



UNIVERSIDADE DE COIMBRA

INSTITUTO DE SISTEMAS E ROBÓTICA

DEPARTAMENTO DE ENGENHARIA ELECTROTÉCNICA E DE COMPUTADORES
UNIVERSIDADE DE COIMBRA
3030-290 COIMBRA, PORTUGAL



TÉCNICO
LISBOA

INSTITUTO DE SISTEMAS E ROBÓTICA

INSTITUTO SUPERIOR TÉCNICO
UNIVERSIDADE DE LISBOA
AV. ROVISCO PAIS, 1 – 1049-001 LISBOA, PORTUGAL

ARTICLE:

3D Reconstruction with Low Resolution, Small Baseline and High Radial Distortion Stereo Images

Authors:

Tiago DIAS, E-Mail: tdias@isr.uc.pt

Helder ARAUJO, E-Mail: helder@isr.uc.pt

Pedro MIRALDO, E-Mail: pmiraldo@isr.tecnico.ulisboa.pt

Abstract – In this paper we analyze and compare approaches for 3D reconstruction from low-resolution (250x250), high radial distortion stereo images, which are acquired with small baseline (approximately 1mm). These images are acquired with the system NanEye Stereo manufactured by CMOSIS/AWAIBA. These stereo cameras have also small apertures, which means that high levels of illumination are required. The goal was to develop an approach yielding accurate reconstructions, with a low computational cost, i.e., avoiding non-linear numerical optimization algorithms. In particular we focused on the analysis and comparison of radial distortion models. To perform the analysis and comparison, we defined a baseline method based on available software and methods, such as the Bouguet toolbox [2] or the Computer Vision Toolbox from Matlab. The approaches tested were based on the use of the polynomial model of radial distortion, and on the application of the division model. The issue of the center of distortion was also addressed within the framework of the application of the division model. We concluded that the division model with a single radial distortion parameter has limitations.

Preprint: submitted to the Proc. Int'l Conf. on Distributed Smart Camera (ICDSC '16)
September, 2016.

3D Reconstruction with Low Resolution, Small Baseline and High Radial Distortion Stereo Images

Tiago Dias and Helder Araujo^{*}
Institute of Systems and Robotics,
University of Coimbra
Pinhal de Marrocos – Polo II,
3030 Coimbra - Portugal
{tdias&helder}@isr.uc.pt

Pedro Miraldo[†]
Institute for Systems and Robotics,
Instituto Superior Técnico,
Universidade de Lisboa
Av. Rovisco Pais, 1, 1049-001 Lisboa, Portugal
pmiraldo@isr.tecnico.ulisboa.pt

ABSTRACT

In this paper we analyze and compare approaches for 3D reconstruction from low-resolution (250x250), high radial distortion stereo images, which are acquired with small baseline (approximately 1mm). These images are acquired with the system NanEye Stereo manufactured by CMOSIS/AWAIBA. These stereo cameras have also small apertures, which means that high levels of illumination are required. The goal was to develop an approach yielding accurate reconstructions, with a low computational cost, i.e., avoiding non-linear numerical optimization algorithms. In particular we focused on the analysis and comparison of radial distortion models. To perform the analysis and comparison, we defined a baseline method based on available software and methods, such as the Bouguet toolbox [2] or the Computer Vision Toolbox from Matlab. The approaches tested were based on the use of the polynomial model of radial distortion, and on the application of the division model. The issue of the center of distortion was also addressed within the framework of the application of the division model. We concluded that the division model with a single radial distortion parameter has limitations.

CCS Concepts

•Computing methodologies → Camera calibration; Epipolar geometry; 3D imaging;

Keywords

3D reconstruction; Stereo; Radial distortion

1. INTRODUCTION

3D reconstruction has been subject of significant research [14]. Reconstruction from stereo pairs requires the calibration of both

^{*}T. Dias and H. Araujo were partially funded by QREN Programme "Mais Centro", Grant SCT 2011 02 027 4824

[†]P. Miraldo was partially funded with grant SFRH/BPD/111495/2015, from *Fundação para a Ciência e a Tecnologia*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ICDSC '16, September 12-15, 2016, Paris, France

© 2016 ACM. ISBN 978-1-4503-4786-0/16/09...\$15.00

DOI: <http://dx.doi.org/10.1145/2967413.2967435>



Figure 1: Image of the NanEye stereo pair.

cameras, as well as of the system itself, namely of the relative pose between the cameras [8]. The correspondences between the regions/features/pixels of left and right cameras allow the estimation of the 3D coordinates of the corresponding points. Algorithms for the establishment of correspondences have also been subject of intense research. The "Middlebury Computer Vision Pages" site contains several databases of stereo images for the evaluation of stereo algorithms [1], as well as results of the evaluations. In this paper we focus on modeling the NanEye stereo camera manufactured by CMOSIS/AWAIBA—see Fig. 1. The main specifications of the system, as provided by the manufacturer, are shown in Tab. 1 Our goal is the development of a model and approach so that 3D reconstructions with good accuracy can be performed despite the shortcomings of the system namely:

1. Low resolution;
2. High radial distortion; and
3. Short baseline.

Consider the geometry of a simple fronto-parallel stereo system (Fig. 2). In this configuration, f is the focal length of both cameras, T is the baseline length, Z is the depth and d is the image disparity. As it is well known and for this case disparity d is given by:

$$d = \frac{fT}{Z} \quad (1)$$

and the uncertainty in depth δZ is given by:

$$\delta Z = \frac{Z^2}{fT} \delta d \quad (2)$$

Therefore the short baseline length (T) implies that both small values and uncertainties in the disparity (d) will increase the uncertainties on the depth estimates.

We did not address the problem of estimating the stereo correspondences. This paper focuses on the issues of geometric model-

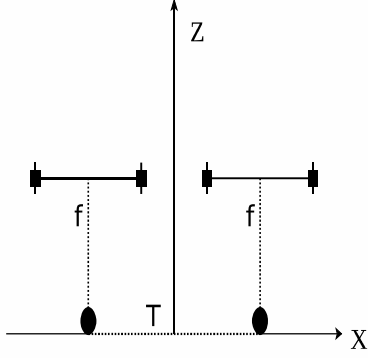


Figure 2: Representation of fronto-parallel stereo geometry.

Resolution (pixels)	250 × 250
Focal length (mm)	0.66
F-number	2.7
Pixel size (μm)	3 × 3
Depth of focus (mm)	5.0– ∞
Size (mm)	2.2 × 1.0 × 1.7

Table 1: Manufacturer’s specifications of NanEye Stereo.

ing a stereo system with low resolution, high radial distortion, and short baseline, so that 3D reconstructions with good accuracy can be obtained.

2. STEREO SYSTEM GEOMETRY

The geometry of a generic stereo system can be represented as shown in Fig. 3. This figure is used for the purpose of definition and representation of the notations used in the paper, such that:

- O^C and O'^C represent the optical centers of the left and right cameras, respectively;
- X^C , Y^C , Z^C and X'^C , Y'^C , Z'^C represent the left and right camera coordinate systems, respectively; and
- x and x' represent the image points, in the left and right cameras, respectively.

Furthermore, we assume that the world coordinate system coincides with the left camera coordinate system. The extrinsic parameters of the stereo system are the relative rotation between the two cameras R and, also, the relative translation T . Therefore, R is the rotation matrix and T is the translation vector of the camera 2 relatively to camera 1.

2.1 Radial Distortion

The images acquired with this system are affected by radial distortion. Radial distortion causes an inward or outward displacement of the pixels (relative to the center of distortion, which may or may not coincide with the image center—in this paper we assumed that they coincide), with the displacement being proportional to the radial distance. In general, radial distortion can be modeled by a polynomial model [3], by a division model [6] or by modeling the camera as a non-central camera [12, 14]. For general parametric non-central models that use (for example) splines, 3D lines corresponding to neighboring pixels are taken into account in the model. In general the polynomial model is applied to the metric coordinates of the pixels (i.e. before the transformation from the image to the sensor plane), while the division model can be applied either to the pixel coordinates in the image plane or in the sensor

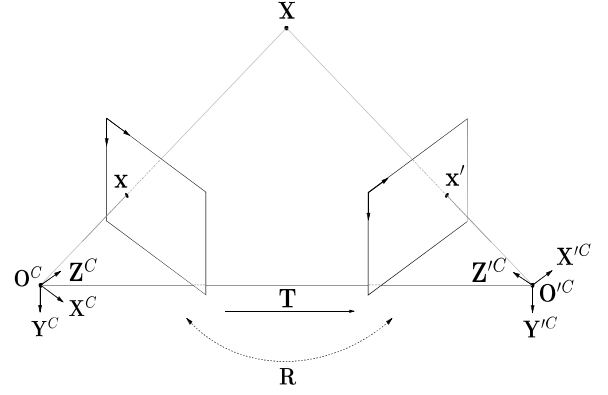


Figure 3: Generic representation of the stereo geometry.

plane [13]. The polynomial model is used in Bouguet’s Camera Calibration Toolbox [2] and in Matlab’s Computer Vision Toolbox. For the purpose of this paper, we considered both the polynomial and the division model.

2.1.1 Polynomial model

Considering the rigid transformation between the world coordinate system and the camera coordinate system T_W^C , then, we have:

$$X^C = {}^C T_W X^W \quad (3)$$

where $X^C = (X^C, Y^C, Z^C)^T$ stands for the point coordinates in the camera coordinate system, and $X^W = (X^W, Y^W, Z^W, 1)^T$ stands for the homogeneous point coordinates in the world coordinate system. The normalized image coordinates are given by:

$$x_n = \begin{bmatrix} X^C/Z^C \\ Y^C/Z^C \end{bmatrix} \quad (4)$$

where $x_n = (x_n, y_n)^T$. The coordinates of the image points affected by radial distortion are obtained from (if we use a model with a polynomial of the 6th degree):

$$x_d = \left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right) \begin{bmatrix} x_n \\ y_n \end{bmatrix} \quad (5)$$

where k_1 , k_2 and k_3 are the radial distortion parameters, $r^2 = x_n^2 + y_n^2$ is the radial distance, and $x_d = (x_d, y_d)^T$ are coordinates of the distorted points. After taking into account radial distortion, the pixel coordinates are computed from:

$$x = K \begin{bmatrix} x_d & y_d & 1 \end{bmatrix}^T \quad (6)$$

where K is the matrix of the intrinsic parameters and $x = (x, y, 1)^T$ are the image coordinates, in pixels.

2.1.2 Rational and Division models

One of the difficulties with the application of the polynomial model is its inversion, since, in general, it is not analytically invertible. One possibility is to use the terms of the Taylor series expansion [10] or perform the numerical inversion. As an alternative it was proposed to model the distortion using either the rational function model [4] or the division model [6]. As already mentioned, and for the purpose of modeling the stereo system described in this paper, we used the division model. The parameters of the division model can be computed using either the fundamental matrix ([15, 11, 6]) or homographies ([6, 9]). This model was originally proposed for uncalibrated images, with a 2nd order polynomial in the

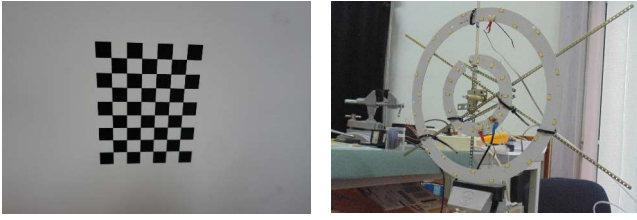


Figure 4: Image of the experimental setup and of one of the checkerboards used in the calibration.

denominator, but it can also be used with calibrated pixels, i.e. with normalized coordinates and with higher order polynomials [13].

This model allows the computation of the undistorted pixels directly from the coordinates of the pixel with distortion, thereby avoiding the inversion difficulty of the polynomial model. Specifically we have:

$$\mathbf{x}_u = \begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{x_d}{1+\lambda r_d^2} \\ \frac{y_d}{1+\lambda r_d^2} \\ 1 \end{bmatrix} = \begin{bmatrix} x_d \\ y_d \\ 1+\lambda r_d^2 \end{bmatrix} \quad (7)$$

where (x_u, y_u) are the coordinates of the undistorted pixel, (x_d, y_d) are the coordinates of the distorted pixel, $r_d^2 = x_d^2 + y_d^2$ is the radial distance, and λ is the distortion coefficient. The division model can use higher order polynomials in the denominator, when additional distortion coefficients have to be estimated.

The computation of λ can be performed using fundamental/essential matrices or homographies. In our case, and given that we used planar checkerboards for calibration, we used homographies to estimate the values of λ for both cameras. If we consider a plane whose images are acquired with projection matrices $\mathbf{P} = [\mathbf{I} | 0]$ and $\mathbf{P}' = [\mathbf{R} | \mathbf{t}]$, the relationship between the images of the points that belong to the plane is given by (up to a scale factor) [9]:

$$\mathbf{x}'_i = \mathbf{H}\mathbf{x}_i \quad (8)$$

where $\mathbf{x}'_i = (x'_i, y'_i, 1)^T$ and $\mathbf{x}_i = (x_i, y_i, 1)^T$ represent the coordinates of the i^{th} corresponding pixels and $\mathbf{H} \in \mathbb{R}^{3 \times 3}$ is the homography matrix ([8]). This relationship between the images of the plane is used to compute the radial distortion parameters for each of the two cameras (corresponding to the division model with only one parameter, i.e. a second order polynomial). Let us denote λ as the parameter for the first camera and λ' for the second camera.

Then, multiplying, on the left, Eq. 8 by the skew symmetric matrix corresponding to $\mathbf{x}'_i$ we obtain:

$$\begin{bmatrix} 0 & -w'_i & y'_i \\ w'_i & 0 & -x'_i \\ -y'_i & x'_i & 0 \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ w_i \end{bmatrix} = \mathbf{0} \quad (9)$$

where $w_i = 1 + \lambda r_d^2$ and $w'_i = 1 + \lambda' r_d'^2$.

From the computation of the planar homography [9]) both distortion coefficients can be estimated. The knowledge of λ and λ' allows the estimation of the undistorted pixels, using the division model (Eq. 11).

2.1.3 Non-Central Model

Non-central models can also be used to model distortion (radial and tangential). One possibility is to use the model proposed and described in [12]. This kind of model can represent generic types of distortions.

Camera 1/Left Camera			
	Dataset 1	Dataset 2	Dataset 3
$f_x(px)$	216.520 ± 1.399	216.307 ± 1.898	216.360 ± 0.579
$f_y(px)$	216.259 ± 1.448	216.421 ± 1.947	216.315 ± 0.576
$u_0(px)$	122.241 ± 1.106	121.847 ± 1.162	122.414 ± 0.758
$v_0(px)$	113.410 ± 0.826	112.981 ± 0.970	111.535 ± 0.646
k_1	-0.349 ± 0.010	-0.353 ± 0.011	-0.369 ± 0.023
k_2	0.148 ± 0.021	0.144 ± 0.021	0.303 ± 0.182
k_3	0.000 ± 0.000	0.000 ± 0.000	-0.366 ± 0.417

Table 2: Intrinsic parameters for the left camera.

Camera 2/Right Camera			
	Dataset 1	Dataset 2	Dataset 3
$f'_x(px)$	213.860 ± 1.903	215.601 ± 2.112	216.258 ± 0.585
$f'_y(px)$	213.325 ± 1.943	215.412 ± 2.161	216.379 ± 0.583
$u'_0(px)$	123.132 ± 1.246	123.486 ± 1.278	124.419 ± 0.739
$v'_0(px)$	124.568 ± 0.906	122.412 ± 1.084	122.283 ± 0.635
k'_1	-0.344 ± 0.011	-0.342 ± 0.012	-0.336 ± 0.004
k'_2	0.143 ± 0.017	0.124 ± 0.020	-0.103 ± 0.208
k'_3	0.000 ± 0.000	0.000 ± 0.000	0.866 ± 0.499

Table 3: Intrinsic parameters for the right camera.

3. STEREO SYSTEM CALIBRATION

To perform the calibration of the stereo system, we used images of a checkerboard, made up of black squares (see Fig. 4). Bouguet's calibration toolbox [2] was used for that purpose. In the specific case of the cameras that we are dealing with, and as a result of the image low resolution and quality, the uncertainties on the estimated parameters depend on the sizes of the checkerboard squares and, therefore, on the average of the distances of the checkerboards from the cameras. For the purpose of calibration, three sets of nineteen images were used. For each set, a different checkerboard was used (squares with different dimensions). All the checkerboards were made up of 9 rows and 7 columns of black and white squares. From each image 48 points were used. An image of the acquisition setup is presented in Fig. 4. The dimensions of the squares and average distances from the cameras were:

Dataset 1: Squares with side lengths of 2.25 mm, located at an average distance of 20 mm from the cameras;

Dataset 2: Squares with side lengths of 4.50 mm, located at an average distance of 50 mm from the cameras; and

Dataset 3: Squares with side lengths of 5.99 mm, located at an average distance of 70 mm from the cameras.

Three calibrations were performed and the intrinsic parameters for both cameras are presented in Tabs. 2 and 3. The intrinsic parameters are:

- f_x and f_y : Focal lengths in pixels along the x and y axes;
- u_0 and v_0 : Coordinates of the principal point; and
- k_1, k_2, k_3 : Radial distortion parameters corresponding, respectively to r^2, r^4, r^6 .

Given the three calibrations, and taking into account the estimated uncertainties on the parameters, we decided to use the values corresponding to the calibration performed with the **Dataset 3**. From these experiments, it is clear that the images acquired by both cameras are strongly affected by radial distortion. This results from the fact that the coefficients of r^6 , where r is the radial distance, are non-negligible. In addition to the intrinsic parameters for both cameras of the stereo system, we also estimated the relative pose between both cameras. This stereo pair is designed and built so that its configuration can be represented by a fronto-parallel system. However inaccuracies occur during the manufacture process.

Relative Pose	
$\omega_x(rad)$	-0.0034
$\omega_y(rad)$	0.0134
$\omega_z(rad)$	0.0341
$T_x(mm)$	1.063 ± 0.040
$T_y(mm)$	-0.203 ± 0.038
$T_z(mm)$	0.118 ± 0.153

Table 4: Relative pose.

	P-average (mm)	D-average (mm)	R (pixels)
Plane 1	2.718	1.267	0.077
Plane 2	3.497	3.2350	0.119
Plane 3	2.805	3.280	0.110
Plane 4	1.777	0.1460	0.313
Plane 5	5.702	5.7960	0.180

Table 5: Results for the baseline approach.

Moreover, since the camera dimensions are extremely small, small deviations from the nominal values may affect the reconstruction accuracy.

The values for relative pose between the two cameras are shown in Tab. 4. The reference coordinate system is the coordinate system of the left camera. Rotation is represented by the angle-rotation values (Rodrigues' representation of rotation) and specifies the rotation between the right and left camera coordinate systems. From the results, we concluded that the rotation can be represented by an angle of approximately 2.1° , around a rotation axis. The translation vector specifies the coordinates of the origin of right camera coordinate system in the left camera coordinate system. As it can be seen from those values, the stereo pair slightly deviates from the fronto-parallel configuration. The world coordinate system coincides with the left coordinate system, and the 3D reconstruction will be performed in the reference coordinate system of the left camera.

4. EXPERIMENTAL RESULTS

In this section, firstly, we define the criteria used for the evaluation, Sec.4.1. Then we tested different reconstruction methods, Sec. 4.2.

4.1 Evaluation Criteria

To evaluate the 3D reconstruction results three error criteria were used:

P-Planarity: The 3D points to be reconstructed lie on planes. Therefore, one of the criteria used to evaluate the quality of the reconstruction is its planarity. For that purpose, for each set of reconstructed points belonging to a plane, a 3D plane was estimated using least-squares. Next, the distance between the 3D reconstructed points and the reconstructed plane was estimated. We repeat this procedure for all the points, and get the average value of the estimated distances. A perfectly planar reconstruction would correspond to a zero average distance (per plane);

D-3D distance: Since the 3D reconstructed points correspond to corners of the squares located in planes, the ground truth 3D distances between consecutive 3D points are known. The distances between consecutive 3D points were measured and the deviations relative to the ground truth (errors) are computed. Their average values (per plane) characterize also the quality of 3D reconstruction; and

R-Reprojection error: The re-projection error is obtained by re-

	P-average (mm)	D-average (mm)	R (pixels)
Plane 1	5.806	2.013	1.512
Plane 2	5.158	4.128	1.215
Plane 3	3.005	3.494	0.405
Plane 4	2.482	0.489	1.171
Plane 5	7.629	14.041	2.117

Table 6: Results for the division model with triangulation.

	P-average (mm)	D-average (mm)	R (pixels)
Plane 1	5.802	2.015	1.512
Plane 2	5.166	4.117	1.215
Plane 3	3.002	3.481	0.405
Plane 4	2.479	0.483	1.171
Plane 5	7.632	14.022	2.117

Table 7: Results for the division model with plane-sweep.

projecting the 3D reconstructed points onto the images and measuring the distances relative to their ground truth images.

4.2 Reconstruction Methods

To evaluate the 3D reconstruction algorithms, we defined a baseline method using the algorithms available in the Bouguet and Matlab Computer Vision toolboxes. For that purpose, the following methods were applied to obtain a reference 3D reconstruction:

- 1-Undistort:** The image coordinates of the points were undistorted by inverting the polynomial distortion model, using a numeric non-linear least-squares optimization; and
- 2-Triangulate:** Using the coordinates of corresponding undistorted image points, the intrinsic parameters of both cameras and the relative pose the coordinates of the 3D points were estimated.

All nineteen images were used to evaluate the reconstruction. However, and for reasons of space, five were chosen to display the results of the evaluation. In each one of the five images, 48 points were used. The reference distance between the 3D points is 5.99 mm.

The results of the evaluation, for the three error criteria previously described, are presented in Tab. 5.

As mentioned in Sec. 2.1, the polynomial model requires a numerical inversion. For that reason, the division model was proposed [6, 9, 15] since it allows the computation of the undistorted coordinates directly from the distorted coordinates, either from the uncalibrated or calibrated images. Using the homography based approach described in Sec. 2.1.2, we computed the distortion parameters for the division model. We computed a single parameter for each image (i.e. we used a second-order polynomial in the denominator). For the computation of the distortion parameters and for the un-distortion of the images, we considered that the center of distortion coincided with the center of the image. Reconstruction was performed using triangulation and plane-sweeping [7, 5], after the un-distortion of the images and using coordinates in pixels. The plane-sweeping method was applied naively, in a direct way. However we plan to apply it adaptively, as a function of the distortion and region of space to be reconstructed. The results, for the same planes, are presented in Tabs. 6 and 7.

These results show that the plane-sweeping approach (as it was applied) is equivalent to triangulation. However, when comparing the removal of distortion using the division model with the baseline approach, it clearly performs worse. The main reason for this behavior stems from the fact that, for each camera, only a single distortion parameter is used.

	P-average (mm)	D-average (mm)	R (pixels)
Plane 1	6.942	0.8080	2.716
Plane 2	4.206	3.7432	0.854
Plane 3	4.864	4.5740	1.656
Plane 4	1.640	0.1930	0.353
Plane 5	833.587	1482	17.741

Table 8: Results for the division model with normalized coordinates, un-distortion and then triangulation.

	P-average (mm)	D-average (mm)	R (pixels)
Plane 1	5.4793	2.015	2.720
Plane 2	4.070	4.1170	0.908
Plane 3	4.433	3.4810	1.710
Plane 4	1.594	0.9110	0.416
Plane 5	26.554	42.788	18.102

Table 9: Results for the division model with normalized coordinates and the triangulating with distorted coordinates Eq. 12.

4.2.1 Center of Distortion

To evaluate the importance of the center of distortion, we applied the distortion model to the normalized images, i.e. After multiplying the image by the inverse of the intrinsic parameters' matrix. In the case of the NanEye stereo cameras, the principal points differ from the center of the images (see Sec. 3). New values for the radial distortions parameters (division model) were computed, also using homographies. A triangulation method that directly estimated the depths, using the distortion parameters, was applied. In this case, reconstruction was performed in a single step. The method performs triangulation directly with the distorted points (with known distortion parameters). Let $\mathbf{x}'_d$ be the normalized distorted coordinates in the right image and $\mathbf{x}_d$ normalized distorted coordinates in the left image with:

$$\mathbf{x}'_d = \begin{bmatrix} x'_d & y'_d & 1 + \lambda' r_d^2 \end{bmatrix}^T \quad (10)$$

and

$$\mathbf{x}_d = \begin{bmatrix} x_d & y_d & 1 + \lambda r_d^2 \end{bmatrix}^T \quad (11)$$

The relationship between the right and left distorted normalized coordinates is given by

$$\alpha_1 \mathbf{x}'_d = \alpha_2 \mathbf{R} \mathbf{x}_d + \mathbf{T} \quad (12)$$

where α_1 and α_2 are the scale factors and $\mathbf{R}$ and $\mathbf{T}$ are the rotation matrix and translation vector, respectively.

Let $\hat{\mathbf{x}}'_d$ be the skew-symmetric matrix corresponding to vector $\mathbf{x}'_d$. Then, multiplying on the left the previous equation by $\hat{\mathbf{x}}'_d$, we obtain:

$$\alpha_2 \hat{\mathbf{x}}'_d \mathbf{R} \mathbf{x}_d + \hat{\mathbf{x}}'_d \mathbf{T} = \mathbf{0} \quad (13)$$

Using this constraint, a system of equations, for all corresponding points in the left and right images, can be determined. The solution of this system of equations can be computed by means of SVD, yielding the α_2 for all points, which constitutes their depths.

The results obtained by first un-distorting the points and then triangulating are presented in Tab. 8, while the results with triangulating directly with the distorted points (using estimates of the distortion, Eq. 12) are represented in Tab. 9.

These results show that the application of the distortion model requires a good estimate of the center of distortion. In the case of these cameras, the principal point differs from the center of distortion, which is one of the reasons for these results. Furthermore

	P-average (mm)	D-average (mm)	R (pixels)
Baseline	0.256	0.6280	0.863
Division Triang	0.008	9.291	57.556
Division Sweep	0.006	9.283	59.374

Table 10: Comparison of distortion models.

	Camera 1/Left Camera		
	Test 1	Test 2	Test 3
$f_x(px)$	217.906±1.009	216.332±0.802	217.396±0.996
$f_y(px)$	217.855±1.004	216.187±0.827	217.121±0.950
$u_0(px)$	121.501±1.096	123.041±1.044	121.813±1.055
$v_0(px)$	113.061±0.954	111.483±0.851	113.324±0.954
k_1	-0.379±0.034	-0.383±0.029	-0.376±0.034
k_2	0.384±0.276	0.370±0.220	0.361±0.182
k_3	-0.588±0.663	-0.425±0.480	-0.575±0.658

Table 11: Intrinsic parameters for the left camera.

the results with the triangulation with the distorted coordinates are equivalent to first un-distorting the points and then triangulating.

4.2.2 Distortion Models: Comparison

To further show that the division model (with one parameter) is insufficient to represent the radial distortion that is parametrized by the polynomial model with three parameters, the following comparison was performed:

- We defined six synthetic 3D points, with known 3D coordinates and a distance of 2.189 mm between neighboring points;
- Using the estimated camera intrinsic parameters, the polynomial distortion coefficients and the relative pose we projected the points into the images; and
- Using these images, three reconstruction methods were performed: the baseline method; the division model & triangulation; and the division model & plane sweeping.

The results are presented in Tab. 10. These results show that the division model with a single coefficient can not correctly model this level of distortion.

4.3 Evaluation Based on Cross-Validation

As it was mentioned in Sec. 3, we used the calibration obtained using the Dataset 3 because of the low uncertainties obtained in the estimated parameters. For that calibration we used all 19 images to calibrate and, then, used images from the same set to evaluate the reconstruction. Using a subset of the calibration images to evaluate the reconstruction may lead to over-fitting. Therefore we decided to use cross-validation to evaluate the reconstruction. Thus, we performed new tests using sets of 10 images for calibration, and the remaining 9 images (in each set) to evaluate the 3D reconstruction. Three tests were made using different combinations/sets of images. Additionally, we used a RANSAC-based plane fitting method to reduce the number of outliers and to improve the planarity results. We used as threshold 15 per cent of the distance between the plane and the origin. Also, a median filter was used to decrease the error due to outliers in 3D distance error.

As it was described in Sec. 3, we used the calibration obtained using the Dataset 3 because of the low uncertainties given by the calibration procedure. For this calibration we used 19 images to calibrate and a subset of the images to evaluate the reconstruction.

The results for the calibration are presented in Tabs. 11 and 12.

For this experiment, the results presented in Tabs. 13,14 and 15 essentially confirm the conclusions previously drawn: the baseline

	Camera 2/Right Camera		
	Test 1	Test 2	Test 3
$f'_x(px)$	215.866±0.993	216.837±0.812	215.473±0.987
$f'_y(px)$	215.902±0.992	217.131±0.839	215.792±0.947
$u'_0(px)$	126.170±1.062	125.429±1.042	125.270±1.012
$v'_0(px)$	120.638±0.925	122.122±0.847	121.327±0.917
k'_1	-0.330±0.032	-0.333±0.032	-0.316±0.034
k'_2	-0.223±0.257	-0.078±0.262	-0.281±0.271
k'_3	1.228±0.605	0.612±0.612	1.356±0.640

Table 12: Intrinsic parameters for the right camera.

		P-average (mm)	D-average (mm)	R (pixels)
Test 1	Plane A1	2.070	1.172	0.114
	Plane B1	3.046	3.000	0.114
	Plane C1	2.323	1.209	0.417
Test 2	Plane A2	2.182	2.401	0.172
	Plane B2	2.857	5.011	0.245
	Plane C2	2.760	1.380	0.455
Test 3	Plane A3	2.818	2.216	0.059
	Plane B3	3.611	3.888	0.088
	Plane C3	3.112	3.536	0.257

Table 13: Results for the baseline model.

method, with three radial distortion parameters generates better results than the variants of the division model we considered (with only one radial distortion parameter). The results for the baseline method are consistently better across all the three criteria, with only a few exceptions.

5. CONCLUSIONS

In this paper we fully characterize the NanEye stereo pair manufactured by CMOSIS/AWAIBA. We show that these cameras are affected by significant radial distortion. We used a polynomial of degree six to model this distortion. We also show that the reconstruction errors are significant and that reconstruction with higher accuracy requires both different radial distortion models and non-linear optimization methods. We also evaluated the division model with a single distortion coefficient and concluded that, while being simpler to use, it yields worse results. We plan to address the problem of 3D reconstruction with these cameras by jointly considering the radial distortion and the small baseline.

6. REFERENCES

- [1] Middlebury computer vision pages. vision.middlebury.edu/stereo/.
- [2] J. Y. Bouguet. Camera calibration toolbox for matlab. www.vision.caltech.edu/bouguetj/calib_doc/.
- [3] D. C. Brown. Decentering distortion of lenses. *Photogrammetric Eng. Remote Sensing*, 1966.
- [4] D. Claus and A. Fitzgibbon. A rational function lens distortion model for general cameras. In *IEEE Proc. Computer Vision and Pattern Recognition (CVPR)*, 2005.
- [5] R. Collins. A space-sweep approach to true multi-image matching. In *IEEE Proc. Computer Vision and Pattern Recognition (CVPR)*, 1996.
- [6] A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *IEEE Proc. Computer Vision and Pattern Recognition (CVPR)*, 2001.
- [7] D. Gallup, J.-M. Frahm, P. Mordohai, Q. Yang, and M. Pollefeys. Real-time plane-sweeping stereo with multiple

		P-average (mm)	D-average (mm)	R (pixels)
Test 1	Plane A1	2.740	1.489	1.792
	Plane B1	2.682	3.063	0.872
	Plane C1	2.004	1.198	1.240
Test 2	Plane A2	3.052	3.503	1.895
	Plane B2	4.547	7.225	2.356
	Plane C2	2.747	2.738	1.227
Test 3	Plane A3	3.273	2.842	1.495
	Plane B3	2.986	4.301	0.743
	Plane C3	2.793	2.728	1.593

Table 14: Results for the division model with triangulation.

		P-average (mm)	D-average (mm)	R (pixels)
Test 1	Plane A1	2.578	1.862	4.497
	Plane B1	2.365	2.946	0.431
	Plane C1	2.077	1.283	0.531
Test 2	Plane A2	4.521	2.899	2.474
	Plane B2	3.928	38.602	9.103
	Plane C2	2.563	2.757	0.602
Test 3	Plane A3	2.396	3.646	2.961
	Plane B3	3.690	3.892	0.705
	Plane C3	3.269	4.054	1.897

Table 15: Results for the division model with normalized coordinates and triangulation.

- sweeping directions. In *IEEE Proc. Computer Vision and Pattern Recognition (CVPR)*, 2007.
- [8] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [9] Z. Kukelova, J. Heller, M. Bujnak, and T. Pajdla. Radial distortion homography. In *IEEE Proc. Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [10] J. Mallon and P. Whelan. Precise radial un-distortion of images. In *IEEE Proc. Int'l Conf. of Pattern recognition (ICPR)*, 2004.
- [11] B. Mičušík and T. Pajdla. Estimation of omnidirectional camera model from epipolar geometry. In *IEEE Proc. Computer Vision and Pattern Recognition (CVPR)*, 2003.
- [12] P. Míraldo and H. Araujo. Calibration of smooth camera models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2013.
- [13] D. Scaramuzza, A. Martinelli, , and R. Siegwart. A flexible technique for accurate omnidirectional camera calibration and structure from motion. In *IEEE Proc. Int'l Conf. of Vision Systems (ICVS)*, 2006.
- [14] R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer, 2010.
- [15] Z. Kukelova, M. Byröd, K. Josephson, T. Pajdla, and K. Åström. Fast and robust numerical solutions to minimal problems for cameras with radial distortion. *Computer Vision and Image Understanding*, 2010.