Online Optimization of Stereo Camera Calibration Accuracy

Kai Cordes, Lin Chen, Sönke Südbeck, and Hellward Broszio

VISCODA GmbH
{cordes,chen,suedbeck,broszio}@viscoda.com
https://www.viscoda.com

Abstract. Depth estimation from stereo cameras is a fundamental *Computer Vision* task with many applications in automated driving, robotics, scene reconstruction, and medical diagnosis. Depth measurements are based on calibration parameters which are computed in a special calibration procedure. In practice, camera orientations may change slightly over time due to mechanical or thermal effects. Then, calibration parameters no longer match the current stereo camera. This results in inaccuracies in the depth measurement and even complete malfunctions.
We propose a method for the evaluation and optimization of the calibration accuracy during use of a stereo camera. It is based on automatically selected image regions and the rectified stereo configuration. The evaluation monitors the actual accuracy of the calibration. The optimization adjusts the calibration parameters. The time-consuming recalibration using calibration patterns is avoided.

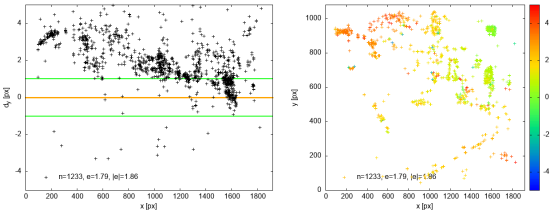
Keywords: Accuracy · Calibration · Stereo · Camera · Optimization · Online

1 Introduction

Stereo cameras are used in a large variety of applications to measure the depth of a 3D scene. For highly accurate measurements, the imaging properties of the stereo camera must be known. These properties are determined in a separate procedure known as calibration using special calibration patterns. Camera calibration is a time-consuming, complicated, and semi-automatic task and there are several approaches for the optimization of the calibration parameters [7,13,6].The accuracy of the calibrations depends on the respective calibration procedure. However, it has been shown that the relative alignment of the cameras change during use due to vibration and mechanical or thermal effects [3]. Thus, accurate stereo calibration cannot be guaranteed over long periods [14,9]. Even well-established datasets have significant systematic errors resulting from sub-optimal calibration parameters for different recording days. In most cases, the camera orientations are subject to changes during data acquisition and the camera parameters are not adapted.

Small errors in the camera parameters lead to significantly decreased reconstruction accuracy [5,15,9]. For given disparity error $\Delta\mathbf{d}$, the range uncertainty increases quadratically with distance [2]. If vertical misalignment in the rectified stereo image occurs, corresponding pixels do not have the same y-coordinate (same *scanline*) and

(a) Rectified stereo image



(b) Vertical disparities d_y for automatically selected corresponding image points in the rectified image pair in Fig. 1a. We show d_y with respect to the x-coordinate (left) and color-coded with (x/y) position in the left stereo image.

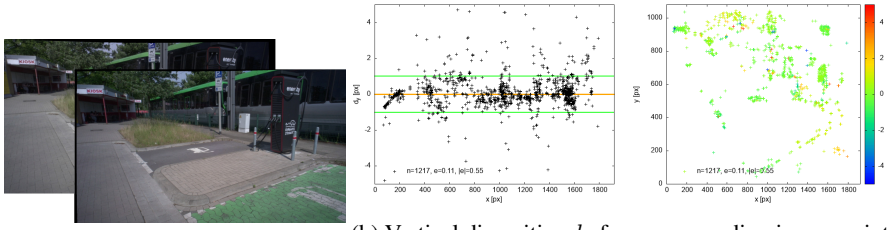
Fig. 1: Visualization of vertical disparities d_y in pixels for the stereo camera in use. For rectified images computed from accurate stereo calibration parameters, the vertical disparity should be zero for all correct point correspondences. They have green color in the right visualization of Fig. 1b. In this example, significant errors with $d_y > 0$ (red color) occur.

traditional stereo matching algorithms such as SGM [8] provide suboptimal results. Hence, vertical misalignment should be avoided.

To quantify the calibration accuracy of stereo camera systems, we evaluate the vertical misalignment using corresponding image points for rectified stereo image pairs. To demonstrate, vertical disparities d_y of corresponding image points are visualized in Fig. 1 for a rectified stereo image pair captured several days after calibrating the cameras. We show d_y for corresponding points with respect to the x-coordinate (Fig. 1b, left) and color-coded with respect to the (x/y) position (Fig. 1b, right) in the left stereo image. The example shows a large systematic error with $d_y > 0$ in the left part of the image. Thus, the calibration parameters no longer match the current camera configuration. The proposed measure for stereo calibration accuracy is used to optimize the calibration parameters. The optimization procedure minimizes the vertical misalignment in the rectified images. We demonstrate that the resulting images have significantly reduced stereo calibration error.

2 Accuracy of Stereo Calibration

Usually, stereo images and cameras are transformed in rectified stereo configuration using a preprocessing step [12]. The rectification computes new extrinsic camera parameters for which both cameras share the same rotation angles and image target plane. It follows, that corresponding image points in left and right image have the same y-coordinate which eases the analysis significantly. Then, the depth z is calculated from the horizontal disparity d_x, the baseline b, and the focallength f as $z = f \cdot \frac{b}{d_x}$. For a reasonable rectification output, accurately calibrated cameras are required. Otherwise, a vertical offset d_y is encountered when comparing corresponding points. Inversely, this offset can be used to quantify the accuracy of the original calibration.



(a) Rectified stereo image

(b) Vertical disparities d_y for corresponding image points in the rectified image pair in Fig. 1a with d_y with respect to the x-coordinate and the (x/y) position in the left stereo image.

Fig. 2: Visualization of the vertical disparities d_y with optimized parameters. Most correspondences have a d_y near zero and have green color in the right visualization of Fig. 2b. The mean vertical disparity is $\epsilon = 0.08\,\text{px}$. The significant error shown in Fig. 1b is eliminated.

For the stereo calibration accuracy measure, we make use of scale invariant keypoint detection and descriptor computation. For keypoint detection, classical approaches still provide higher subpixel accuracy compared to machine learning approaches. As shown in [4], A-KAZE [1] keypoints provide dominant subpixel localization accuracy. To ease the analysis, stereo images and cameras are rectified. We assume a small, but non-zero vertical offset d_y and limit the correspondence analysis to a small search space. Thus, the probability of outliers, i.e., wrongly established correspondences should be small. The disparity $\Delta\mathbf{d}$ for each corresponding feature point pair $\mathbf{p}_l, \mathbf{p}_r$ is defined as $\Delta\mathbf{d} = \mathbf{p}_r - \mathbf{p}_l = (d_x, d_y)^t$. The mean of vertical disparities for all corresponding n features points is [3]:

$$\epsilon = \frac{1}{n}\sum_{i=1}^{n} d_y^{(i)} \tag{1}$$

The mean of the absolute vertical disparities for all corresponding n features points is [3]:

$$|\epsilon| = \frac{1}{n}\sum_{i=1}^{n} |d_y^{(i)}| \tag{2}$$

The mean of the vertical disparities ϵ in equation (1) shows the systematic error resulting from inaccurate calibration parameters and is independent from the feature localization error. The absolute vertical disparities $|\epsilon|$ in equation (2) provides the magnitude of the error. For the example in Fig. 1, $n = 1233$ corresponding keypoints are established. We obtain $\epsilon = 1.79\,\text{px}$ and $|\epsilon| = 1.86\,\text{px}$.

3 Optimization of Stereo Calibration

For the optimization of camera parameters, the mean vertical disparity (equation (2)) is minimized [3]. For each iteration, images and cameras are rectified and correspondences are established to obtain the cost function value ϵ as described in Sect. 2. We optimize 6 parameters (relative position and angles of the right camera).

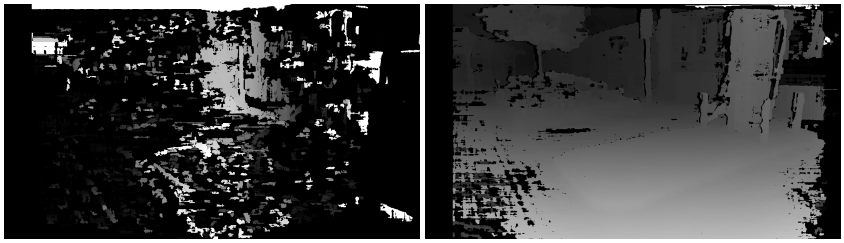


Fig. 3: Depth maps computed by SGM (*Semi-Global-Matching*) for the rectified stereo images from Fig. 1 (left) and the optimized calibration as shown in Fig. 2 (right).

The resulting relative position C_r and angles for the stereo camera in the example experiment are as follows:

$$\begin{aligned} C_r &= (7.91\,mm, -0.24\,mm, -3.67\,mm) \\ (pan, tilt, roll)_r &= (0.09796°, -0.09354°, -0.07998°) \end{aligned} \tag{3}$$

For the new camera parameters, $n = 1217$ keypoints are obtained resulting in $\epsilon = 0.11$ px and $|\epsilon| = 0.55$ px. The corresponding rectified images are shown in Fig. 2a. For validation, the verical disparities are shown in Fig. 2b. Compared to the original versions (Fig. 1), a significant decrease of the vertical disparities is visible. The systematic error diminishes. In Fig. 3 we compare depth maps computed by SGM (*Semi-Global-Matching*) [8] as implemented in OpenCV. Similar to the experimental results in [10], a significant improvement is achieved for the camera parameters optimized with our approach (right) compared to the original stereo calibration (left). Currently, it is unclear, how much these calibration errors affect machine learning models for stereo depth estimation. Initial tests with RAFT-Stereo [11] indicate that these approaches appear less sensitive to small calibration errors. But, an accuracy analysis requires much more effort since the performance is dependent on both train and test datasets and their calibrations. In [3], it is shown that stereo vision datasets have various calibration error structures.

For classical depth estimation methods, cf. Fig. 3, our experiments show the practicability of the parameter optimization. Since no calibration patterns are needed, this procedure can be applied to adjust the camera parameters during acquisition as *Online-calibration*.

4 Conclusions

Stereo cameras in use are subject to thermal and mechanical stresses that lead to fluctuations in the relative alignment of the two cameras. With the developed method, a consistently high accuracy of depth estimation can be ensured during use. Therefore, a keypoints correspondence analysis with high localization accuracy is employed. From the keypoints, vertical disparities are computed. The accuracy measure provides the possibility for error control and indicates the need for a recalibration.

The method is applied to self-recorded stereo images of a test vehicle. The calibration errors occurred under heavy mechanical stress caused by vibrations, e.g. from

driving on cobblestones, are corrected. The proposed methodology enables *Online Calibration* since calibration patterns are not needed.

This work was partially supported by the German Federal Ministry of the Environment, Nature Conservation, Nuclear Safety and Consumer Protection (GreenAutoML4FAS project no. 67KI32007A).

References

1. Alcantarilla, P.F., Solutions, T.: Fast explicit diffusion for accelerated features in nonlinear scale spaces. IEEE Trans. Patt. Anal. Mach. Intell **34**(7), 1281–1298 (2011) 3
2. Bansal, M., Jain, A., Camus, T., Das, A.: Towards a practical stereo vision sensor. In: Conference on Computer Vision and Pattern Recognition (CVPR) -Workshops. pp. 63–63. IEEE (2005) 1
3. Cordes, K., Broszio, H.: Accuracy evaluation and improvement of the calibration of stereo vision datasets. In: European Conference on Computer Vision (ECCV) - VCAD workshop (2024) 1, 3, 4
4. Cordes, K., Grundmann, L., Ostermann, J.: Feature evaluation with high-resolution images. In: Azzopardi, G., Petkov, N. (eds.) Computer Analysis of Images and Patterns. pp. 374–386. Springer International Publishing (2015) 3
5. Dang, T., Hoffmann, C., Stiller, C.: Continuous stereo self-calibration by camera parameter tracking. IEEE Transactions on image processing **18**(7), 1536–1550 (2009) 1
6. Furgale, P., Rehder, J., Siegwart, R.: Unified temporal and spatial calibration for multi-sensor systems. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 1280–1286. IEEE (2013) 1
7. Geiger, A., Moosmann, F., Car, O., Schuster, B.: A toolbox for automatic calibration of range and camera sensors using a single shot. In: International Conference on Robotics and Automation (ICRA) (2012) 1
8. Hirschmüller, H.: Accurate and efficient stereo processing by semi-global matching and mutual information. In: Conference on Computer Vision and Pattern Recognition (CVPR). vol. 2, pp. 807–814. IEEE (2005) 2, 4
9. Kumar, A., Mannan, F., Jafari, O.H., Li, S., Heide, F.: Flow-guided online stereo rectification for wide baseline stereo. In: Conference on Computer Vision and Pattern Recognition (CVPR). pp. 15375–15385 (2024) 1
10. Ling, Y., Shen, S.: High-precision online markerless stereo extrinsic calibration. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 1771–1778. IEEE (2016) 4
11. Lipson, L., Teed, Z., Deng, J.: Raft-stereo: Multilevel recurrent field transforms for stereo matching. In: International Conference on 3D Vision (3DV). pp. 218–227. IEEE (2021) 4
12. Loop, C., Zhang, Z.: Computing rectifying homographies for stereo vision. In: Conference on Computer Vision and Pattern Recognition (CVPR). vol. 1, pp. 125–131. IEEE (1999) 2
13. Scharstein, D., Hirschmüller, H., Kitajima, Y., Krathwohl, G., Nešić, N., Wang, X., Westling, P.: High-resolution stereo datasets with subpixel-accurate ground truth. In: German Conference on Pattern Recognition (GCPR). pp. 31–42. Springer (2014) 1
14. Warren, M., McKinnon, D., Upcroft, B.: Online calibration of stereo rigs for long-term autonomy. In: IEEE International Conference on Robotics and Automation. pp. 3692–3698 (2013) 1
15. Xu, D., Zeng, Q., Zhao, H., Guo, C., Kidono, K., Kojima, Y.: Online stereovision calibration using on-road markings. In: 17th International IEEE Conference on Intelligent Transportation Systems (ITSC). pp. 245–252. IEEE (2014) 1