

Robust Camera Calibration and Evaluation Procedure Based on Images Rectification and 3D Reconstruction

Rachid Guerchouche and François Coldefy

France Télécom R&D 2, Av. Pierre Marzin 22307 Lannion Cedex France
{rachid.guerchouche, francois.coldefy}@orange-ftgroup.com

Abstract. This paper presents a robust camera calibration algorithm based on contour matching of a known pattern object. The method does not require a fastidious selection of particular pattern points. We introduce two versions of our algorithm, depending on whether we dispose of a single or several calibration images. We propose an evaluation procedure which can be applied for all calibration methods for stereo systems with unlimited number of cameras. We apply this evaluation framework to 3 camera calibration techniques, our proposed robust algorithm, the modified Zhang algorithm implemented by J. bouguet and Faugeras-Toscani method. Experiments show that our proposed robust approach presents very good results in comparison with the two other methods. The proposed evaluation procedure gives a simple and interactive tool to evaluate any camera calibration method.

1 Introduction

Camera calibration aims at estimating the parameters of the relationship binding the 3D world reference space and the 2D camera coordinates system. Camera calibration is essential to many computer vision applications such as images rectification, 3D reconstruction and objects tracking. The camera calibration consists of estimating the *intrinsic* and *extrinsic* parameters representing respectively the internal camera characteristics and the camera pose in the world reference. Many methods are proposed in the literature. Several states of the art can be found in [1–7].

The objective of this paper is twofold. We first present an accurate calibration method based on robust estimation of the *Perspective Projection Matrix* PPM which was already presented in [7], but to which we add some improvements. Second, we propose a new experimental procedure to evaluate camera calibration methods, which is based on images rectification and on 3D reconstruction.

The paper is organized as follows: in the next section, the pinhole camera model is presented. Section 3 details our improved robust camera calibration. Section 4 presents the experimental procedure to evaluate the calibration methods. Experimental results are discussed in Section 5. Finally, Section 6 concludes the paper and opens perspectives of future work.



2 Camera Model

We focus on the pinhole camera model, which is widely used in computer vision. It assumes that the camera performs a perfect perspective transformation \mathbf{P} from the 3D scene coordinates $[X \ Y \ Z]$ to image plane coordinates (u, v) :

$$[\eta u \ \eta v \ \eta]^t = \mathbf{P} [X \ Y \ Z \ 1]^t, \quad (1)$$

where η is an homogeneous factor and $[.]^t$ is the transpose operator. The PPM \mathbf{P} , a 3×4 matrix, is defined as the product $\mathbf{P} = \mathbf{Q} [\mathbf{R} \ | \ \mathbf{t}]$ where \mathbf{Q} and $[\mathbf{R} \ | \ \mathbf{t}]$ are respectively the intrinsic and extrinsic matrices. \mathbf{R} is an 3×3 orthogonal matrix representing the camera orientation and \mathbf{t} the position vector of the camera in the 3D space. More details on this model can be found in [8].

Cameras usually exhibit significant lens radial and tangential distortions [9]. Radial distortion is point-symmetric at the optical center of the lens and causes an inward or outward shift of image points from their initial perspective projection. Imperfect centering of the lens components and/or other manufacturing defects leads to *decentering* effects that can be modeled by tangential distortion.

Our proposed camera calibration algorithm does not include distortion parameters estimation, which can be a drawback when using high distorted optical lens of cheap cameras as webcams or short focal length lens.

3 Robust Camera Calibration

Our algorithm for camera calibration is based on robust estimation of the PPM. The goal is to design a method needing minimal user interaction and using a simple object as a calibration pattern.

The considered object is a cube of known dimensions with six differently colored faces. We developed two versions of our algorithm, the first one running with a single image, and the second one using several images acquired at different depths and positions in the camera field of view, in order to estimate the parameters more accurately. In this last case, all the acquired images should show the same two adjacent faces to the camera. The user will then have to manually identify these two faces on one or another of the images to initialize our calibration algorithm.

An image segmentation procedure is applied to detect the 6 vertices of the two selected adjacent faces. The 6 detected points are used to compute a coarse estimation of the PPM by applying the Faugeras-Toscani method [10]. We refine the estimation by minimizing the Chamfer distance between the projection of the cube edges and the corresponding detected image contours.

In the case of multi-images calibration, the parameters estimation accuracy is performed by a nonlinear optimization technique (Levenberg-Marquardt) over all the acquired images.



3.1 Cube Vertices Detection

The detection of the 6 cube vertices is achieved by first segmenting the calibration images in order to identify the two adjacent faces. The segmentation is performed by a meanshift algorithm applied to the color Luv space because of its well-known nice properties. The user selects, on one image, the both cube faces, whose corresponding segmented region areas are then identified and whose mean colors are computed. When several calibration images are used, the identification of the two faces in the remaining images is performed by determining the segmented regions whose color is most similar with the previously computed cube face colors.

In order to locate the cube's corners on the image, we create for each cube's face a binary image null everywhere except for the pixels belonging to the considered face, and on which we apply a Harris based corner detector. More precisely, at each pixel, we compute the determinant of the Hessian of a smoothed version of the binary image. A pixel p is labeled as a corner if the determinant value at p is a local minimum with absolute value lower than a predefined threshold. However, because the contours of the segmented face region may be insufficiently smooth, some false alarms may occur. In order to eliminate such false detections, we compute the morphological skeleton of the binary face region. As the projected face is a parallelogram, the morphological skeleton has an 'X' form, whose extremities correspond to its corners. All the pixels previously labeled as corner points located too far from the obtained skeleton are removed. For each image, we thus detect six corners corresponding to the vertices of the two adjacent faces. Figure 1 shows two examples of our corner detector corresponding to a simple and a complex scenes. In the first situation, the cube is the sole object present in the scene, and in the second, the cube has been placed in more complex environment which includes other objects and textures. Note that in both examples, the corner detection performs quite well, without needing any parameter tuning neither for the segmentation process nor for the corners detection algorithm.

3.2 PPM Robust Estimation

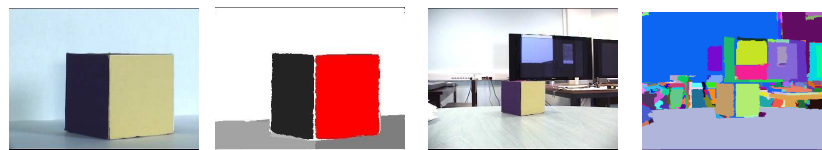
Using the 6 detected vertices, a first coarse estimation of the PPM is computed with Faugeras-Toscani algorithm [10]. However, the corner detector does not provide accurate positions of the cube vertices. A refinement of the PPM estimation is then achieved by minimizing the Chamfer distance between the projected cube edges and the image contours. The image contours are obtained using EDISON system [11–13] (*Edge Detection and Image SegmentatiON*¹).

The cube's 3D edges are first discretized by sampling them uniformly. Let E_i^j be the j^{th} 3D point of the i^{th} discrete cube edge and $e_i^j = P_\theta E_i^j$ its projection onto the image, where P_θ stands for the PPM and θ presents the extrinsic and intrinsic parameters. We aim at minimizing in θ the following functional:

$$U(\theta) = \sum_{(i,j) \in \mathcal{E}} \rho(f_{i,j}), \quad (2)$$

¹ <http://www.caip.rutgers.edu/riul/research/code/EDISON/>





(a) Initial images and color-based segmentation.



(b) Detected faces.



(c) Skeletons of the detected regions.



(d) Detected face corners.

Fig. 1. Semi automatic vertex detection of the cube; (a) two different image acquisitions; the original image (left) and the obtained segmentation (right); (b) the binary images of each cube face; (c) the image skeleton of the cube's faces regions; (d) detected corners(white circles).

where \mathcal{E} is the set of visible projected edge points, $f_{i,j} = D(e_i^j)$ is the value of the Contour Chamfer map D at the point e_i^j and ρ is the Tukey function. The Tukey function thresholds the influence of outliers which corresponds in this case to badly detected image contours. The minimization of θ is performed by an Iterative Reweighted Least Square Algorithm (IRLS) [14] procedure. More details can be found in [7].

3.3 Camera Parameters Estimation Using Several Images

In order to estimate more accurately the camera parameters, several images can be acquired with the cube placed at different depths and positions in the camera field of view. Let us denote by $[\mathbf{R}_k | \mathbf{t}_k]$ the cube orientation and position matrix expressed in the camera world coordinate system for the image k . Let us also denote by $E_{i,j}^k$, the j^{th} cube's discretized of edge i seen in image k and in of the

two selected adjacent faces and $e_{i,j}^k$ its projection onto the image k . Let K be the total number of calibration images. The goal is to determine the intrinsic parameter matrix \mathbf{Q} and the position and orientation matrices $[\mathbf{R}_k | \mathbf{t}_k]_{k=1\dots K}$ which minimize the following functional:

$$V(Q, [R_k | t_k]_{k=1\dots K}) = \sum_{k=1}^K \sum_{(i,j) \in \mathcal{E}_k} D^k(Q[R_k | t_k]E_{i,j}^k), \quad (3)$$

where $D^k(\cdot)$ is the Chamfer distance of the face contour on image k , and \mathcal{E}_k is the set of the visible edge points of the cube on the image k . The minimization of (3) is non linear and is performed by applying the Levenberg-Marquardt algorithm. A first initialization of \mathbf{Q} can be obtained by applying the Faugeras-Toscani method.

4 Evaluation Procedure

This section describes the proposed procedure to evaluate and compare calibration methods. Objective evaluation of such methods is affected by the lack of criteria to compare the final camera parameters estimation obtained by the different methods. Only a few authors compare camera calibration methods [4, 2]. González *et al.* [5] present a comparative analysis of eight camera calibration methods in which they focused on the stability of the camera parameters: (i) the stability of the intrinsic parameters when the camera setup is constant and the calibration pattern is relocated (ii) the stability of the extrinsic parameters when the pattern is constant and the configuration of the camera varies. The conclusion of their study is that the values of the camera parameters estimates depends on the calibration procedure, *i.e* the calibration method, the nature of the calibration pattern and its location in the acquired images. In the case of a constant camera setup, estimated intrinsic parameters values should theoretically not change. However, in practice, these values vary from one calibration process to another. In the case of a fixed camera and pattern, with variation of the camera configuration (focus and/or zoom) the extrinsic parameter estimated values are not constant as they should be.

To cope with this problem, we propose to evaluate and compare calibration methods through image pair rectification and 3D reconstruction of a pattern object. Among all the camera calibration algorithms proposed in the past years, we decide to compare our method to two well known methods: Faugeras-Toscani [10] and modified Zhang algorithm [15] proposed by J. Bouguet [16] (which will be referred as Zhang-Bouguet method).

In order to compare and evaluate the three methods, we set up a vision system composed of three fixed -horizontally, *mvBlueFOX*[®] cameras², with image resolution 1042×768 .

The 3 left images in Figure 2 show sample patterns used for each method. Both the Zhang-Bouguet's and the Faugeras-Toscani's calibration reference points

² <http://www.matrix-vision.com/products/hardware/mvbluefox.php>

were obtained using the corners detector proposed in the Camera Calibration Toolbox for *Matlab*[®] [16], which, after a rectangular zone delimitation by four points, automatically detects the corners of the black and white squares on the patterns. In [17], Krüger *et al.* proposed a fully automatic detector of these points, but it is not used in this work. The cited *Matlab*[®] Toolbox is also used to calibrate the 3 cameras for Zhang-Bouguet method³. The 3 right images in Figure 2 show the images of our planar pattern acquired by our three cameras set up. This known pattern contains 21×21 black and white squares. It will be used to evaluate the different calibration methods through the measure of the consistency of the obtained image pair rectifications and of the 3D planar pattern reconstruction. Note that these images were not used in the calibration process. The image rectification is computed for each camera pair $(i, j)_{i,j=1,2,3}^{i < j}$ using the Fusiello *et al.* [18] compact algorithm.

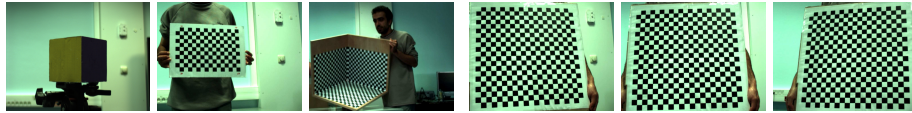


Fig. 2. The 3 left images: patterns used in Robust, Zhang-Bouguet and Faugeras-Toscanni calibration. The 3 right images: the rectification test images.

We also use the Matlab Toolbox corners detector to detect the intersection between the black and the white squares of the planar test pattern. The inner 19×19 squares produce, for each image acquired by the camera $i_{i=1,2,3}$ a set of $N = 400$ organized points $c_n^i = (x_n^i \ y_n^i)_{\{n=1 \dots N\}}$. In the rectification process of an image pair (i, j) , the points c_n^i and c_n^j are respectively mapped on the rectified images to the points $p_n^i = (u_n^i \ v_n^i)$ and $p_n^j = (u_n^j \ v_n^j)$ which should be on a line parallel to the u axis, *i.e.* $v_n^i = v_n^j$.

To evaluate the rectification errors, we compute the Rectification Mean Square Error (*RMSE*) $e^{i,j}$ for each image pair (i, j) as: $e^{i,j} = \sqrt{\frac{1}{N} \sum_{n=1}^{n=N} (v_n^i - v_n^j)^2}$. In the case of multi-images calibration, we dispose of an estimation of the intrinsic parameters matrix Q and of the extrinsic parameters $[R_k | t_k]_{k=1 \dots K}$ for each image k . The *RMSE* is then computed as a global mean over all the couples (p_n^i, p_n^j) obtained from all the estimated $(PPM_k^i, PPM_k^j)_{k=1 \dots K}$ couples for cameras i and j . To compare the accuracy of each calibration method, we compute the global mean error M associated two the three cameras system as: $M = \frac{1}{3} \sum_{i,j=1,2,3}^{i < j} e^{i,j}$.

³ This Matlab Toolbox is an implementation of the Zhang method with some improvements added by its author J. Bouguet (see [16]).

5 Experimental Results

In order to compare the performances of the three methods, we realized 5 calibrations. Two were performed for the Zhang-Bouguet method using 10 and 30 images. Two calibrations were also achieved using our robust calibration: the first using a single image with the cube at the center of the acquired image; the second is performed with 8 additional images of the cube positioned in order to cover the whole camera field of view. The last calibration is performed using the Faugeras-Toscani method with a single image, using the implementation proposed by González⁴. The *RMSE* errors are shown in Table 1.

Method	Number of images	$e^{1,2}$	$e^{1,3}$	$e^{2,3}$	M
Zhang-Bouguet	10	27.30	3.22	19.61	16.71
	30	5.84	2.52	6.67	5.01
Faugeras-Toscani	1	2.83	3.32	5.8	3.74
Robust	1	1.69	2.64	3.75	2.69
	9	5.30	11.21	7.33	7.94

Table 1. Rectification Mean Square Error (in pixels) of the three calibration methods.

All the experiments of Table 1 show large error values except for the calibrations using one image (our robust method and Faugeras-Toscani), and for the stereo pair (1, 3) with Zhang-Bouguet method. The use of several images in our robust algorithm is supposed to improve the calibration accuracy, and thus to minimize the *RMSE*, ... but it is not what we observed! Accuracy with Zhang-Bouguet calibration using 30 images is improved in comparison with the use of only 10 images, but the *RMSE* remains unsatisfying as the mean error is about 5 pixels. The most probable cause of this result is the failure of our multi-images robust method and Zhang-Bouguet's one (except for the (1,3) camera pair) to properly estimate PPM for some image pairs. The estimation of the intrinsic parameters is then biased. It is sometimes due to an inadequate position of the test pattern out of the camera focus which provides blurred images.

We developed a robust algorithm to detect and remove those image pairs. For each stereo pair (i, j) , for all the m calibration image pairs, we estimate the baseline $(B_{i,j})_{k=1..m}$. Theoretically, the baseline is constant. We look for baseline outlier estimates, *e.g.* values which are more than 1.5σ away from the mean (σ is the estimated standard deviation of the baseline). We then suppress the corresponding calibration image pairs, and reestimate the camera parameters. We reiterate until no image pair is removed. Figure 3 shows the plot of the baseline values (*points*), the baseline mean (*continue line*), and the $\pm 1.5\sigma$ deviations (*discontinue line*) found at the first iteration for the Zhang-Bouguet method using 30 images for camera pair (1, 2).

⁴ http://mozart.dis.ulpgc.es/Gias/josep/source_code.htm



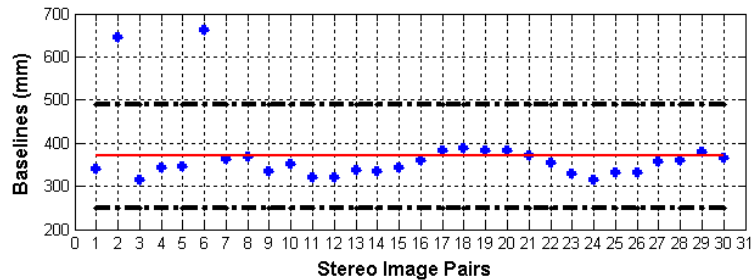


Fig. 3. Plot of the baseline values $B_{1,2}$ for the 30 image pairs in the case of Zhang-Bouguet method; points: baseline values; continue line: the baseline mean; discontinue line: the $\pm 1.5\sigma$ deviations; the two outlier values correspond to the pairs 2 and 6.

Two outlier values are detected for the stereo image pairs 2 and 6. After aberrant image pairs removal, Zhang-Bouguet calibration is performed with only 28 images, Zhang-Bouguet calibration with 10 images is performed with 10 images by replacing the aberrant pairs by other coherent ones, and our robust calibration with 9 images is performed with 7 images (also two aberrant pairs are removed). The obtained results after outliers removal are shown in Table 2.

Method	Number of images	$e^{1,2}$	$e^{1,3}$	$e^{2,3}$	M
Zhang-Bouguet	10	3.33	3.38	1.95	2.88
	28	3.27	2.93	1.63	2.54
Robust	7	1.59	2.41	3.44	2.48

Table 2. Rectification errors (in pixels) after removal of the aberrant image pairs.

The *RMSE* values (about 3 pixels for images 768 pixels height) are now more acceptable. Zhang-Bouguet method using 28 images ($M = 2.54$), our robust one using 7 images ($M = 2.48$) or a single image ($M = 2.69$) show similar accuracy whereas Faugeras-Toscani provides significantly greater error ($M = 3.74$). Another remark concern the introduction of distortion parameters estimation in the case of Zhang-Bouguet calibration.

5.1 3D Reconstruction

In case of multi-images calibration (Zhang-Bouguet and our robust methods), we have as much as PPM estimations as image pairs. For each calibration method and each camera pair, we then select the PPM providing the smallest *RMSE* to perform the 3D reconstruction of our test pattern grid (Fig.1). Table 3 shows for each camera pair the observed minimum *RMSE* among all the calibration image pairs. Zhang-Bouguet and our robust methods show similar accuracy.

Calibration	$e^{1,2}$	$e^{1,3}$	$e^{2,3}$	M
Zhang-Bouguet	1.53	1.96	1.03	1.50
Robust	1.06	1.88	1.83	1.59

Table 3. Minimum *RMSE* among all the calibration image pairs for each camera pair.

Methods	(Roust,Zh-Bou)			(Robust,Faug-Tos)			(Zh-Bou,Faug-Tos)		
Coordinates	x	y	z	x	y	z	x	y	z
Pair (1,2)	3.03	-1.25	-8.29	-33.06	74.61	-85.91	-36.10	75.86	-77.61
Pair (1,3)	1.23	0.32	-1.94	-32.12	73.55	-84.31	-33.35	73.22	-82.36
Pair (2,3)	-0.67	0.24	4.12	32.91	71.96	-83.12	32.24	72.21	-87.25

Table 4. The mean difference of 3D reconstruction obtained from the three different parameters estimation, for each camera pair. The columns shows respectively the mean difference of the reconstruction between our method and Zhang-Bouguet (first column), our method and Faugeras-Toscani (second column) and Zhang-Bouguet and Faugeras-Toscani (third column).

The 3D reconstruction of the 400 pattern grid points, obtained from the rectification applying Fusiello *et al.* algorithm [18], is performed with each couple of camera (i, j) . The world reference has its origin at the middle of the baseline, the x axis parallel to the baseline and the focal axis given from the mean of the focal axis of each camera (see [18] for more details). Table 4 shows for each camera pair the mean difference position in millimeter between the 3D reconstruction of the test pattern grid points obtained from Zhang-Bouguet (calibration with 28 images), Faugeras-Toscani and our robust method (calibration with 7 images). The columns shows respectively the mean difference of the reconstruction between our method and Zhang-Bouguet (first column), our method and Faugeras-Toscani (second column) and Zhang-Bouguet and Faugeras-Toscani (third column). Faugeras-Toscani reconstruction is far away from the two others. Our robust method and Zhang-Bouguet's reconstruction essentially differ on the z axis. The weakest difference occurs for pair (1,3) which corresponds to the most convergent camera pairs.

6 Conclusion

In this paper, we proposed an accurate camera calibration method based on robust estimation of the perspective projection matrix. In order to evaluate our method and compare it to others, we also proposed an efficient experimental procedure based on images rectification and multistereo 3D reconstruction. This procedure can be applied for all stereo systems with unlimited number of cameras. Experimental results show the usefulness of our method compared to



two well known calibration methods: Faugeras-Toscani and modified Zhang algorithm proposed by Bouguet. The proposed experimental procedure offers a simple and interactive tool to evaluate and compare calibration methods. We also demonstrate that in the case of multi-images calibration, the existence of aberrant images considerably affect the accuracy of calibration. We then developed an efficient technique to detect and exclude those images from the calibration process. In future work, we will use distortion parameters estimation in the camera model and we will compare it to photometric camera.

References

1. Clarke, T.A., Fryer, J.G.: The Development of Camera Calibration Methods and Models. *Photogrammetric Record*, **16(91)** (April 1998) 51–66
2. Salvi, J., Armangué, X., Batlle, J.: A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, **35-7** (2002) 1617–1635
3. Hemayed, E.E.: A Survey of Camera Self-Calibration Methods. *Advanced Video and Signal based Surveillance*, **00** (2003) 351–357
4. Zollner, H., Sablatnig, R.: Comparison of Methods for Geometric Camera Calibration using Planar Calibration Targets. *DIME, Proc. of the 28th Workshop of the Austrian Association for Pattern Recognition (OAGM/AAPR)*, **179** (2004) 237–244
5. Isern-González, J., Cabrera-Gámez, J., Guerra-Artal, C., Naranjo-Cabrera, Á.: Stability Study of Camera Calibration Methods. *VI Workshop en Agentes Físicos, WAF'2005, I Congreso Español Informática (CEDI'2005)*
6. Remondino, F., Fraser, C.: Digital Camera Calibration Methods: Consideration and Comparison. *ISPRS Commission V Symposium*, **36(5)** (2006) 266–272
7. Guerchouche, R., Coldefy, F., Zaharia, T.: Accurate camera calibration algorithm using a robust estimation of the perspective projection matrix. *Proc. of SPIE MDIPRCEA IX* **6315** (2006)
8. Horaud, R., Monga, O.: *Vision par Ordinateur: Outils Fondamentaux*. Hermes *2nd edition*(1995)
9. Wei, G., De Ma, S.: Implicit and Explicit Camera Calibration: Theory and Experiments. *IEEE Trans. Pattern Anal. Mach. Intell.*, **16(5)** (1994) 469–480
10. Faugeras, O.D., Toscani, G.: Camera calibration for 3D computer vision. *Proc. of Int. Workshop on Machine Vision and Machine Intelligence*, (1987) 240–247
11. Meer, P., Georgescu, B.: Edge detection with embedded confidence. *IEEE Trans. Pattern Anal. Machine Intell.*, **23** (December 2001) 1351–1365
12. Comanicu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Machine Intell.*, **24** (May 2002) 603–619
13. Christoudias, C., Georgescu, B., Meer, P.: Synergism in low-level vision. *16th International Conference on Pattern Recognition*, **04** (August 2002) 150–155
14. Huber, P.: *Robust Statistics*. New York Wiley, (2000)
15. Zhang, Z.: A Flexible New Technique for Camera Calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, **22-(11)** (2000) 1330–1334
16. http://www.vision.caltech.edu/bouguetj/calib_.doc/
17. Kruger, L. E., Whler, C., Wurz-Wessel, A., Stein, F.: In-factory calibration of multiocular camera systems. *Proc. of the SPIE OMPE*, **5457** (2004) 126–137
18. Fusiello, A., Trucco, E., Verri, A.: A compact algorithm for rectification of stereo pairs. *Machine Vision Applications*, **12(01)** (2000) 16–22

