

This paper presents our experimental approach to the problem of a mobile robot stereo vision system calibration under a hardware imperfection. In our experiments, we used crawler-type mobile robot «Servosila Engineer» (Fig. 1). The left camera of the stereo pair was displaced after the robot had felt from a significant height during an experiment. Our goal was to verify whether it is possible to achieve acceptable stereo system calibration results in a

presence of such hardware imperfection. We implemented and verified checkerboard and circle grid based calibration methods. The two methods comparison demonstrated that a circle grid based calibration should be preferred over a classical checkerboard calibration approach.



**Fig. 1.** «Servosila Engineer» crawler robot.

The rest of the paper is organized as follows. Section 2 presents mobile robot stereo vision system. Section 3 is dedicated to camera calibration approaches. Experiments on calibration with a checkerboard and a circle grid based methods are presented in Section 4. Finally, we conclude in Section 5.

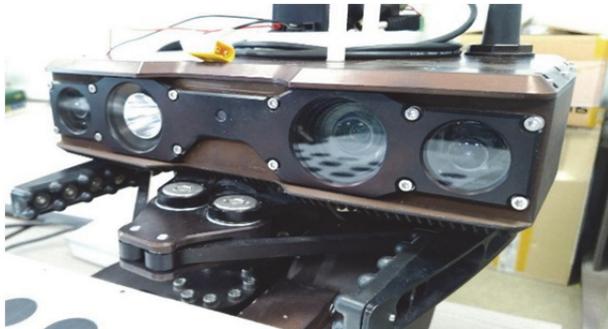
## 2 Robot's vision-related features

Russian crawler-type mobile robot «Servosila Engineer» [5] is equipped with four cameras:

\* Corresponding author: [safin.ramil@it.kfu.ru](mailto:safin.ramil@it.kfu.ru)

- a forward-view optical zoom camera;
- a wide angle of view rear-view camera;
- a pair of cameras for stereo vision.

Stereoscopic vision system consists of two monocular cameras which are located on the front of the robot’s head (Fig. 2). Supported pixel formats and frame resolutions of the stereo system are described in Table 1.



**Fig. 2.** «Servosila Engineer» robot stereo vision system.

**Table 1.** Stereoscopic vision system features of «Servosila Engineer».

Camera model	Image format	Supported camera resolution
DFM 22BUC03-ML	Bayer RGB	744x480 and 640x480

### 3 Camera calibration

#### 3.1 Problem statement

Camera calibration is used in order to estimate extrinsic and intrinsic parameters and distortion coefficients of the camera systems (i.e., monocular or stereo) [6, 7].

Intrinsic parameters, which specify a camera itself, include internal optical characteristics of the camera:

- a principal point (location of an image center)
- a focal length (a distance between the camera lens and an image plane)
- a skew coefficient
- a radial and a tangential distortion coefficients of the camera lens

Extrinsic, or external, parameters relate the world coordinate system to the camera centered coordinate system. It is defined by a translation vector and a rotation matrix. Both intrinsic and extrinsic parameters enable us to transform objects of 3D space into a 2D image and vice versa. This process is referred as a perspective projection and it maps each 3D point to a 2D image plane [6].

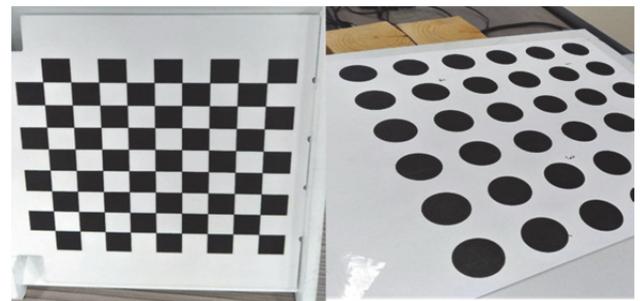
For stereoscopic vision systems, intrinsic parameters are calculated separately for both cameras. However, external parameters are estimated for an entire stereo vision system, and thus spatial relations between the cameras are determined [8]. Obtained parameters are used to align images from both cameras along a same plane. This process is called stereo rectification and it is a crucial step in a case of misalignment of the cameras.

#### 3.2 Calibration using 2D apparatus

To calculate camera parameters, a proper mapping between 3D world points and their corresponding 2D image points is required. This mapping could be obtained applying several images of a calibration pattern (e.g., a checkerboard). With regard to calibration object dimensions, approaches could be classified into four groups [3]:

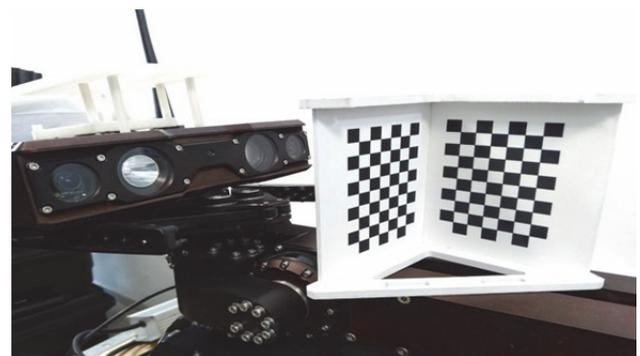
- self-calibration, or zero-dimensional calibration
- one-dimensional object-based calibration
- two-dimensional plane-based calibration
- three-dimensional object-based calibration

The process of calibration using a 2D plane includes tracking of a planar pattern, which is located at different orientations relatively to the camera [9], and various patterns could be used (Fig. 3).



**Fig. 3.** Examples of 2D calibration patterns: a checkerboard (left) and a circle grid (right).

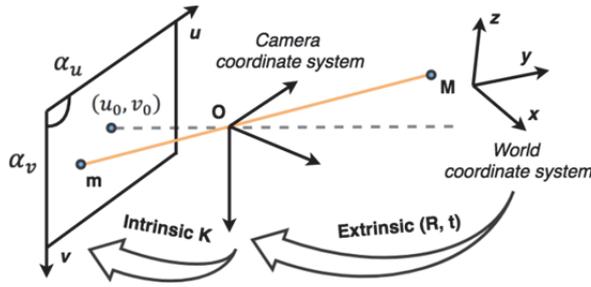
Comparing to other calibration approaches, calibration with 2D apparatus is easier to implement as 2D calibration pattern is easy to create. For example, self-calibration techniques are generally less accurate and require solving more complicated mathematical problem resulting in a more difficult calibration process. In contrast with 3D object-based calibration (Fig. 4), which uses up to three orthogonal to each other planes, 2D object-based calibration does not require an elaborate setup and has a relatively good calibration quality.



**Fig. 4.** 3D reference object-based calibration pattern.

#### 3.3 Pinhole camera model

Most calibration methods are based on a pinhole camera model (Fig. 5) and deal with radial distortion model. The pinhole camera parameters are represented in a matrix that is referred as camera matrix.



**Fig. 5.** Pinhole camera model.

In Figure 5 an optical center of the camera  $O$  denotes the origin of the camera’s coordinate system and the axes of this coordinate system define an image plane. Position of a 3D point  $M$  and the coordinates of its corresponding projection to the image plane  $m$ , representing a 2D point, are denoted as

$$M = [X, Y, Z]^T, \quad (1)$$

$$m = [u, v]^T. \quad (2)$$

The ray, which connects point  $M$  and the camera’s optical center  $O$ , intersects the image plane at point  $m$ . The corresponding augmented vectors of 2D and 3D points are denoted respectively as

$$\tilde{M} = [X, Y, Z, 1]^T, \quad (3)$$

$$\tilde{m} = [u, v, 1]^T. \quad (4)$$

Intrinsic, or internal, parameters of the camera are defined as the following 3-by-3 matrix (intrinsic matrix):

$$K = \begin{bmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (5)$$

where  $\alpha_u$  and  $\alpha_v$  denote focal length in  $u$  and  $v$  axes directions respectively,  $(u_0, v_0)$  are the coordinates of the center of the image plane, and  $\gamma$  is a skew coefficient. If the image plane axes are orthogonal to each other, then the skew coefficient is equal to zero.

Extrinsic, or external, parameters of the camera represent a transformation from a 3D world coordinate system to a 3D coordinate system of the camera. It includes of rotation and translation:

$$[R \ t] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix}, \quad (6)$$

where  $R$  is a rotation matrix,  $t$  is a translation vector.

A 3-by-4 matrix  $C$ , consisting of both extrinsic and intrinsic parameters (Eq. 5, 6), is called camera projection matrix, or camera matrix, and it could be defined as follows:

$$C = K [R \ t]. \quad (7)$$

Using Eq. 3, 4, 7 it is possible to define the relationship between 3D point  $M$  and its image projection  $m$ :

$$s\tilde{m} = K [R \ t] \tilde{M} \equiv C \tilde{M}, \quad (8)$$

where  $s$  is a arbitrary scale factor.

This way, 3D world points are transformed into 3D camera model coordinates using the external parameters. Given 3D points camera coordinates are projected into the image plane using the internal camera parameters (Eq. 5). To estimate the camera parameters, 3D world points and their corresponding 2D image points are used. In order to get those correspondences (samples), multiple images of a calibration pattern could be used, such as a checkerboard or circle grid (Fig. 3). Thus, the camera parameters can be determined.

## 4 Calibration experiments

### 4.1 Hardware imperfection issues

In our experimental work the goal was to calibrate the stereo vision system of Russian crawler-type mobile robot «Servosila Engineer» (Fig. 2). Even though initially the stereo pair was mounted correctly by the manufacturer, the left camera was displaced from its original position as a result of the robot turning upside down and falling from a significant height during an experiment. Figures 6-7 demonstrate that the images from the left and from the right cameras are displaced relative to each other (the left camera’s frame is higher), so their principal points are not on the same horizontal line; no cropping or any other image processing were applied to the original images. Thus, approximately 20% of visual information will be discarded during the rectification procedure. In addition, barrel lens distortion affects the images.



**Fig. 6.** An image from the right camera of «Servosila Engineer» robot stereo vision system.



**Fig. 7.** An image from the left camera of «Servosila Engineer» robot stereo vision system.

## 4.2 Calibration with a checkerboard

The first calibration method that we used in our experiments was based on checkerboard calibration pattern approach [13]. This method requires the dimensions of the pattern to be known and to be constant during the calibration process. The algorithm detects chessboard corners (as an intersection of two black squares) from left and right images being captured by the left and the right cameras of the stereo vision system respectively. In order to estimate camera system parameters, images are obtained at different orientations of the calibration pattern.

Calibration was conducted using robot operating system (ROS). To capture video stream from the stereo system of the robot we employed our own ROS package for simultaneous video streaming from several different cameras [14]. Stereo camera calibration ROS package was used for calibration purposes [15]. It required a checkerboard template with known dimensions (number of squares and their size in meters). It should be noted that the pattern must be hold horizontally and the number of its columns should exceed the number of rows. When the calibration process successfully terminates, the camera system parameters are displayed and applied.

Four different checkerboard patterns were tested in order to obtain an optimal solution (Table 2). Calibration pattern with dimensions of 8 by 6 and a square side of 26 mm showed acceptable results (Fig. 8, 9). The process of calibration took approximately 5 minutes on our machine.

**Table 2.** Checkerboard patterns performance analysis.

Square size (one side), mm	Columns x rows	Epipolar error (pixels)	Results
26	8x6	0.33	good
26	9x6	0.35	good
46	8x6	1.09	poor
46	11x7	0.48	fair



**Fig. 8.** Calibration with the checkerboard (8x6, cell size is 26 mm, epipolar error is 0.33).



**Fig. 9.** Undistorted images after camera calibration with 8x6 26 mm cell checkerboard.

In an attempt to achieve more accurate calibration results, we tried increasing the number of checkerboard

columns by one column. However, the pattern of 9 by 6 with cells demonstrated almost the same results (Fig. 10).



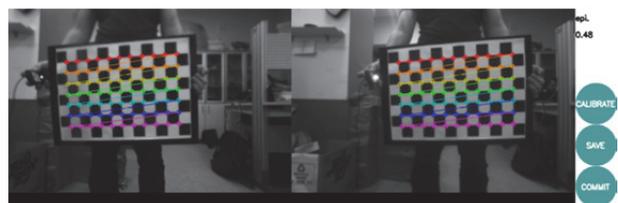
**Fig. 10.** Calibration with the checkerboard (9x6, cell size - 26 mm, epipolar error - 0.35).

In contrast to the calibration pattern with a cell size of 26 mm, a checkerboard with a cell size of 46 mm and dimensions of 8 columns by 6 rows demonstrated poor calibration quality results, producing an epipolar error that exceeded one pixel (Fig. 11). Some of the experimental trials did not converge at all and the obtained distortion coefficients were zero.



**Fig. 11.** Calibration with the checkerboard (8x6, cell size - 46 mm).

In the next experiment, we used a checkerboard with 46 mm cells and 11 columns by 7 rows and successfully estimated distortion (Fig. 12). The obtained epipolar error was greater than the one for the 26 mm cell size pattern (0.48 vs 0.33) and camera parameters estimation time also increased with the size of the pattern.



**Fig. 12.** Calibration using 11x7 checkerboard pattern (cell size - 46 mm, epipolar error - 0.48).

When the pattern size was increased (number of rows and columns), it became harder to accomplish the calibration process. Nevertheless, after successful calibration, images had less common pixels due to cameras displacement (Fig. 13) and some information about the surrounding environment was lost.



**Fig. 13.** Loss of information due to left camera displacement.

### 4.3 Calibration with a circle grid

Circle grid based calibration is another option of camera calibration [16], which could be more accurate than classical checkerboard based approaches. We experimented with robot stereo vision system calibration applying a 6 by 6 circle grid pattern with a circle diameter of 23 mm (Fig. 14). ROS stereo camera calibration package was used. Calibration resulted in estimated distortion coefficients and rectified images (Fig. 15). The obtained epipolar error was 0.18 pixels, which is almost two times better than the result of the best checkerboard based calibration.

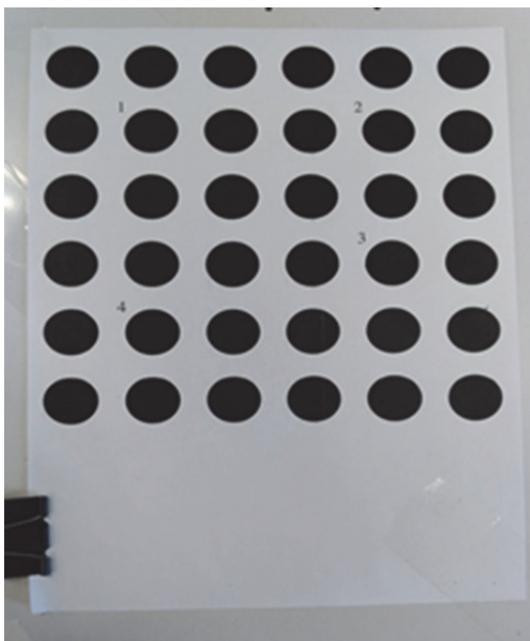


Fig. 14. Circle calibration grid (6x6, circle diameter – 23mm).



Fig. 15. Calibration with the circle grid (6x6, circle diameter – 23 mm).

### 4.4 Monocular calibration experiments

In addition to main set of stereo vision system calibration experiments, we attempted to calibrate only one of the robot cameras and to apply the estimated distortion directly on the other camera. For calibration we selected the right camera as it had not been displaced from its original location. Based on our previous conclusion on a better calibration approach, the circle grid was used. As a result, the error remained, approximately, the same, as if stereo calibration with circle grid pattern was used (Fig. 16).

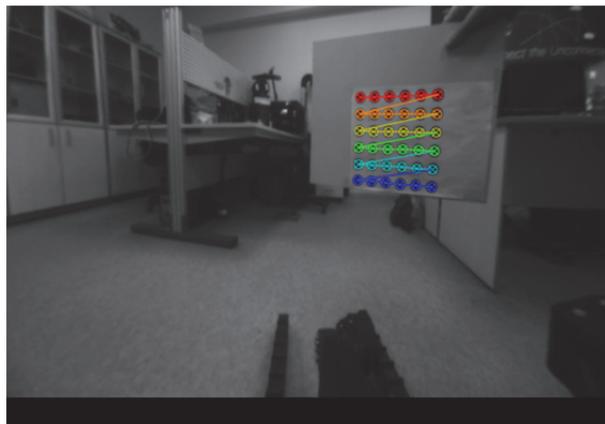


Fig. 16. Calibration of the right camera of the stereo vision system (with a circle grid).

## 5 Conclusions

In this paper, we reviewed and experimentally compared two popular stereo vision system calibration approaches, the checkerboard and the circle grid based approaches. Moreover, the experiments were performed for a mobile robot stereo vision system calibration under a hardware imperfection. We used a crawler-type mobile robot «Servosila Engineer». Stereo system cameras of the robot were displaced relative to each other, causing loss of surrounding environment information. We implemented and verified checkerboard and circle grid based calibration methods. The two methods comparison demonstrated that a circle grid based calibration should be preferred over a classical checkerboard calibration approach. Although fairly good calibration results were achieved, the rectification process discarded almost 20% of the surrounding environment information, which decreases robot’s stereo vision perception information.

As a part of our future work, we plan to implement calibration methods, which are based on fiducial markers. These approaches allow camera calibration in the presence of marker occlusions and rotations [17] and such features are especially important for autonomous self-calibration of a robot in real world operation conditions.

This work was partially supported by the Russian Foundation for Basic Research (RFBR) and Ministry of Science Technology & Space State of Israel (joint project ID 15-57-06010). Part of the work was performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

## References

1. C. Brand, M. J. Schuster, H. Hirschmüller, M. Suppa, In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, 1846–1853 (2014)
2. A. Buyval, I. Afanasyev, E. Magid, Ninth Int. Conf. on Machine Vision, Int. Society for Optics and Photonics, **10341** (2017)

3. Z. Zhang, IEEE Transactions on pattern analysis and machine intelligence, **26** (7), 892–899 (2004)
4. W. Zhao, N. Nandhakumar. Pattern Recognition, **29** (12), 2115–2126 (1996)
5. N. Alishev, R. Lavrenov, Y. Gerasimov, Int. Conf. on Artificial Life and Robotics, 204–207 (2018)
6. J. Suriensky, M. Cmarada, Annals & Proceedings of DAAAM International, (2012)
7. M. Fiala, C. Shu. *Fully automatic camera calibration using self-identifying calibration targets* (Techn. Rep. 48306, 2005)
8. J. Heikkila, O. Silven, Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1106–1112 (1997)
9. P.F. Sturm, S.J. Maybank, Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 432–437, (1999)
10. J. Wang, F. Shi, J. Zhang, Y. Liu, Pattern Recognition, **41** (2), 607–615 (2008)
11. C.B. Duane. Photogramm. Eng, **37** (8), 855–866 (1971)
12. T.A. Clarke, J.G. Fryer, The Photogramm. Record, **16** (91), 51–66 (1998)
13. Z. Zhang, IEEE Transactions on pattern analysis and machine intelligence, **22** (11), 1330–1334 (2000)
14. R. Safin, R. Lavrenov. Int. Conf. on Artificial Life and Robotics, 220–223 (2018)
15. Camera calibration at ROS Wiki. Ros.org, Available at: [http://wiki.ros.org/camera\\_calibration](http://wiki.ros.org/camera_calibration)
16. P. F. Luo, J. Wu. Optical Engineering, **47** (3), 033607 (2008)
17. A. Sagitov, K. Shabalina, L. Sabirova, H. Li, E. Magid. Int. Conf. on Informatics in Control, Automation and Robotics, 182–191 (2017)