



# Article Sensor-Aided Calibration of Relative Extrinsic Parameters for Outdoor Stereo Vision Systems

Jing Wang<sup>1</sup>, Banglei Guan<sup>2</sup>, Yongsheng Han<sup>3</sup>, Zhilong Su<sup>1,4,5</sup>, Qifeng Yu<sup>2</sup> and Dongsheng Zhang<sup>1,4,5,\*</sup>

- <sup>1</sup> Shanghai Key Laboratory of Mechanics in Energy Engineering, School of Mechanics and Engineering Science, Shanghai University, Shanghai 200444, China
- <sup>2</sup> College of Aerospace Science and Engineering, National University of Defense Technology, Changsha 410073, China
- <sup>3</sup> Shandong Water Conservancy Vocational College, Rizhao 276826, China
- <sup>4</sup> Shaoxing Institute of Technology, Shanghai University, Shaoxing 312074, China
- <sup>5</sup> Shanghai Institute of Aircraft Mechanics and Control, Zhangwu Road, Shanghai 200092, China
- \* Correspondence: donzhang@staff.shu.edu.cn

**Abstract:** Calibration of the stereo vision systems is a crucial step for precise 3D measurements. Restricted by the outdoors' large field of view (FOV), the conventional method based on precise calibration boards is not suitable since the calibration process is time consuming and the calibration accuracy is not guaranteed. In this paper, we propose a calibration method for estimating the extrinsic parameters of the stereo vision system aided by an inclinometer and a range sensor. Through the parameters given by the sensors, the initial rotation angle of the extrinsic parameters and the translation vector are pre-established by solving a set of linear equations. The metric scale of the translation vector is determined by the baseline length provided by the range sensor or GNSS signals. Finally, the optimal extrinsic parameters of the stereo vision systems are obtained by nonlinear optimization of inverse depth parameterization. The most significant advantage of this method is that it enhances the capability of the stereo vision measurement in the outdoor environment, and can achieve fast and accurate calibration results. Both simulation and outdoor experiments have verified the feasibility and correctness of this method, and the relative error in the outdoor large FOV was less than 0.3%. It shows that this calibration method is a feasible solution for outdoor measurements with a large FOV and long working distance.

Keywords: camera calibration; inclinometer; stereo vision system; 3D displacement measurement

# 1. Introduction

Stereo vision systems based on digital image correlation (DIC) have been widely used in industrial applications due to its advantages such as non-contact, high precision, and full field of view [1,2]. DIC is an efficient algorithm for obtaining deformation by tracking the image surface features of an object [3,4]. To guarantee the accuracy of the measurements, camera calibration is a critical preparatory step. Camera calibration is a process of determining a series of parameters related to the imaging system, including intrinsic parameters and extrinsic parameters. The intrinsic parameters include equivalent focal lengths, principal point coordinates, and image distortion factors, which reflect the relationship between the image coordinate system and the camera coordinate system. They are inherent parameters of the camera and usually constant regardless of environmental changes. The extrinsic parameters reflect the relationship between the coordinate systems. Since it is generally assumed that the world coordinate system coincides with the left camera coordinate system in the stereo vision unit, it is transformed into solving the relative rotation and translation between the dual camera coordinate systems, which usually changes with the camera's position.



Citation: Wang, J.; Guan, B.; Han, Y.; Su, Z.; Yu, Q.; Zhang, D. Sensor-Aided Calibration of Relative Extrinsic Parameters for Outdoor Stereo Vision Systems. *Remote Sens.* 2023, *15*, 1300. https://doi.org/ 10.3390/rs15051300

Academic Editors: Belen Riveiro and Mario Soilán

Received: 17 January 2023 Revised: 21 February 2023 Accepted: 24 February 2023 Published: 26 February 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

Camera calibration methods can be divided into conventional calibration, selfcalibration, and active vision calibration methods. The conventional calibration methods establish the relationship between the pixel coordinates and world coordinates according to the 3D features provided by the calibration objects. Representative works include the classic two-step camera calibration method proposed by Tasi [5] and the widely used calibration method based on the plane chessboard proposed by Zhang [6]. They are especially suitable for a small indoor scene due to the limited size of the calibration board which is precise and expensive. In contrast, it is not necessary to use a specific calibration object for the self-calibration method. Faugeras et al. first proposed the concept of self-calibration in 1992 [7] based on the Kruppa equation. This provided a solution for camera calibration in a large FOV without use of calibration objects, while the accuracy was not always guaranteed due to the low robustness in calculation. Active vision calibration refers to the use of known movements of the camera in camera calibration [8–10]. Although the calibration target is not required, it is expected the camera to be precisely controlled in motion, which potentially causes problems such as expensive devices and operational difficulties. Recently, based on the above principles, many attempts have been made for development of camera calibration for large FOVs. Miyata et al. proposed a multi-camera calibration method using an omnidirectional camera to obtain the camera features along with the stereo vision FOV for calibration, although the position and attitude between the omnidirectional camera and the stereo vision system need to calibrated in advance [11]. Gao et al. proposed a dual-camera calibration method using a zoom lens to change the FOV [12], allowing the calibration of extrinsic parameters with a conventional calibration board. Sun et al. used a one-dimensional bar with feature points in the FOV to achieve stereo vision calibration [13]; although it is light compared with the calibration board, its size still limits the application scenarios. Wang et al. proposed a method of using two points in obtaining the angle information between cameras based on the assumption that the intrinsic parameters are invariant [14]. Liu at al. used a combination of multiple small calibration objects in a large FOV to generate reference points distributed across the entire field for calibration [15]. Zhang et al. designed a new spherical calibration target and proposed an improved separation parameter calibration method for large-field stereo vision measurements [16]. A common difficulty for these methods is that they are time-consuming, and the robustness needs to be improved. Therefore, there are still significant challenges in camera calibration for stereo vision systems in outdoor large FOV conditions. Additionally, the demand continues to increase [17–19].

Therefore, in order to achieve a fast, accurate camera calibration in outdoor large FOV conditions, a novel calibration method has been proposed based on sensor units. The paper is organized as follows: in Section 2, the principle of this calibration method is described. The simulation and experiments in Section 3 are provided to confirm the advantages of the proposed method. The key procedures in improving accuracy are discussed in Section 4 and a conclusion is drawn in Section 5.

## 2. Methodology

In this paper, the stereo vision system for large FOV applications consists of two workstations. Each consists of a measurement unit, including a video camera, an inclinometer, a range sensor, and a computer.

In order to achieve 3D displacement/motion measurement, the binocular stereo vision uses a common arrangement, as shown in Figure 1. The perspective pinhole model is widely adopted to describe the relationship between the word coordinates ( $x_w$ ,  $y_w$ ,  $z_w$ ) and the image coordinates (u, v). Briefly, the projection relationship between the projection

points  $P_l$  and  $P_r$  on the left and right images and the world coordinate point P can be written as Equation (1).

$$s_1 \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = K_1[I|0] \begin{vmatrix} x_w \\ y_w \\ z_w \\ 1 \end{vmatrix}, s_2 \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = K_2[R|t] \begin{vmatrix} x_w \\ y_w \\ z_w \\ 1 \end{vmatrix}$$
(1)

where *s* is the scale factor, and *K* is the intrinsic parameter matrix. The subscripts 1 and 2 denote the left and right image units. *I* is an identity matrix, *R* and *t* are the rotation matrix and translation vector from the world coordinate system, which are aligned with the left camera coordinate, to the right camera coordinate system. The constraint relation of the binocular stereo vision shown in Figure 1 can be described as an epipolar geometry.



Figure 1. Binocular stereo vision model.

The aim of camera calibration is to determine all these coefficients in Equation (1) to establish an analytical model between the image coordinates and world coordinates. Conventional calibration methods based on plane boards are not available due to the size limitation. Camera calibration for large FOVs is generally divided into a two-step process. Since the intrinsic parameters including equivalent focal lengths, principal point coordinates system and the camera coordinate system, they are constant if the camera and the lens are determined. They can be predetermined by means of a calibration board in laboratory conditions. After this operation, camera calibration becomes a job of determining the extrinsic parameters of the essential matrix, which is combination of a rotation matrix and translation vector.

Epipolar constraint allows a spatial point and its left and right image points to remain in the same plane for a dual camera system. This relationship is widely used to determine the extrinsic parameters of imaging systems by generating a set of equations using the corresponding point pairs. Since the intrinsic parameters are predetermined, only the essential matrix matrix E, which is a combination of rotation matrix R and translation vector t, is to be determined.

$$\boldsymbol{E} = [\boldsymbol{t}]_{\times} \boldsymbol{R} \tag{2}$$

Note that  $[t]_{\times}$  is the skew-symmetric matrix of *t*, as shown in Equation (3).

$$[\mathbf{t}]_{\times} = \begin{pmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{pmatrix}$$
(3)

Since the scale of the translation vector is uncertain in the epipolar geometric relationship, only three independent rotation components and two independent translation components in Equation (2) are to be determined. Generally speaking, at least five pairs of matching points can be used to determine E [20]. However, considering the nonlinearity of the essential matrix E, eight pairs of matching points are required for estimating the extrinsic parameters [21]. During calculation, R and t are generally determined by performing singular value decomposition (SVD) on E. Since the SVD algorithm is sensitive to errors introduced in E, a general practice is to involve a large number of correspondence pairs to improve the robustness [22]. This causes uncertainty and potentially influences the measurement.

Since the rotation matrix actually contains circular functions about the relative orientation between the two camera axes, determination of the relative angles between the dual cameras becomes an alternative for constructing the rotation matrix. If the rotation matrix between cameras is estimated by means of sensors, the translation vector components are calculated by solving a set of linear equations. Thus, the initial values of extrinsic parameters are obtained with finite correspondence pairs, and Bundle Adjustment (BA) is followed to refine the extrinsic parameters. Based on such an innovation, the composition of the measurement unit is shown in Figure 2. The benefit of such routine is obvious. Firstly, it is not necessary to find hundreds of correspondence pairs between left and right images. Secondly, it improves the robustness of camera calibration by setting the initial parameters with the aid of sensors. In particular, the initial parameters are usually accurate enough with high precision sensors, and this method is applicable if the number of correspondence pairs is limited.



Figure 2. The composition of the measurement unit with a camera, an inclinometer, and a range sensor.

Feng et al. used an inertial measurement unit (IMU) which was fixed on the camera to obtain the relative angles to aid camera calibration [23]. However, since the acceleration and attitude angle obtained through the IMU are based on the inertial reference system, the angle output is deviated from the influence of the earth's rotation and the magnetic field. In order to eliminate this influence, the attitude angle is obtained with a differential algorithm to improve the accuracy, which leads to additional operations in stereo camera calibration. In this study, the inclinometer, which is based on the "physical pendulum" structure, is adopted to directly output the pitch and roll angles relative to the horizontal plane without any movement.

In a general imaging unit, the angle between the camera optical axis and the horizontal plane is defined as the pitch angle. The angle between the optical axis and North in the horizontal plane is defined as the attitude angle. Additionally, the roll angle is defined as the motion of a camera about its optical axis. In order to obtain the information about the camera pose, an inclinometer is attached and aligned with the camera, as shown Figure 2.

This layout has been successfully applied in monocular video deflectometers [24], although the attitude angle is disabled. While in stereo vision, besides the pitch and roll angles, the relative attitude angle is essential to construct the rotation matrix R. In a common condition, the stereo imaging station is placed on the ground with a distance between each other facing the object, as shown in Figure 3. In the vertical view, the optical axes of the dual imaging stations intercept at position Q. Thus, a triangle is established by projecting the optical centers  $O_{cl}$ ,  $O_{cr}$ , and Q in the horizontal plane. The relative attitude angle is actually identical to the angle  $\beta$  in the triangle, and it can be determined if all side lengths in the triangle are known.



**Figure 3.** Relative attitude angle calculation: (**a**) using laser rangefinder and inclinometer; (**b**) using GNSS.

There are generally two ways to measure the side length in this triangle. One is to use a laser rangefinder, which utilizes a ruby laser in combination with an optical telescope to aim a laser beam and a photomultiplier for detecting the laser beam reflected from the target with precision in centimeters. Considering the stereo imaging stations are displaced dozens even hundreds of meters in outdoor applications, this method is accurate enough for the side length measurement. Note that two imaging stations and the object usually do not remain in one horizontal plane. Fortunately, the variations in pitch angles are obtained with the inclinometer attached to each imaging unit. Corrections are required by outputting the readouts from the laser rangefinder as well as the inclinometer. In this condition, the side lengths in the horizontal plane are calculated based on a simple circular function, as shown in Figure 3a. The other way is to use global navigation satellite system (GNSS) in calculating the side lengths [25]. With the use of Real-time kinematic (RTK) positioning, it provides up to centimeter-level accuracy by measuring the phase of the signal's carrier wave in addition to the information content of the signal and relying on a single reference station. By placing a GNSS antenna at the left and right imaging stations, and the object in turn, the precise longitudes, latitudes and altitude of O<sub>cl</sub>, O<sub>cr</sub>, and Q are acquired. Thus, the accurate positions of the apex of a horizontal triangle are determined from the readouts of the GNSS signals by ignoring the altitude. As shown in Figure 3b, the side lengths of the horizontal triangle are obtained directly from calculating the distance among the apex with known longitudes and latitudes. Additionally, the relative attitude angle  $\beta$  is calculated according to Equation (4).

$$\beta = \cos^{-1} \left( l_l^2 + l_r^2 - l^2 \right) / \left( 2 \times l_l \times l_r \right)$$
(4)

In the meantime, the scale is known by determining the distance between two imaging stations by using either the rangefinder or GNSS techniques. The accuracy of an inclinometer is about 0.01°, and it provides the pitch and roll information of the camera in this layout. However, since the angular outputs by the inclinometer are based on the horizontal plane, they need to be corrected before they are used to calculate the relative rotation angle. Based on the regulation of the coordinate system defined in Figure 2, the geometric relationship shown in Figure 4 can be drawn,  $\gamma$  and  $\alpha'$  are the angles in  $x_i$  and  $z_i$  axis, respectively, output by the inclinometer relative to the horizontal plane. Regarding  $\gamma$  as the rotation angle of axis  $x_i$  rotating around the axis  $z'_i$ , the angle  $\alpha'$  should be corrected to  $\alpha$  for ensuring the continuity of rotation. To estimate the corrected counterpart  $\alpha$ , the right triangle a, spanned by the vector  $x_i$  with the related inclination angle  $\gamma$ , is translated to plane b to make its vertical edge is aligned with that of the right triangle spanned by axis  $z_i$  with angle of  $\alpha'$ . According to trigonometric cosine in space geometry, the corrected angle  $\alpha$  can be derived from  $\alpha'$ , as shown in Equation (5).



Figure 4. Geometric relationship between the measured angle and rotation angle.

In the above equation,  $\alpha'$  and  $\gamma$  are the pitch and roll angles directly outputted from the inclinometer. Follow such operation, the Euler angles of each camera, i.e.,  $(\alpha, \beta, \gamma)$  are determined.

Since the pitch and roll angles of each camera, and the relative attitude angle between the cameras have been obtained, the rotation matrix *R* can be obtained by multiplying the three rotation matrices around the *z-y-x* axes, as shown in Figure 5. It follows a certain rotation order: firstly from the right camera coordinate system to the right horizontal coordinate system  $\mathbf{R}_{hr-cr}^{-1}$ , secondly to the left horizontal coordinate system  $\mathbf{R}_{hr-hl}$ , and lastly to the left camera coordinate system  $\mathbf{R}_{hl-cl}$ , as expressed in Equation (6).

$$R = R_{hr-cr}^{-1} R_{hr-hl} R_{hl-cl}$$

$$R_{hr-cr} = R(\gamma_r) R(\alpha_r), \quad R_{hr-hl} = R(\beta), \quad R_{hl-cl} = R(\gamma_l) R(\alpha_l)$$
(6)

(5)

(7)





Figure 5. Conversion relationship between coordinate systems.

In the meantime, since a set of equations can be established based on the epipolar constraint relationship, the essential matrix is known as shown in Equation (2). With the established rotation matrix between the cameras, the translation vector components  $(t_1, t_2, t_3)$  are calculated by solving a set of linear equations. After establishing the initial values for the extrinsic parameters, nonlinear bundle adjustment (BA) optimization is performed using limited correspondence pairs of image feature points. In the BA optimization, the world coordinates of a constructed point can be expressed based on the relationship between the intrinsic parameters and the depth of the left camera, as written as Equation (8).

$$\begin{bmatrix} \frac{u_i - c_x}{f_x} d & \frac{v_i - c_y}{f_y} d & d \end{bmatrix}^T$$
(8)

where  $u_i$  and  $v_i$  are the image coordinates;  $c_x$  and  $c_y$  are the principal point coordinates;  $f_x$  and  $f_y$  are the equivalent focal lengths; and d is the depth of the left camera. Consider that the large depth brings considerable errors in outdoor large FOVs and long working distance applications. In order to improve the stability and accuracy of BA, an inverse depth parameter expression is used, as written in Equation (9).

$$\begin{bmatrix} \frac{u_i - c_x}{w f_x} & \frac{v_i - c_y}{w f_y} & \frac{1}{w} \end{bmatrix}^T$$
(9)

where *w* is defined as 1/d.

### 3. Experiments and Results

In Section 2, a sensor-aided method for camera calibration has been proposed. After the rotation matrix is estimated according to the relative pose between the stereo cameras, the translation vector is calculated from solving a set of linear equations based on the epipolar constraint relationship. Since this solution differs from existing calibration methods, propagation of errors and accuracy are major concerns in applications. In order to verify the reliability of the proposed method, simulation and real experiments were carried out.

### 3.1. Simulation

Sensors such as inclinometer, rangefinders, and GNSS, were used to estimate the rotation matrix in the proposed method. The influence of errors induced in the sensors was analyzed via simulation. The intrinsic parameters of a stereo-camera system are assumed to be known. In detail, the image resolution is assumed as  $5000 \times 5000$  pixels; the lens focal length is 16 mm; and the pixel size is  $1.85 \mu$ m. In a simulated arrangement, the pitch and roll angles of the left camera are  $1^{\circ}$  and  $-1^{\circ}$ , respectively. Additionally, the pitch and roll angles of the right camera are  $1^{\circ}$  and  $-1^{\circ}$ , respectively. The baseline distance is assumed as 1030.82 mm. Then, the ground truth of the extrinsic parameters is directly calculated according to Equation (6), by determining the rotation vector as  $(-0.52^{\circ}, 29.99^{\circ}, -0.52^{\circ})$ , and translation vector as (-1000 mm, -10 mm, 250 mm). The FOV is thus formed as  $600 \text{ mm} \times 600 \text{ mm}$ , in which a virtual ruler with a length of 40.349 mm is used as a benchmark to evaluate the measurement accuracy.

During simulation, errors were introduced to the inclinometer readouts and were divided into 20 groups ranging from 0.05° to 1° with a step of 0.05°, and the attitude angle errors were divided into 20 groups ranging from 0.025° to 0.5° with a step of 0.025°. In each group, image feature points were randomly generated with given errors. One thousand simulations were carried out in each condition, and the mean error was recorded to evaluate the performance.

The influence of the angle error in the inclinometer on the measurements is shown in Figure 6. It is indicated that for the general placement stereo vision measurement system, the measured length error remains at a low level if the errors of the inclinometer readouts remain within 1°. This requirement is satisfied in most cases, even considering the misalignment between the camera and the inclinometer. In contrast, the measured length error increased with the attitude angle error, which is considered to be a key error source in the proposed calibration method. Since the initial attitude angle is obtained by measuring the side length in the horizontal triangle by means of laser rangefinder or RTK GNSS according to Equation (4), the accuracy of length measurement is accurate enough as the side length often ranges from dozens to hundreds of meters in outdoor applications.



Figure 6. Influence of angle input error: (a) inclinometer error; (b) attitude error.

In the simulation, the number of feature points in BA optimization and the corresponding distribution were also studied. The number of feature points was increased from 10 to 100 with a step of 5, and the distribution of feature points was analyzed with the proportion in the image from 10% to 100%. During the simulation, the errors induced from the inclinometer were assumed as  $\pm 0.1^{\circ}$ , errors in the attitude angle were assumed as  $\pm 0.1^{\circ}$ . The positioning errors of the feature points obey the Gaussian distribution of (0, 0.5), and 50 pairs of feature points were utilized. The distribution of the feature points was also studied. The feature points were assumed to be randomly distributed in a square region centered at the image center. The proportion, which was defined as the area divided by the image size, was assumed from 10% to 100% with a step of 5%. A set of examples of randomly distributed feature points is provided in Figure 7.



**Figure 7.** Examples of random feature point locations on the left camera image (the proportion of the image is 20%, 40%, 60%, and 80%).

As shown in Figure 8a, increasing the number of feature points was helpful in reducing measurement errors. However, when the number exceeded 40, the errors reached a plateau indicating that it is not necessary to involve many feature points in camera calibration. Meanwhile the evaluation of feature point distribution in Figure 8b indicates that if the proportion in the image was greater than 40%, camera calibration becomes steady. These findings demonstrate that the proposed method is not sensitive to errors induced by the sensors, and is not required to use many feature points which is beneficial for fast and accurate camera calibration.



**Figure 8.** The influence of feature point factors: (**a**) the number of feature points; (**b**) the proportion of the FOV occupied by the feature points.

#### 3.2. Experimental Validation

Two experiments were conducted to validate the proposed method. The first was conducted under laboratory conditions in order to compare the proposed method with the classic Zhang's calibration method. The second was conducted in outdoor conditions on a wind turbine blade model.

# 3.2.1. Experimental Procedure

The layout of the indoor experimental setup is shown in Figure 9. A stereo vision system which was composed of two imaging units was set up to measure the corner-featured target movement driven by a micrometer with a FOV of 600 mm  $\times$  450 mm. Each unit included a digital monochromatic camera with spatial resolution of 4024  $\times$  3036 pixels, and an inclinometer with its axis aligned with the camera's optical axis. In such a configuration, the pitch and roll angles of a camera are determined by readout from the aligned inclinometer. The distance between the two units was measured as 0.895 m by means of a laser rangefinder. The working distances of the left and right cameras were measured as 1.757 m and 1.773 m, respectively. The relative attitude angle of the system was determined based on Equation (4) in the horizontal triangle with known side lengths. Before displacement measurements, camera calibration was conducted with the classic Zhang's calibration method using precise chessboard and the proposed method in order.



Figure 9. Indoor experimental setup.

In Zhang's calibration, a checkerboard composed of  $11 \times 8$  squares with side lengths of 40 mm was placed in the FOV with different poses. In total, thirty image pairs were collected, and the image coordinates of these corner points were extracted for construction of the intrinsic and extrinsic parameters of the stereo vision system.

In the proposed calibration, the intrinsic parameters of the cameras were predetermined according to the results from Zhang's calibration. The rotation matrix was constructed based on Equation (6) with the known roll, pitch, and attitude angles of the camera. Since the size of the square corner target was 50 mm, it provided scale information which was useful in determining the translation vector. In addition, specially designed features were also placed in the field, and the corresponding images were simultaneously captured with the stereo imaging system. An example was presented in Figure 10. A total of one hundred pairs of features were collected for extraction of the image coordinates of the feature points by means of the Shi–Tomasi corner detection algorithm [26]. By solving a series of linear equations, the extrinsic parameters of the stereo vision system were determined after generation equations based on the epipolar relationship.



Figure 10. Corner targets in the FOV.

Two groups of calibration parameters were first used to reconstruct the 3D position of the chessboard which was used in Zhang's calibration at an arbitrary pose as shown in Figure 11. By extracting the corresponding image coordinates of the grids, the world coordinates of all cross points in the chessboard were reconstructed. Additionally, the side length of grid was accordingly estimated.



Figure 11. Checkerboard calibration plate in the FOV.

A translation experiment was conducted to evaluate the proposed method. The square corner target, which was mounted on the precision translation stage, was controlled to move 17.5 mm with a step of 3.5 mm in the x, y, and z directions. The movement was measured with the stereo vision system by means of the calibration parameters obtained by Zhang's method and the proposed method

## 3.2.2. Experimental Results

The intrinsic parameters of the stereo vision system obtained by Zhang's calibration are presented in Table 1. In the meantime, the extrinsic parameters of the system were also estimated, as shown in Table 2. In contrast, the pitch and roll angles of the left camera were obtained as  $0.867^{\circ}$  and  $-0.843^{\circ}$ , respectively. Additionally, the pitch and roll angles of the right camera were obtained as  $0.641^{\circ}$  and  $0.088^{\circ}$ , respectively. The attitude angle was calculated as  $29.36^{\circ}$  with known side lengths in the horizontal triangle. Based on the proposed method, the rotation matrix between the cameras was constructed directly from these output angles. Additionally, the translation vector was determined by solving a series of linear equations which was generated with 50 pairs of matched points with the scale

factor. The extrinsic parameters were also obtained by means of the rotation and translation vectors, and are listed in Table 2.

Intrinsic Parameters	Left Camera	<b>Right Camera</b>
$f_x$ (pixels)	8701.27	8715.25
$f_{y}$ (pixels)	8697.10	8713.49
$u_0$ (pixels)	2003.92	1970.53
$v_0$ (pixels)	1447.46	1477.75
$\overline{k_1}$	-0.091	-0.153
$k_2$	-0.006	2.694

Table 1. Intrinsic parameters of the stereo vision system.

Table 2. Extrinsic	parameters of	the stereo	vision system	resulting	from the	e two methods.

<b>Extrinsic Parameters</b>	Proposed Method	Zhang's Method
Rotation vector (°)	(-0.46, 29.36, -0.37)	(-0.27, 29.32, -0.56)
Translation vector (mm)	(-857.10, -10.63, 255.27)	(-854.80, -9.49, 259.81)

Reprojection errors were estimated with the two groups of calibration parameters. It was found they were at the same level with 0.12 pixels for Zhang's calibration and 0.11 pixels for the proposed method.

The world coordinates of all cross points in the chessboard were plotted with the two groups of calibration parameters, as shown in Figure 12a. Figure 12b uses the image coordinates of all corner points on the left image in Figure 11 to show the average error distribution of the distance reconstructions of the side length connected to each corner point. Considering all checkerboard distances, the average errors by Zhang's calibration and the proposed method were 0.012 mm and 0.014 mm, respectively. The root mean square errors were 0.015 mm and 0.014 mm, respectively. It was demonstrated that the sensor-aided calibration method for estimating the extrinsic parameters of the stereo vision system is reliable and the accuracy is at the same level as Zhang's method.



**Figure 12.** Reconstruction of the checkerboard: (**a**) reconstruction of the chessboard grid; (**b**) error distribution of the chessboard grid.

In the translation experiment, the displacement of the feature point was measured by means of the two groups of calibration parameters at every motion step. As shown in Table 3, the measured displacement components with the two groups of calibration parameters were consistent with errors less than 0.02 mm.

Trans	lation	Proposed	Method	Zhang's l	Method
Motion	Direction	Measured	Error	Measured	Error
	Х	3.509	0.009	3.514	0.014
3.500	Y	3.511	0.011	3.515	0.015
	Z	3.512	0.012	3.492	-0.008
	Х	6.989	-0.011	7.018	0.018
7.000	Y	7.013	0.013	7.011	0.011
	Z	7.016	0.016	6.992	-0.008
	Х	10.508	0.008	10.507	0.007
10.500	Y	10.488	-0.012	10.506	0.006
	Z	10.492	-0.008	10.505	0.005
	Х	14.014	0.014	14.013	0.013
14.000	Y	13.085	-0.015	14.009	0.009
	Z	13.987	-0.013	14.011	0.011
	Х	17.509	0.009	17.013	0.013
17.500	Y	17.017	0.017	16.085	-0.015
	Z	17.508	0.008	17.515	0.015
Mear	n error	0.01	12	0.01	11

Table 3. Measured displacement of the feature point (unit in mm).

#### 3.3. Outdoor Large FOV Experiment

# 3.3.1. Experimental Setup

An outdoor experiment was conducted on the rotating wind turbine blades, as shown in Figure 13. Three blades with a length of 2.15 m were driven by a stepping motor. The stereo vision system was placed in front of the rotating wind turbine blades. The camera with spatial resolution of  $5120 \times 5120$  pixels was employed for image collection. The field of view was about  $5 \times 5$  m when the lens with a 70 mm focus length was used. An inclinometer was aligned with the optical axis of the camera in order to identify the pose information. With the use of RTK-GNSS sensors, the baseline between two imaging stations was measured as 4.21 m, and the distances from the turbine center to the left and right cameras in the horizontal plane were 15.21 m and 15.46 m, respectively. During the experiment, the wind turbine was rotating at 7.5 rpm, and synchronized images were acquired at 20 fps for about one minute with the developed stereo vision system.

In camera calibration, the typical two-step routine was carried out. In detail, the intrinsic parameters of the cameras were predetermined by means of a precise chessboard, as shown in Table 4. As the speckles on the blade surface were used as feature points, they were evenly distributed in the image field during rotation. In total, 282,539 pairs of features points were successfully extracted from the sequential image pairs. Camera calibration was then conducted in two ways. As a traditional method, all matched features were involved in establishing equations based on the epipolar constraint relationship. The rotation matrix and translation vector were obtained from the resulting essential matrix by means of singular value decomposition (SVD) [27]. In the proposed method, about 100 features points were randomly selected from the feature collections for camera calibration, followed by BA with an inverse depth parameter expression.



Figure 13. Outdoor experimental setup.

Table 4. Intrinsic	parameters of	the stereo vision	system [	predetermined	with Zhang	s's method.
					(	,

<b>Intrinsic Parameters</b>	Left Camera	<b>Right Camera</b>
$f_x$ (pixels)	15,953.42	16,067.16
$f_{y}$ (pixels)	15,948.73	16,060.78
$u_0$ (pixels)	2613.21	2471.86
$v_0$ (pixels)	2515.96	2478.96
$\overline{k_1}$	0.042	0.027
<i>k</i> <sub>2</sub>	-0.681	-0.3817

To verify the accuracy of the measurement system, the delicate calibration chessboard (Figure 11) was placed at different locations in the measurement field and the length of the grids was evaluated using the resulting calibration parameters by means of the two methods. As shown in Figure 14a, the chessboard was placed multiple times in a horizontal line and the corresponding images pairs were acquired by the stereo vision system simultaneously. The overall grid length was evaluated with two groups of calibration parameters. Then, the 3D motion of the selected features on a blade was evaluated, as shown in Figure 14b.



Figure 14. (a) Placement of calibration board; (b) selected features on the blade.

#### 3.3.2. Experimental Results

In the experimental setup, the initial angular information was directly provided by the inclinometers which were aligned with the cameras. In detail, the pitch and roll angles of the left camera were given as  $13.244^{\circ}$  and  $-0.379^{\circ}$ , respectively, while the pitch and roll angles of the right camera were given as  $13.465^{\circ}$  and  $1.646^{\circ}$ , respectively. The corrected pitch angles were derived according to Equation (5). According to Equation (4), the relative attitude angle was determined as  $16.24^{\circ}$  in the horizontal plane. Thus, the rotation matrix between the cameras was determined. By introducing limited feature points, the translation vector components were obtained by solving a series of linear equations. The extrinsic parameters are listed in Table 5, and compared with the traditional calibration method.

Table 5. Extrinsic parameters obtained with the proposed and traditional methods.

Extrinsic Parameters	Proposed Method	Traditional Method
Rotation vector (°)	(-1.31, 16.55, -1.12)	(-1.22, 16.64, -1.11)
Translation vector (mm)	(-4145.09, 83.75, 718.65)	(-4145.54, 106.12, 726.15)

The reprojection errors were evaluated as 0.32 pixels for the proposed method and 0.35 pixels for the traditional method, which suggested that the proposed method is effective with equivalent accuracy to the traditional method.

The side length of one hundred and fifty seven grids on the calibration board was estimated at each location shown in Figure 14a, with the camera parameters determined by the proposed method. It was found that the average side length was measured as  $40 \pm 0.092$  mm, which showed good agreement with the benchmark.

As indicated in Figure 14b, eight feature points along the radial direction of the blade were selected. The image correlation algorithm which considered rotation and translation was employed [28]. The 3D trajectory and the corresponding components in the world coordinate system aligned with the left camera were shown in Figure 15.



**Figure 15.** 3D motion of the blade: (**a**) 3D trajectory; (**b**) displacement in x direction; (**c**) displacement in y direction; (**d**) displacement in z direction.

# 4. Discussion

In this study, a method for estimation of the relative extrinsic parameters has been proposed. The merits of this method include that the rotation matrix is directly determined by means of sensors, and the translation vector is estimated by solving a series of linear equations satisfying the epipolar constraint relationship. This is especially suitable for outdoor measurements with large FOVs in conditions with limited feature points. With the proposed procedure, camera calibration lasted only for minutes with fair accuracy. It provides a convenient and robust way to achieve accurate 3D displacement measurement for engineering applications.

Although the proposed method is easy to apply, two key points must be considered in determining the rotation matrix between stereo vision stations. The first is to accurately determine the Euler angles between cameras. In the proposed method, the optical axis of the camera is required to be aligned with the inclinometer. A precise tuning device is expected to provide an accurate adjustment in integration. An inclinometer usually provides a pair of pitch and roll angles. However, the output pitch angle should be corrected according to Equation (5) based on the geometric relationship between the output angle of the inclinometer and the rotation regulation. The second is to determine the

relative attitude angles between cameras. In a conventional stereo vision system, the two imaging units are placed towards the target in a triangular arrangement. The relative attitude angle is defined as the vertex angle at the target in the horizontal plane shown in Figure 3. Typically, two convenient methods are helpful in determining this angle. The first is to use the global navigation satellite system (GNSS) in determining the position information of the target, as well as the left and right cameras [25]. Since it directly provides coordinates of the triangle in the horizontal plane with accuracy in centimeters, it provides an accurate estimation of the relative attitude angle. The other is to use an optical range finder to measure the distance between the stereo imaging stations, and the distances between the target and the individual cameras. Note that the accuracy of the distance measurement using the range finder is in decimeters. Errors introduced in determining the side lengths in the triangle are greater than those while using GNSS. In addition, since the imaging units and the target might not in a horizontal plane in a general situation, the pitch angle of the range finder must be considered. From this point of view, GNSS is suggested in determining the relative attitude angle in the proposed method.

### 5. Conclusions

In this paper, a convenient and robust camera calibration method for estimating the extrinsic parameters of the stereo vision system has been proposed, which is assisted by an inclinometer and a range sensor for outdoor large FOV measurements. In brief, the rotation matrix is estimated by determining the Euler angles between the cameras. Additionally, the translation vector components are reduced from solving a series of equations established based on the epipolar constraint relationship with finite matched feature pairs. The nonlinear optimization of inverse depth parameterization is used as an effective method if feature points are insufficient. The influences of output errors from the inclinometer and the relative attitude angle, and the feature number and distribution of image features on 3D reconstruction were numerically analyzed. Comparative experimental results have verified the reliability and robustness of the proposed method. The proposed method has been applied to the measurement of a rotary blade, and the 3D motion components were accurately quantified.

**Author Contributions:** Conceptualization, D.Z. and Q.Y.; methodology, Z.S.; software, J.W.; validation, J.W., B.G. and Y.H.; formal analysis, B.G.; investigation, J.W.; resources, Y.H.; data curation, Z.S.; writing—original draft preparation, J.W.; writing—review and editing, D.Z.; visualization, Z.S.; supervision, D.Z.; project administration, D.Z.; funding acquisition, D.Z. and Q.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (NSFC), grant numbers 12072184, 12002197, and 11727804.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

#### References

- 1. Zhang, S.; Mao, S.; Arola, D.; Zhang, D. Characterization of the strain-life fatigue properties of thin sheet metal using an optical extensiometer. *Opt. Lasers Eng.* **2014**, *60*, 44–48. [CrossRef]
- Su, Z.; Pan, J.; Zhang, S.; Wu, S.; Yu, Q.; Zhang, D. Characterizing dynamic deformation of marine propeller blades with stroboscopic stereo digital image correlation. *Mech. Syst. Signal Process.* 2022, 162, 108072. [CrossRef]
- 3. Brown, D. Close-Range Camera Calibration. Photogramm. Eng. 2002, 37, 855–866.
- 4. Luhmann, T.; Fraser, C.; Maas, H.-G. Sensor modelling and camera calibration for close-range photogrammetry. *ISPRS J. Photogramm. Remote Sens.* **2016**, *115*, 37–46. [CrossRef]
- 5. Tsai, R. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE J. Robot. Autom.* **1987**, *3*, 323–344. [CrossRef]
- 6. Zhang, Z. A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. 2000, 22, 1330–1334. [CrossRef]
- Faugeras, O.D.; Luong, Q.T.; Maybank, S.J. Camera self-calibration: Theory and experiments. In Proceedings of the Computer Vision—ECCV'92, Berlin/Heidelberg, Germany, 19–22 May 1992; pp. 321–334.

- Brückner, M.; Bajramovic, F.; Denzler, J. Intrinsic and extrinsic active self-calibration of multi-camera systems. *Mach. Vis. Appl.* 2014, 25, 389–403. [CrossRef]
- Yamazaki, S.; Mochimaru, M.; Kanade, T. Simultaneous self-calibration of a projector and a camera using structured light. In Proceedings of the CVPR 2011 WORKSHOPS, Colorado Springs, CO, USA, 20–25 June 2011; pp. 60–67.
- Li, C.; Su, R. A Novel Stratified Self-calibration Method of Camera Based on Rotation Movement. J. Softw. 2014, 9, 1281–1287. [CrossRef]
- 11. Miyata, S.; Saito, H.; Takahashi, K.; Mikami, D.; Isogawa, M.; Kojima, A. Extrinsic Camera Calibration Without Visible Corresponding Points Using Omnidirectional Cameras. *IEEE Trans. Circuits Syst. Video Technol.* **2018**, *28*, 2210–2219. [CrossRef]
- 12. Gao, Z.; Gao, Y.; Su, Y.; Liu, Y.; Fang, Z.; Wang, Y.; Zhang, Q. Stereo camera calibration for large field of view digital image correlation using zoom lens. *Measurement* **2021**, *185*, 109999. [CrossRef]
- 13. Sun, J.; Liu, Q.; Liu, Z.; Zhang, G. A calibration method for stereo vision sensor with large FOV based on 1D targets. *Opt. Lasers Eng.* **2011**, *49*, 1245–1250. [CrossRef]
- Wang, Y.; Wang, X. An improved two-point calibration method for stereo vision with rotating cameras in large FOV. J. Mod. Opt. 2019, 66, 1106–1115. [CrossRef]
- Liu, Z.; Li, F.; Li, X.; Zhang, G. A novel and accurate calibration method for cameras with large field of view using combined small targets. *Measurement* 2015, 64, 1–16. [CrossRef]
- 16. Zhang, Y.; Liu, W.; Wang, F.; Lu, Y.; Wang, W.; Yang, F.; Jia, Z. Improved separated-parameter calibration method for binocular vision measurements with a large field of view. *Opt. Express* **2020**, *28*, 2956–2974. [CrossRef]
- 17. Gorjup, D.; Slavič, J.; Boltežar, M. Frequency domain triangulation for full-field 3D operating-deflection-shape identification. *Mech. Syst. Signal Process.* **2019**, 133, 106287. [CrossRef]
- 18. Poozesh, P.; Sabato, A.; Sarrafi, A.; Niezrecki, C.; Avitabile, P.; Yarala, R. Multicamera measurement system to evaluate the dynamic response of utility-scale wind turbine blades. *Wind Energy* 2020, 23, 1619–1639. [CrossRef]
- 19. Jiang, T.; Cui, H.; Cheng, X. A calibration strategy for vision-guided robot assembly system of large cabin. *Measurement* **2020**, *163*, 107991. [CrossRef]
- Nister, D. An efficient solution to the five-point relative pose problem. *IEEE Trans. Pattern Anal. Mach. Intell.* 2004, 26, 756–770. [CrossRef] [PubMed]
- 21. Hartley, R.I. In defense of the eight-point algorithm. IEEE Trans. Pattern Anal. Mach. Intell. 1997, 19, 580–593. [CrossRef]
- Fathian, K.; Gans, N.R. A new approach for solving the Five-Point Relative Pose Problem for vision-based estimation and control. In Proceedings of the 2014 American Control Conference, Portland, OR, USA, 2014, 4–6 June 2014; pp. 103–109.
- Feng, W.; Su, Z.; Han, Y.; Liu, H.; Yu, Q.; Liu, S.; Zhang, D. Inertial measurement unit aided extrinsic parameters calibration for stereo vision systems. Opt. Lasers Eng. 2020, 134, 106252. [CrossRef]
- 24. D'Alfonso, L.; Garone, E.; Muraca, P.; Pugliese, P. On the use of the inclinometers in the PnP problem. In Proceedings of the 2013 European Control Conference (ECC), Zurich, Switzerland, 17–19 July 2013; pp. 4112–4117.
- Zhang, D.; Yu, Z.; Xu, Y.; Ding, L.; Ding, H.; Yu, Q.; Su, Z. GNSS Aided Long-Range 3D Displacement Sensing for High-Rise Structures with Two Non-Overlapping Cameras. *Remote Sens.* 2022, 14, 379. [CrossRef]
- Shi, J.; Tomasi, C. Good features to track. In Proceedings of the 1994 IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 21–23 June 1994; pp. 593–600.
- Feng, W.; Zhang, S.; Liu, H.; Yu, Q.; Wu, S.; Zhang, D. Unmanned aerial vehicle-aided stereo camera calibration for outdoor applications. *Opt. Eng.* 2020, 59, 014110. [CrossRef]
- Yang, D.; Su, Z.; Shuiqiang, Z.; Zhang, D. Real-time Matching Strategy for RotaryObjects using Digital Image Correlation. *Appl. Opt.* 2020, 59, 6648–6657. [CrossRef] [PubMed]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.