Real-Time Stereo Camera Calibration Using Stereo Synchronization and Erroneous Input Image Pair Elimination

M. S. Shashi Kumar and N. Avinash

Abstract With the increasing usage of stereo cameras in electronic gadgets, it becomes necessary to have fast, accurate, and automatic stereo calibration during production. In this paper we propose a novel method to achieve fast, accurate and automatic stereo calibration on the video feed using with chessboard pattern. Stereo calibration requires optimal stereo pair images at various orientations without delay between capturing of left and right cameras when objects are moving. In this paper we have developed a novel software based approach to capture synchronized frames from the stereo camera setup with very minimum delay between left and right camera of stereo setup. An approach to reject unmatched stereo pairs is developed based on z-score method, so that only valid optimal image pairs are used in stereo calibration. The optimal sets of stereo synchronized error free images are used for stereo camera calibration and calibration results are stored. The entire process runs in one shot real time without human intervention thus speeding up the stereo camera calibration process.

Keywords: Z-score · Stereo calibration · Stereo synchronization

1 Introduction

Stereo calibration is used to estimate geometric relation between cameras in the stereo setup. Stereo calibration requires input image pairs with different views. If we use time variant video streams for calibration then we need to select correct

M. S. Shashi Kumar (✉) · N. Avinash
Wittybot Technologies, Bangalore, India
e-mail: shashikumar@wittybot.com

N. Avinash
e-mail: avinash@wittybot.com

P. P. Swamy and D. S. Guru (eds.), *Multimedia Processing, Communication and Computing Applications*, Lecture Notes in Electrical Engineering 213,
DOI: 10.1007/978-81-322-1143-3_5,

stereo pairs which are captured at same instance of time. Hence, it requires specialized system architecture to select stereo image pair simultaneously. This process is economically expensive because of additional hardware. In this paper, we handle this issue using software technique where we use timestamps while capturing the images from the stereo camera setup. Stereo image pair selection is based on minimum time difference between two images of stereo cameras. Some of the stereo synchronized image pair may be blurred or may not cover the field of view. Hence these image pairs have to be removed from the input set of images for stereo calibration. Z-score technique is used to eliminate the wrong image pairs and retain appropriate image pairs for stereo calibration. The above process is used to stereo synchronized image pairs from video feed, selecting optimal stereo pairs from accumulated pairs and calibrating these optimal stereo pairs is automated as a one step process for calibrating the stereo setup. These calibration parameters can be further used as needed by application.

Stereo calibration is the process of computing the geometrical relationship between the two cameras in space. Stereo calibration depends on finding the rotation matrix (R) and translation vector (T) between the two cameras. Further Rx, Ry, Rz and Tx, Ty, Tz define the R and T about the x, y, z axes of the 3d Cartesian space. Cipolla [1] proposes a method to that uses rigidity constraints of parallelism and orthogonality and uses vanishing point technique to find intrinsic and extrinsic parameters. Faugeras and Toscani [2] proposes a technique to use least squares method to obtain a transformation matrix which relates 3D points with their 2D projections. The advantage here is the simplicity of the model which consists in a simple and rapid calibration. In this work we use the technique explained in [3, 4, 8] to find R and T between two cameras.

Stereo Rectification is the process of aligning image planes to a common plane so that the epipolar lines become collinear without rotating the actual cameras [5]. The epipolar lines become parallel to the horizontal axis of image after rectification, so that it is possible to scan along corresponding rows from two images. In this paper we use stereo rectification to verify the stereo calibration parameters.

Most of the work carried out on automatic stereo calibration is about detecting the chess board corners. Arturo de la Escalera and Jose María Armingol [6] propose a method to detect number of chess board corners using Hough lines. Very less work has emphasized about stereo synchronization of frames in stereo setup and elimination of error pairs. Stereo synchronization problems arise when the calibration object is moving and there is delay in capturing between left and right image of the stereo setup. Error in calibration can also happen due to blurring of image, calibration object might have moved in one of the images due to image synchronization problems, calibration object may not be visible in one or both images of the stereo setup to achieve optimal calibration parameters. Filtering of all these kinds of images is required to generate optimal set of images for accurate calibration.

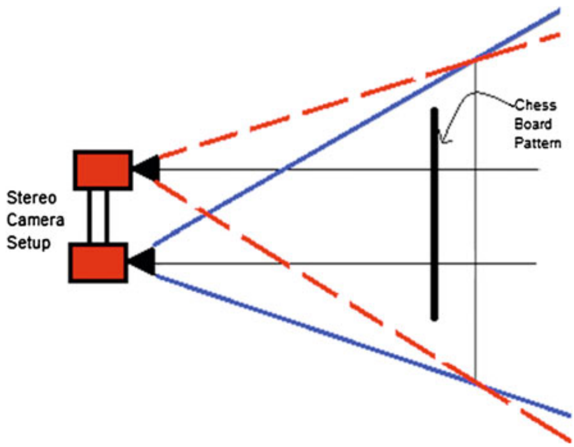
In this paper we emphasize on accurate stereo calibration by selecting image pair with minimal delay between capturing of images from left and right cameras of stereo setup. Followed by error pair elimination we find and remove unmatched image pairs by applying z-score technique to find optimal set of image pairs for stereo calibration.

This paper is organized as follows: In Sect. 2, overview of the system is described. In Sect. 3, detailed descriptions of proposed methodology are described. In Sect. 4, experimental results are presented and finally in Sect. 5 we add concluding remarks.

2 Overview of the System

The experimental setup consists of two cameras rigged side by side as shown in the Fig. 1. The chess board pattern is held in front of the camera such that entire chess board is inside the field of view of both left and right cameras of stereo setup. Once the program is started the user continuously changes the orientation of the chess board till the calibration is successfully completed. Internally in the program stereo synchronization followed by error pair elimination to generate optimal image pairs for calibration which are calibrated. The calibration results are checked for consistency over 1 sigma standard deviation over few trials. If not inside the 1 sigma limit one more stereo synchronized frame is added to the queue. Stereo calibration is performed including the new frame added and standard deviation is checked for 1 sigma limits. Like this more stereo pairs are added in the queue till stereo calibration parameters are converged to 1 sigma limits. Once calibration results get converged these calibration parameters are stored for further usage.

Fig. 1 Stereo camera setup



3 Proposed Methodology

Stereo setup as shown in Fig. 1 gives the input video feed for stereo calibration. Stereo synchronization selects left and right camera images with minimum time delay. Error pair elimination removes any unmatched pair of images for calibration. Using these above steps sufficient number of optimal stereo pairs are collected from the video feed. From the stereo calibration process we calculate the R and T vectors and further rectification matrix. This rectification matrix is computed which can be used to rectify the images (row aligned images) as shown in Fig. 2.

The proposed system consists of following steps:

- Stereo synchronization
- Error Pair elimination
- Stereo calibration
- Image rectification.

3.1 Stereo Synchronization

Video stream from stereo camera are used as input for calibration, and there is always time delay between capturing images for left and right cameras. We need to choose proper image pair that has very minimal delay (nearly equal to zero) between capturing for right and left cameras. This is achieved by using time stamp assigned for each captured image by camera driver. The images with minimum time stamp difference are only selected for camera calibration.

The stereo synchronization process for pairing of images in Fig. 3 and algorithm follows. Here two left and right queues are maintained to fill images from left and right camera in two different threads.

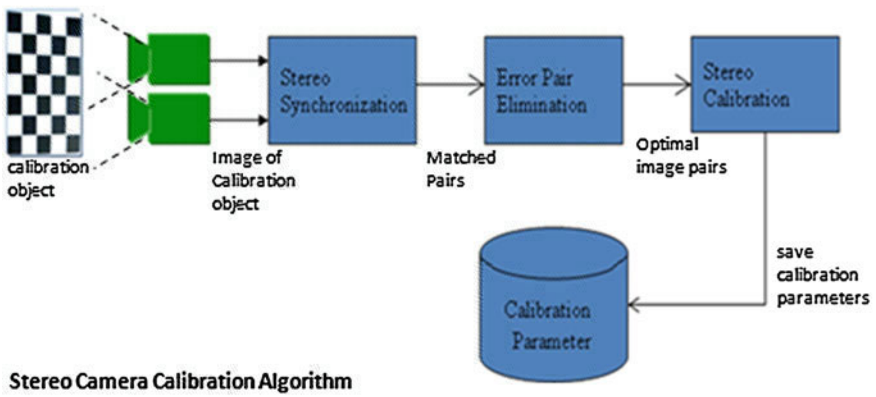


Fig. 2 System block diagram

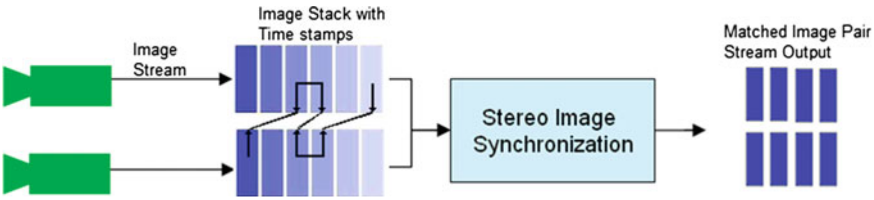


Fig. 3 Stereo synchronization sample output depiction

Start with Left queue as reference queue and 1st frame of left queue as reference frame. Find the time difference with all the frames in the right queue. Consider 3rd frame of right queue has the least time difference. Image Pair 1st frame of left and 3rd frame of right queue. Remove 1st frame in left and 1st, 2nd, 3rd frame from right queue. Now make right queue as reference queue and 1st frame in the right queue (after removal of frames) as reference frame and find the time difference with all the frames in the right queue. Consider 2nd frame in the left queue matches. Pair 1st frame in right queue and 2nd frame in the left queue. Remove 1st frame in right and 1st and second frame in left queue. Now once again make left queue as reference queue. This procedure continues…

- Step 1: Capture images simultaneously from left and right cameras in two different threads into left and right queue in parallel.
- Step 2: Assign left queue for left camera images and right queue for right camera images.
- Step 3: Start with Left queue as reference queue and 1st Frame as the reference frame.
- Step 4: Find the time difference with all the frames in the opposite queue.
- Step 5: Make the next frame in the reference queue as reference frame. Repeat step 4 for all the frames in the queue.
- Step 6: Find the frame with least time difference.
- Step 7: If below threshold, pair these images and remove all the frames above and including paired images in both the queues.
- Step 8: Swap the reference queue to the opposite queue (left to right or right to left).
- Step 9: Go to step3 till all the frames in a queue are emptied.
- Step 10: Go to step 1 till sufficient number of stereo image pairs are generated.
- Step 11: Considering an example with respect to Fig. 3.

3.2 Error Pair Elimination

All the stereo pair images obtained from stereo synchronization may not be useful for stereo calibration. In few stereo paired images the calibration object may be located in the field of view, some are blurred, object may not be visible in one or both of the images, the object might have moved in one of the images due to delay in capturing. Error is introduced by these kinds of image pairs in calibration parameters and in rectification process. In error pair elimination we find and remove unmatched image pairs by applying z-score technique.

The z-score is a common yard stick for different type of data. Each z-score corresponds to a point in a normal distribution and as such is sometimes called a normal deviate since a z-score will describe how much a point deviates from a median or specification point. The z-score is calculated by subtracting your sample median from a data point and dividing by the standard deviation. This value is a measure of the distance in standard deviations of a sample from the median.

The R and T parameters of individual images will give cue that object is clearly visible in the given image or not. These extrinsic parameters of the camera are obtained by using Zhang's monocular calibration method [3]. The detailed algorithm for error pair elimination is given below.

Following is the algorithm for error pair elimination.

Median of absolute deviation (MAD) can be calculated using the following formula. For a data set X_1, X_2… X_n, the MAD is defined as the median of the absolute deviations from the data's median:

$$MAD = median_i(|X_i - median_j(x_j)|) \tag{1}$$

Z-score can be calculated using the following formula

$$x = \frac{\mathrm{ZSCORE\_CONST[Xi\ \ - median(Xj)]}}{\mathrm{MAD(Xj)}} \tag{2}$$

where ZSCORE_CONST is a normalizing constant corresponding to the z-score partitioning of the normal distribution.

Step 1: Apply monocular calibration and use the same to find Stereo R and T vector from left & right for the stereo pair.
Step 2: Find Median of Absolute Deviation (MAD) for all the {Rx, Ry, Rz, Tx, Ty, Tz} for each of the image pair.
Step 3: Calculate z-score for all the extrinsic parameters {Rx, Ry, Rz, Tx, Ty, Tz} for each image pair.
Step 4: Calculate cumulative z-score for each image pair.
Step 5: If z-score of given stereo image pairs greater than z-score threshold then eliminate the corresponding image pairs, use the image pair for store it for stereo calibration.

3.3 Stereo Calibration

Stereo image pairs in the image queue are the optimal pairs of images pairs selected by error pair elimination. Stereo calibration is performed on these optimal pairs. Stereo calibration is the process of computing the geometrical relationship between the two cameras in space. Here in this process of calibrating two cameras at the same time and will be looking to relate them together through a rotation matrix and a translation vector. We obtained the intrinsic parameters of two cameras using Zhang's method [3] as represented below.

$$M_{Lold} = \begin{pmatrix} f_l & 0 & C_{lx} \\ 0 & f_l & C_{ly} \\ 0 & 0 & 1 \end{pmatrix}, \quad M_{Rold} = \begin{pmatrix} f_r & 0 & C_{rx} \\ 0 & f_r & C_{ry} \\ 0 & 0 & 1 \end{pmatrix}$$

Let P(x, y, z) be a point in world co-ordinate observed by stereo cameras. The corresponding image points $P_l(u, v, 1)$ and $P_r(u, v, 1)$ in left and right image planes. In stereo calibration we obtain the relation (with respect to left camera) how much right camera is rotated and translated using essential matrix [5].

$$P_l^T \epsilon\, P_r = 0 \tag{3}$$

where, $\varepsilon = [t_x]R$ is a 3×3 essential matrix with three degrees of freedom of the rotation matrix R and the three degrees of freedom of the translation vector T.

3.4 Image Rectification

Image rectification is the process of "correcting" the individual images so that they appear as if they had been taken by two cameras with row-aligned image planes [3]. After Rectification, the optical axes (or principal rays) of the two cameras are parallel and so we say that they intersect at infinity thus making it stereo disparity calculation between two images simple.

For stereo rectification we used Bouguet's algorithm [7] to minimize the amount of change in reprojection produces for each of the two images while maximizing common viewing area. It rotates each camera half a rotation, so their principal rays each end up parallel. As a result, such a rotation puts the cameras into coplanar alignment as shown in Fig. 4.

4 Results and Analysis

Analysis is concentrated on the following experiments.

1. To evaluate the number of frames required and time to converge.
2. To evaluate the correctness by measuring 3D distance.

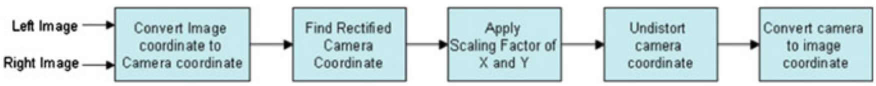


Fig. 4 Block diagram of image rectification algorithm

The experimental setup consists of two Microsoft Lifecam 5,000 webcams placed side by side as shown in the proposed methodology in Fig. 1. Size of captured Image is set same for both the cameras at 640 × 480 pixels. Here we have used a desktop with configuration–Intel Core 2 Duo Processor 2 GHz; 512 MB DDR2 RAM for calibration of stereo setup.

Experiment 1: Following is the algorithm used for obtaining the calibration parameters of the stereo calibration setup. Here we add one frame each time and check for the standard deviation of each of the calibration parameters. Consistent calibration parameters for 1sigma standard deviation is chosen as it covers 34.1 % from mean value so as to achieve accurate calibration parameters.

Step 1: Start the live feed keep changing the orientation of the chess board continuously. Capture first 3 frames at different instances.
Step 2: Perform stereo calibration and store results.
Step 3: Add one more (next) frame to calibration image set and perform stereo calibration and store results.
Step 4: Check if 5set of readings of stereo calibration results are available. If not, go to step 3.
Step 5: If 5 set of readings is available, check if all the extrinsic parameters (Rotation and translation parameters) of the last 5 readings of stereo calibration results are within 1 sigma limits.
Step 6: If not inside 1 sigma limits. Go to step 3 to add one more pair of image to calibration data set.
Step 7: If under 1 sigma limits save calibration details of the last stereo calibration results.
Step 8: Repeat the experiment few times to verify results.

From the Table 1 and Fig. 5 we infer that on an average our algorithm requires about 11–15 frames to calibrate the stereo setup within a time of 15 s where as the system without stereo synchronization and error pair elimination is inconsistent. The accuracy of the calibration is validated in experiment 2.

Experiment: 2 To evaluate the correctness of the proposed algorithm the following experiment is performed. In this experiment stereo rectified image pair is randomly picked and 3D co-ordinates is calculated for each of the chessboard corners. The root mean square error for all the corners in the chess board id calculated. The experimentation steps are as given below.

Step 1: Select one stereo rectified pair.
Step 2: Find the all the corners of chess board in left and right images.
Step 3: Find 3D points for each of the corners. Find the root mean square error in measurement.
Step 4: Repeat steps 1, 2, 3.

Table 2 shows the root mean square error for the proposed method. Here we can observe that the RMS error of 3D distance measured. A chess board of 5 × 8 is used in this experimentation whose box size is of 45 × 45 mm.

Table 1 Table showing number of frames required for calibration over 8 iterations

Trial no	With stereo synchronization and error pair elimination		Without stereo synchronization and error pair elimination	
	Number of frames	Time taken in sec	Number of frames	Time taken in sec
1	13	15	68	220
2	12	14	66	186
3	14	15	60	178
4	13	15	42	198
5	11	12	78	196
6	13	14	92	248
7	11	12	72	188
8	11	11	59	172

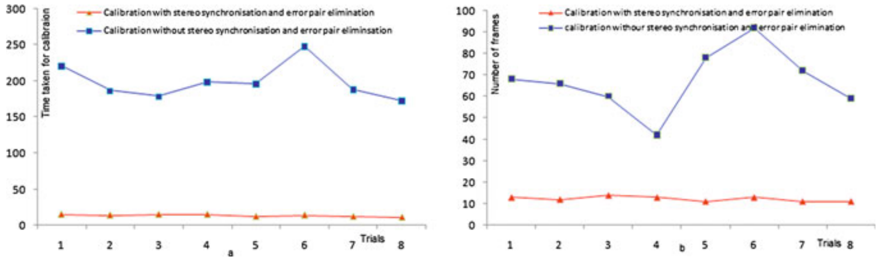


Fig. 5 Graph plotted for (**a**) Time required for with and without stereo synchronization and error pair elimination (**b**) Number of frames required for with and without stereo synchronization and error pair elimination

Table 2 RMS error of 3D distance measurement with and without z-score method

Trial no	RMS error with z-score method in mm	RMS error without z-score method in mm
1	2.098	13.368
2	1.478	6.235
3	1.621	8.628
4	1.416	9.998
5	1.996	12.160
6	1.849	14.963
7	1.288	11.818
8	1.385	5.314

In this experimentation we are measuring error in 3D distance measurement between successive 40 chess board corners. From Fig. 6 we can observe that the RMS error with stereo synchronization and error pair elimination is at a maximum of 2.1 mm over 8 trials whereas without stereo synchronization and error pair elimination is at a maximum of 15 mm.

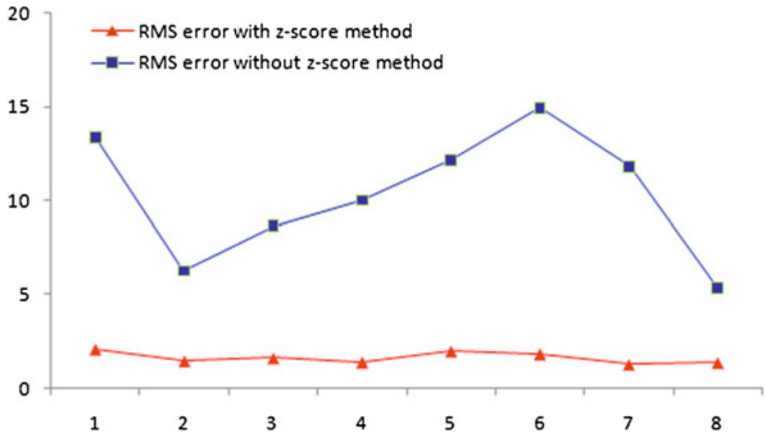


Fig. 6 Graph plotted for RMS error of 3D distance measured with and without z-score method

The inferences obtained from the experiments carried out in this section are as follows. From experiment 1, with proposed method we can calibrate the stereo setup without human intervention for selecting the image pairs or to remove the error pairs. This proposed method converges within 1 sigma standard deviation with in as little as 15 frames. On our experimental setup it takes a maximum of 15 s. On a faster computer it takes significantly less. Also from the experiment 2 it can be noted that the measurement error with the proposed method is significantly lesser as compared to conventional method without stereo synchronization and error pair elimination. Also we can see that the error in the proposed method is quiet consistent.

5 Conclusion

This paper presents a real time efficient and accurate calibration method without human intervention to calibrate a stereo camera setup. The advantage of this method is that it automatically selects stereo synchronized images with very less time difference between image pair, selects optimal pairs for by calibration eliminating bad pairs and calibrates the optimal image pairs selected. All these steps happen in a single step which saves time by fast calibration. Money is saved as extra hardware for stereo synchronization is replaced by stereo approach.

References

1. Cipolla R, Drummond T, Robertson D (1999) Camera calibration from vanishing points in images of architectural scenes. pp 382–391 BMVC
2. Faugeras OD, Toscani G (1986) The calibration problem for stereo. In: Proceedings of the IEEE computer vision and pattern recognition, pp 15–20
3. Zhang Z (2000) A flexible new technique for camera calibration. IEEE Trans Pattern Anal Mach Intell 22(11):1330–1334
4. Opencv 2.1 Reference manual [18th March 2010]
5. Loop C, Zhang Z (1999) Computing rectifying hymnographies for stereo vision. In: Proceedings of IEEE computer society conference on computer vision and pattern recognition, vol 1, pp 125–131
6. Escalera A de la, Armingol JM (2010) Automatic chessboard detection for intrinsic and extrinsic camera parameter calibration sensors (Basel). 10(3), pp 2027–2044
7. Guerchouche R, Coldefy F (2007) Robust camera calibration and evaluation procedure based on images rectification and 3D reconstruction. In: The 5th international conference on computer vision systems, pp 2967–2977
8. Bradski GR, Kaehler A (2008) Learning openCV: computer vision with the openCV library. O'REILLY