



## Article

# Camera Self-Calibration with GNSS Constrained Bundle Adjustment for Weakly Structured Long Corridor UAV Images

Wei Huang<sup>1</sup> , San Jiang<sup>2,\*</sup> and Wanshou Jiang<sup>1,3</sup>

<sup>1</sup> State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China; hw1006@whu.edu.cn (W.H.); jws@whu.edu.cn (W.J.)

<sup>2</sup> School of Computer Science, China University of Geosciences, Wuhan 430074, China

<sup>3</sup> Collaborative Innovation Center of Geospatial Technology, Wuhan University, 129 Luoyu Road, Wuhan 430079, China

\* Correspondence: jiangsan@cug.edu.cn

**Abstract:** Camera self-calibration determines the precision and robustness of AT (aerial triangulation) for UAV (unmanned aerial vehicle) images. The UAV images collected from long transmission line corridors are critical configurations, which may lead to the “bowl effect” with camera self-calibration. To solve such problems, traditional methods rely on more than three GCPs (ground control points), while this study designs a new self-calibration method with only one GCP. First, existing camera distortion models are grouped into two categories, i.e., physical and mathematical models, and their mathematical formulas are exploited in detail. Second, within an incremental SfM (Structure from Motion) framework, a camera self-calibration method is designed, which combines the strategies for initializing camera distortion parameters and fusing high-precision GNSS (Global Navigation Satellite System) observations. The former is achieved by using an iterative optimization algorithm that progressively optimizes camera parameters; the latter is implemented through inequality constrained BA (bundle adjustment). Finally, by using four UAV datasets collected from two sites with two data acquisition modes, the proposed algorithm is comprehensively analyzed and verified, and the experimental results demonstrate that the proposed method can dramatically alleviate the “bowl effect” of self-calibration for weakly structured long corridor UAV images, and the horizontal and vertical accuracy can reach 0.04 m and 0.05 m, respectively, when using one GCP. In addition, compared with open-source and commercial software, the proposed method achieves competitive or better performance.

**Keywords:** digital photogrammetry; camera self-calibration; Brown model; polynomial model; aerial triangulation



**Citation:** Huang, W.; Jiang, S.; Jiang, W. Camera Self-Calibration with GNSS Constrained Bundle Adjustment for Weakly Structured Long Corridor UAV Images. *Remote Sens.* **2021**, *13*, 4222. <https://doi.org/10.3390/rs13214222>

Academic Editors: Francesco Nex and Marco Piras

Received: 14 September 2021

Accepted: 18 October 2021

Published: 21 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

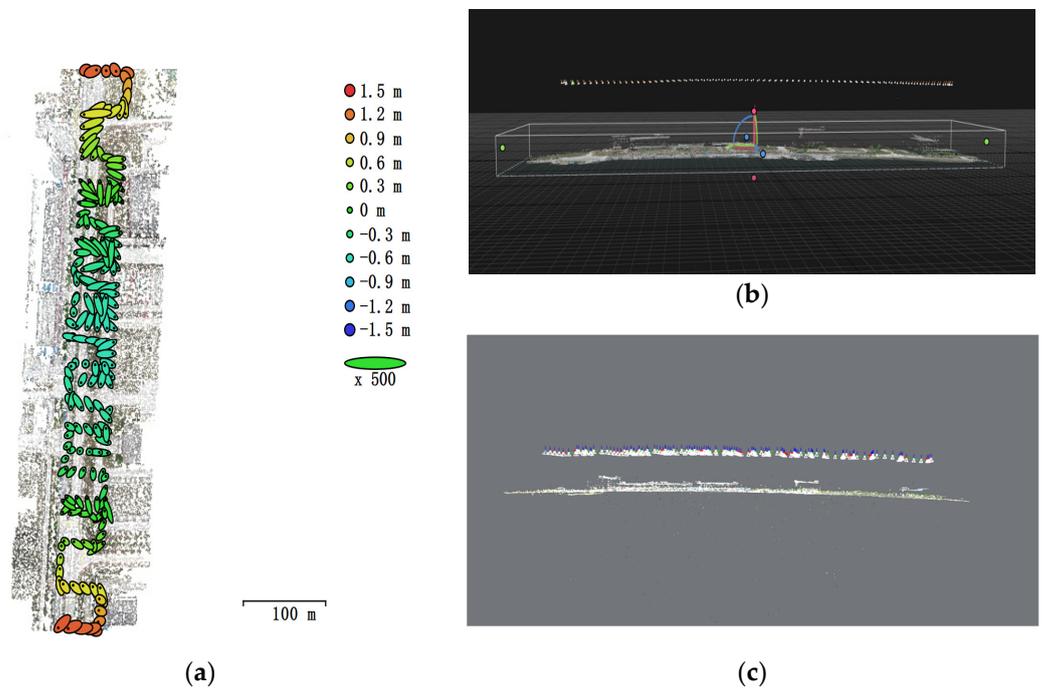


**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With the advantages of flexible data acquisition and ease of use, UAV has become one of the most important remote sensing platforms for the photogrammetry and remote sensing community [1]. The UAVs have the characteristics of small size, autonomous vertical take-off and landing with low site requirements, high flight safety performance, and the flexibility to adjust the direction of flight, making them widely used for the regular inspection of high-voltage transmission line corridors [2–6]. However, the UAV platforms are often equipped with consumer-grade, non-metric digital cameras, mainly due to the limitations of the platform's load capacity. These cameras have non-ignored lens distortions when compared with metric sensors, which influences the robustness and precision of AT. The long corridor structure of UAV images is a critical configuration, and the reconstructed model would be bending with the inaccurately estimated distortion parameters. For example, Figure 1 illustrates the “bowl effect” of self-calibration with commercial software and open-source software for long corridor UAV images because of the failure in camera

distortion parameter estimation. Thus, it is very critical to accurately estimate the distortion parameters of cameras for high-precision AT.



**Figure 1.** The “bowl effect” of self-calibration BA for long corridor UAV images. (a) The result of commercial software Agisoft Photoscan V1.4.1, Z error is represented by the ellipse color, and X, Y errors are represented by the ellipse shape. (b) The result of commercial software Reality Capture V1.2, and (c) the result of open-source software AliceVision.

In photogrammetry and computer vision fields, the purpose of camera self-calibration is to solve the internal orientation parameters and camera lens distortion parameters, which affect the precision of 3D reconstruction. At present, the existing camera calibration methods can be categorized into two groups: pre-calibration and self-calibration. The pre-calibration method usually depends on either the indoor calibration board with fixed patterns [7,8] or the outer door large-scale 3D calibration test field [9]. However, this kind of method has many disadvantages applied in UAV image calibration. On the one hand, the indoor calibration errors of the calibration board with special patterns would be expanded with the increase of UAV flying height, and the focal length of the camera is usually set to be fixed depending on the flying height of UAVs, which is not suitable for indoor camera pre-calibration; on the other hand, although the outer door 3D calibration test field could improve the precision of camera calibration, it requires a lot of manpower and consumes time. Compared with the drawbacks of pre-calibration, the self-calibration process is simpler and more convenient without any known calibration targets.

With the above considerations, the relevant scholars have performed in-depth research on the camera self-calibration methods with different camera distortion models, including the physical model and mathematical model. Among the camera distortion models, the Brown model [10] and its improved model [11] are the most classical physical models. However, it is a challenge for camera self-calibration to occur in the critical configuration, due to the high correlation between the distortion parameters of the Brown model [12]. In the field of computer vision, the division model is another kind of physical model commonly used [13], which can fit simple camera distortion. Recently, many researchers have combined the division model with the fundamental matrix or the essential matrix to solve the camera distortion parameters by establishing polynomial equations [14–17]. However, it cannot fit complex camera distortion and is not suitable for UAV camera self-calibration. The physical camera model cannot describe the complex distortion precisely and may not

work when the pattern of camera distortion is not apparent and the precise knowledge on distortion is unavailable. Based on this consideration, the mathematical model tries to use function approximation theory to accurately fit complex camera distortion, such as the quadratic orthogonal polynomial [18] and quartic orthogonal polynomial model [19]. Tang et al. [20,21] proposed the orthogonal polynomial models based on Legendre and Fourier polynomials and applied the models in an aerial camera for self-calibration. Subsequently, Babapour et al. [22] presented the Chebyshev–Fourier and Jacobi–Fourier camera models, which significantly improved the horizontal and elevation accuracy of aerial images. However, few studies have applied the Legendre- and Fourier-based distortion models to camera self-calibration for long corridor UAV images. This is the first key research content of this paper.

For the camera self-calibration with a long corridor structure, the related research can be divided into three categories: the research on the theoretic analysis [23,24], the research on the strategies of self-calibration [25,26], and the accuracy verification with such structures [27–32]. Wu et al. [23] analyzed the motion field of images with radial distortion and proved the ambiguous reconstruction with the “bowl effect” of camera self-calibration under weak structures and configuration through mathematical theory. Zhou et al. [24] discussed the impact of the focal length parameter estimation of camera self-calibration with a flat, corridor configuration. The research on the theoretic analysis of camera self-calibration with a long corridor structure has focused on investigating the causes and influencing factors, but has not presented any solutions to solve the problems. For the long corridor structure, Tournadre et al. [25] presented a 7th-order polynomial combined with radial camera distortion model (F15P7) and verified the accuracy of the orientations with a weak configuration using ground control points (GCP). Although this method can alleviate the “bowl effect”, it relied on more than three GCPs for absolute image orientation. Polic et al. [26] proposed an uncertainty-based camera model selection method to reduce the “bowl effect”, but this method did not consider the newest mathematical-based distortion models and high-precision GNSS observations. Griffiths et al. [27] analyzed the accuracy of 3D reconstruction from long corridor structure UAV images in detail, and experiments show that the more complex distortion model can improve the accuracy of camera self-calibration. The related works of [28–32] are mainly focused on accessing the accuracy of the DSM (Digital Surface Model), DTM (Digital Terrain Model), the influence of the distribution of GCPs, and on giving suggestions for data collection without improvement of the strategies for camera self-calibration. Compared to the existing literature about self-calibration with a long corridor structure, the proposed paper extends the scope of research on camera distortion models and investigates the accuracy of the recently proposed orthogonal polynomial model with the strategies for camera parameter initialization and high-precision GNSS fusion in a long corridor structure.

For the traditional aerial photogrammetry in surveying and mapping, the UAV platform generally collects image data with the regular region and often has multiple parallel and overlapping stable structures. However, in the application of UAV inspection for power lines, only rectangle or S-shaped strip flight trajectory is adopted to collect image data, due to cost considerations. Since the constraints between the long corridor structure are reduced to the minimum, the correlation of camera intrinsic parameters and external parameters cannot be restricted with the stability structure of images, which leads to the “bowl effect” phenomenon and affects the relative and absolute accuracy of 3D reconstruction. At present, most UAVs are equipped with centimeter-level high-precision differential GNSS, which can provide better initial position parameters to constrain the camera projection centers [33,34]. Traditional technology for fusing high-precision GNSS locations and oriented images of SfM is to minimize a weighted sum of image and GNSS errors. However, when the structure of images is degenerated and unstable, the oriented images bend after camera self-calibration. In this situation, the traditional technology of fusing GNSS locations and SfM would not align the projection centers of the image to the GNSS locations, which cannot eliminate the “bowl effect”. How to use the high-precision

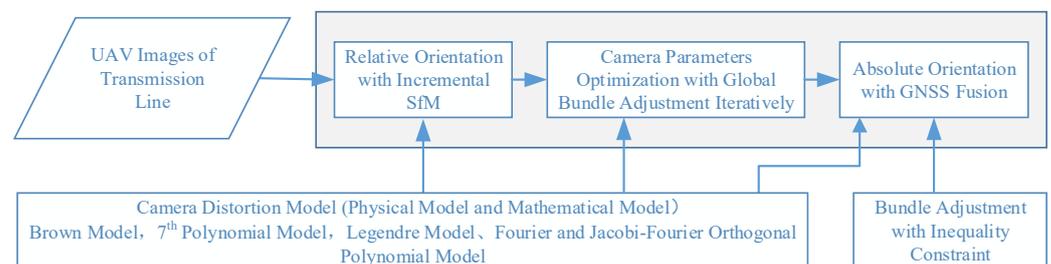
GNSS information to significantly alleviate the “bowl effect” is the second key research content of this paper.

To overcome the problem, this paper firstly investigates the classical physical model and orthogonal polynomial model for camera self-calibration. Then, a new strategy combined with the parameter initialization of UAV images and high-precision GNSS observation fusion is proposed for camera self-calibration with the physical model and orthogonal polynomial model. Finally, four UAV image datasets are used in the experiments for camera self-calibration, which illustrates the feasibility of this strategy. Compared with the related works of long corridor camera self-calibration, the proposed method has the following contributions: (1) the accuracy of the newest mathematical camera distortion models is investigated and verified, which can achieve better accuracy in the vertical direction; (2) a new strategy of camera self-calibration for long corridor UAV images is proposed, which can alleviate the “bowl effect” with the high-precision GNSS locations for direct georeferencing; (3) compared with the traditional method, which needs more than three GCPs to solve the problem of “bowl effect”, the proposed method achieves competitive accuracy with only one GCP constraint, which is meaningful for the UAV photogrammetric community.

This paper is organized as follows. Section 2 presents the camera distortion model and the proposed camera self-calibration method with the camera parameters’ initialization and high-precision GNSS fusion in detail. In Section 3, UAV datasets and experimental results are presented and discussed. Section 4 concludes the results of this study and presents future work.

## 2. Methodology

The proposed method mainly studies camera self-calibration in the incremental SfM framework, which is used to accurately estimate the image orientation and camera intrinsic parameters. Firstly, the most commonly used camera distortion models are analyzed in Section 2.1, including physical models and mathematical models. Secondly, the BA with inequality constraint is investigated in detail in Section 2.2. Finally, the camera self-calibration strategy is introduced for the long corridor structure of UAV images in Section 2.3. Figure 2 describes our main research contents. The camera distortion model introduced in Section 2.1 is applied in the proposed camera self-calibration method, and the bundle adjustment with inequality constraint described in Section 2.2 is applied in the absolute orientation of the proposed camera self-calibration method.



**Figure 2.** The camera self-calibration workflow for long corridor UAV images.

### 2.1. Camera Distortion Model

The classical Brown camera model in the field of photogrammetry and computer vision, the 7th polynomial model, and the Legendre, Fourier, and Jacobi–Fourier orthogonal polynomial models are compared and analyzed. The mathematical form of each camera distortion model is described below.

#### 2.1.1. Brown Model

The Brown distortion model [10] has been widely used in the photogrammetry and computer vision field. The parameters of the Brown model mainly include symmetrical

radial distortion, tangential distortion, and in-plane distortion [35,36]. The mathematical equations are as shown below.

$$\begin{aligned}\Delta x &= \bar{x}(k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1(r^2 + 2\bar{x}^2) + 2p_2\bar{x}\bar{y} + b_1\bar{x} + b_2\bar{y} \\ \Delta y &= \bar{y}(k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1(r^2 + 2\bar{y}^2) + 2p_2\bar{x}\bar{y} \\ r &= \sqrt{(x - x_0)^2 + (y - y_0)^2} = \sqrt{\bar{x}^2 + \bar{y}^2}\end{aligned}\quad (1)$$

where  $x_0, y_0$  is the principal point of the image,  $k_1, k_2, k_3$  are the radial distortion coefficients,  $p_1, p_2$  are the tangential distortion coefficients, and  $b_1, b_2$  are the in-plane distortion coefficients, which are named as the affinity and the shear terms. The in-plane coefficients represent different scaling factors and non-orthogonal pixels in the image along the x- and y-axis.

### 2.1.2. 7th Polynomial Model

The 7th polynomial camera distortion model is provided by the open-source software MicMac [25], and it contains a total of 66 parameters. The distortion functions are as follows:

$$\begin{aligned}\Delta x &= a_0\bar{x} + a_1\bar{y} - 2a_2\bar{x}^2 + a_3\bar{x}\bar{y} + a_4\bar{y}^2 + \\ &a_6\bar{x}^3 + a_7\bar{x}^2\bar{y} + a_8\bar{x}\bar{y}^2 + a_9\bar{y}^3 + \\ &a_{14}\bar{x}^4 + a_{15}\bar{x}^3\bar{y} + a_{16}\bar{x}^2\bar{y}^2 + a_{17}\bar{x}\bar{y}^3 + a_{18}\bar{y}^4 + \\ &a_{24}\bar{x}^5 + a_{25}\bar{x}^4\bar{y} + a_{26}\bar{x}^3\bar{y}^2 + a_{27}\bar{x}^2\bar{y}^3 + a_{28}\bar{x}\bar{y}^4 + a_{29}\bar{y}^5 + \\ &a_{36}\bar{x}^6 + a_{37}\bar{x}^5\bar{y} + a_{38}\bar{x}^4\bar{y}^2 + a_{39}\bar{x}^3\bar{y}^3 + a_{40}\bar{x}^2\bar{y}^4 + a_{41}\bar{x}\bar{y}^5 + a_{42}\bar{y}^6 + \\ &a_{50}\bar{x}^7 + a_{51}\bar{x}^6\bar{y} + a_{52}\bar{x}^5\bar{y}^2 + a_{53}\bar{x}^4\bar{y}^3 + a_{54}\bar{x}^3\bar{y}^4 + a_{55}\bar{x}^2\bar{y}^5 + a_{56}\bar{x}\bar{y}^6 + a_{57}\bar{y}^7 \\ \Delta y &= -a_0\bar{y} + a_1\bar{x} + a_2\bar{x}\bar{y} - 2a_3\bar{y}^2 + a_5\bar{x}^2 + \\ &a_{10}\bar{x}^3 + a_{11}\bar{x}^2\bar{y} + a_{12}\bar{x}\bar{y}^2 + a_{13}\bar{y}^3 + \\ &a_{19}\bar{x}^4 + a_{20}\bar{x}^3\bar{y} + a_{21}\bar{x}^2\bar{y}^2 + a_{22}\bar{x}\bar{y}^3 + a_{23}\bar{y}^4 + \\ &a_{30}\bar{x}^5 + a_{31}\bar{x}^4\bar{y} + a_{32}\bar{x}^3\bar{y}^2 + a_{33}\bar{x}^2\bar{y}^3 + a_{34}\bar{x}\bar{y}^4 + a_{35}\bar{y}^5 + \\ &a_{43}\bar{x}^6 + a_{44}\bar{x}^5\bar{y} + a_{45}\bar{x}^4\bar{y}^2 + a_{46}\bar{x}^3\bar{y}^3 + a_{47}\bar{x}^2\bar{y}^4 + a_{48}\bar{x}\bar{y}^5 + a_{49}\bar{y}^6 + \\ &a_{58}\bar{x}^7 + a_{59}\bar{x}^6\bar{y} + a_{60}\bar{x}^5\bar{y}^2 + a_{61}\bar{x}^4\bar{y}^3 + a_{62}\bar{x}^3\bar{y}^4 + a_{63}\bar{x}^2\bar{y}^5 + a_{64}\bar{x}\bar{y}^6 + a_{65}\bar{y}^7\end{aligned}\quad (2)$$

where  $\bar{x}, \bar{y}$  are the same as the definition in the Brown model, and  $a_0, a_1, \dots, a_{65}$  are the coefficients of the 7th polynomial. Six coefficients are eliminated to reduce the correlation between those coefficients.

### 2.1.3. Legendre Model

The Legendre model is composed of a series of orthogonal polynomials, which greatly reduce the correlation between distortion coefficients. This model has been used in the camera self-calibration for professional digital mapping cameras. However, the feasibility of a consumer-grade camera applied in camera self-calibration needs to be verified. Therefore, the Legendre orthogonal polynomial model with 66 parameters is introduced in this paper. The mathematical expression is presented in Formula (3), where  $p_{m,n} = 0^{-6}l_m(\bar{x})l_n(\bar{y})$ ,  $m, n \in [0, 5]$ ;  $\bar{x} = \frac{x-w/2}{w}$ ,  $\bar{y} = \frac{y-h/2}{h}$ ;  $x, y$  are the pixel coordinates in the image; and  $w$  and  $h$  represent the width and height of the image, respectively.  $l_m(\bar{x})$  and  $l_n(\bar{y})$  mean the Legendre polynomials and  $a_0, a_1, \dots, a_{65}$  are the coefficients. Similar to the 7th polynomial distortion model, six coefficients are eliminated in the Legendre model.

$$\begin{aligned}
\Delta x &= a_0p_{1,0} + a_1p_{0,1} + a_2p_{2,0} + a_3p_{1,1} + a_4p_{0,2} + a_5p_{3,0} + a_6p_{2,1} + \\
& a_7p_{1,2} + a_8p_{0,3} + a_9p_{4,0} + a_{10}p_{3,1} + a_{11}p_{2,2} + a_{12}p_{1,3} + a_{13}p_{0,4} + \\
& a_{14}p_{5,0} + a_{15}p_{4,1} + a_{16}p_{3,2} + a_{17}p_{2,3} + a_{18}p_{1,4} + a_{19}p_{0,5} + a_{20}p_{5,1} + \\
& a_{21}p_{4,2} + a_{22}p_{3,3} + a_{23}p_{2,4} + a_{24}p_{1,5} + a_{25}p_{5,2} + a_{26}p_{5,3} + a_{27}p_{3,4} + \\
& a_{28}p_{2,5} + a_{29}p_{5,3} + a_{30}p_{4,4} + a_{31}p_{3,5} + a_{32}p_{5,4} + a_{33}p_{4,5} + a_{34}p_{5,5} \\
\Delta y &= a_1p_{1,0} - a_0p_{0,1} + a_{35}p_{2,0} - a_2p_{1,1} - a_3p_{0,2} + a_{36}p_{3,0} + a_{37}p_{2,1} + \\
& a_{38}p_{1,2} + a_{39}p_{0,3} + a_{40}p_{4,0} + a_{41}p_{3,1} + a_{42}p_{2,2} + a_{43}p_{1,3} + a_{44}p_{0,4} + \\
& a_{45}p_{5,0} + a_{46}p_{4,1} + a_{47}p_{3,2} + a_{48}p_{2,3} + a_{49}p_{1,4} + a_{50}p_{0,5} + a_{51}p_{5,1} + \\
& a_{52}p_{4,2} + a_{53}p_{3,3} + a_{54}p_{2,4} + a_{55}p_{1,5} + a_{56}p_{5,2} + a_{57}p_{5,3} + a_{58}p_{3,4} + \\
& a_{59}p_{2,5} + a_{60}p_{5,3} + a_{61}p_{4,4} + a_{62}p_{3,5} + a_{63}p_{5,4} + a_{64}p_{4,5} + a_{65}p_{5,5}
\end{aligned} \tag{3}$$

#### 2.1.4. Fourier Model

The mathematical formula of 16 parameters in the first-order orthogonal polynomial distortion model based on the two-dimensional Fourier series [21] is as follows:

$$\begin{aligned}
\Delta x_f &= a_0c_{1,0} + a_1c_{0,1} + a_2c_{1,-1} + a_3c_{1,1} + a_4s_{1,0} + a_5s_{0,1} + a_6s_{1,-1} + a_7s_{1,1} \\
\Delta y_f &= a_8c_{1,0} + a_9c_{0,1} + a_{10}c_{1,-1} + a_{11}c_{1,1} + a_{12}s_{1,0} + a_{13}s_{0,1} + a_{14}s_{1,-1} + a_{15}s_{1,1}
\end{aligned} \tag{4}$$

where  $c_{m,n} = 10^{-6} \cos(m\bar{x}_f + n\bar{y}_f)$ ,  $s_{m,n} = 10^{-6} \sin(m\bar{x}_f + n\bar{y}_f)$ ,  $\bar{x}_f = \frac{x-w/2}{w}\pi$ , and  $\bar{y}_f = \frac{y-h/2}{h}\pi$ ;  $x, y$  are the pixel coordinates in the image;  $w$  and  $h$  represent the width and height of the image, respectively; and  $a_0, a_1, \dots, a_{15}$  are the coefficients. When there is significant radial distortion in the image, it needs to be employed together with the radial distortion model. Therefore, the radial distortion and quadratic polynomial are applied.

$$\begin{aligned}
\Delta x_{rg} &= \bar{x}_r(k_1r^2 + k_2r^4 + k_3r^6) + \\
& b_0\bar{x}_g + b_1\bar{y}_g - 2b_2\bar{x}_g^2 + b_3\bar{x}_g\bar{y}_g + b_4\bar{y}_g^2 \\
\Delta y_{rg} &= \bar{y}_r(k_1r^2 + k_2r^4 + k_3r^6) - \\
& b_0\bar{y}_g + b_1\bar{x}_g + b_2\bar{x}_g\bar{y}_g - 2b_3\bar{y}_g^2 + b_5\bar{x}_g^2
\end{aligned} \tag{5}$$

where  $\bar{x}_r, \bar{y}_r$  are consistent with  $\bar{x}, \bar{y}$  in the Brown model;  $\bar{x}_g = \bar{x}_f/\pi$ ,  $\bar{y}_g = \bar{y}_f/\pi$ ;  $k_1, k_2, k_3$  are the radial coefficients; and  $b_0, b_1, \dots, b_5$  are the coefficients of a quadratic polynomial. The final hybrid distortion model is as follows:

$$\begin{aligned}
\Delta x &= \Delta x_f + \Delta x_{rg} \\
\Delta y &= \Delta y_f + \Delta y_{rg}
\end{aligned} \tag{6}$$

#### 2.1.5. Jacobi–Fourier Model

Compared with the Fourier model, the Jacobi–Fourier model has a higher horizontal and vertical accuracy [22]. In this paper, the Jacobi–Fourier model is adopted, and the function is presented in Formula (7), where  $J_n(\alpha, \beta, r)$  is the Jacobi polynomial and the mathematical expression is as shown in Formula (8);  $\bar{x}, \bar{y} \in [0, 1]$ , represented as normalized image coordinates;  $r$  is the distance from the normalized pixel coordinate to the origin,  $r^2 = \bar{x}^2 + \bar{y}^2$ ;  $N_J, M_F, N_F$  are the variable parameters of Jacobi and Fourier, respectively; and  $a_{i,m,n}, b_{i,m,n}, a'_{i,m,n}, b'_{i,m,n}$  are the coefficients of the polynomial.

$$\begin{aligned}
\Delta x_{jf} &= \sum_{i=0}^{N_J} \sum_{m=0}^{M_F} \sum_{n=1}^{N_F} a_{i,m,n} J_i(\alpha, \beta, r) \sin(m\pi\bar{x} + n\pi\bar{y}) \\
& + \sum_{i=0}^{N_J} \sum_{m=0}^{M_F} \sum_{n=1}^{N_F} b_{i,m,n} J_i(\alpha, \beta, r) \cos(m\pi\bar{x} + n\pi\bar{y}) \\
\Delta y_{jf} &= \sum_{i=0}^{N_J} \sum_{m=0}^{M_F} \sum_{n=1}^{N_F} a'_{i,m,n} J_i(\alpha, \beta, r) \sin(m\pi\bar{x} + n\pi\bar{y}) \\
& + \sum_{i=0}^{N_J} \sum_{m=0}^{M_F} \sum_{n=1}^{N_F} b'_{i,m,n} J_i(\alpha, \beta, r) \cos(m\pi\bar{x} + n\pi\bar{y})
\end{aligned} \tag{7}$$

$$\begin{aligned}
 J_n(\alpha, \beta, \tau) &= \sqrt{\frac{\omega(\alpha, \beta, \tau)}{b_n(\alpha, \beta) \cdot \tau}} G_n(\alpha, \beta, \tau) \\
 G_n(\alpha, \beta, \tau) &= \frac{n!(\beta-1)!}{(\alpha+n-1)!} \sum_{s=0}^n (-1)^s \frac{(\alpha+n+s-1)!}{(n-s)!s!(\beta+s-1)!} \tau^s \\
 b_n(\alpha, \beta) &= \frac{n![(\beta-1)!]^2(\alpha-\beta+n)!}{(\beta+n-1)!(\alpha+n-1)!(\alpha+2n)} \\
 \omega(\alpha, \beta, \tau) &= (1-\tau)^{\alpha-\beta} \tau^{\beta-1}
 \end{aligned} \tag{8}$$

In Formula (8),  $\alpha, \beta$  are set to 7 and 3, respectively, according to the suggestion of [22];  $\tau \in [0, 1]$ , and  $G_n, b_n, \omega$  are the polynomial, normalizing constant, and weighting function, respectively. Similar to the Fourier model, the Jacobi–Fourier model is mixed with radial distortion and quadratic polynomial as given in the following.

$$\begin{aligned}
 \Delta x &= \Delta x_{jf} + \Delta x_{rg} \\
 \Delta y &= \Delta y_{jf} + \Delta y_{rg}
 \end{aligned} \tag{9}$$

## 2.2. GNSS-Aided Bundle Adjustment with Inequality Constraint

GNSS-aided BA is the commonly used method in the photogrammetry and computer vision fields. By minimizing the reprojection error, the traditional BA can optimize the internal and external parameters of the camera together with the 3D coordinates of the tie points. The error equation of GNSS-aided BA considers the deviation between the image projection center  $X_c$  and GNSS phase center  $X_{gps}$ . The jointly optimized error function is as described in Formula (10), where  $w$  is the weight of GNSS.

$$e_u = \sum_j \rho_j \left( \left\| \pi(P_c, X_k) - x_j \right\|_2^2 \right) + \sum_n \rho_n \left( \left\| w(X_c - X_{gps}) \right\|_2^2 \right) \tag{10}$$

Different from the traditional weighted GNSS-aided BA, Maxime et al. [33] proposed an inequality constrained bundle adjustment (IBA) method with GNSS fusion to reduce the deviation error accumulation between the image projection center  $X_c$  and the GNSS location  $X_{gps}$  for long image sequences. The basic idea of IBA is to improve the absolute accuracy with GNSS-aided BA on the premise of appropriately increasing the reprojection error. Let  $X^* = (X_c^T, X_a^T, X_k^T)$  be the optimal solution of standard BA without consideration of GNSS information, where  $X_c, X_a, X_k$  represent the projection center of image, rotation angle, and 3D coordinates of tie points, respectively. Further, let  $e(X^*)$  be the minimum sum of square reprojection errors with the optimal solution, i.e.,  $\forall X, e(X^*) \leq e(X)$ . Suppose  $e_t$  be a threshold that is slightly larger than the minimum reprojection error  $e(X^*)$ , i.e.,  $e(X^*) < e_t$ . IBA assumes that the GNSS error is bounded and the reprojection error  $e(X)$  of BA with GNSS constraint should be less than  $e_t$ , that is,  $e(X) \leq e_t$ . Under this condition, the optimized image projection center should be as close to the GNSS phase center as possible, i.e.,  $X_c \approx X_{gps}$ .

Let  $X_2 = (X_a, X_k)$ , then the unknowns of BA can be expressed as  $X = (X_c^T, X_2^T)$ . Let  $P = (I, 0)$  be such that  $X_c = PX$ . IBA establishes the optimization equation by combining penalty function and inequality constraint, as shown in Formula (11).

$$e_I(X) = \frac{\gamma}{c_I(X)} + \|PX - X_{gps}\|^2 \tag{11}$$

where  $\gamma$  is a user-defined weight and a positive number,  $c_I(X) = e_t - e(X)$ , and  $c_I(X) > 0$ . The objective function is iteratively minimized by this inequality and penalty function  $\gamma/c_I(X)$  constraint. The penalty function value is close to positive infinity in the neighborhood of  $c_I = 0$ . The algorithm in C style is shown in Algorithm 1 and the parameter  $\gamma$  is set  $\gamma = \frac{e_t - e(X^*)}{10} \|PX^* - X_{gps}\|^2$  according to [33]. More details should be referenced in [33].

**Algorithm 1** The procedure of IBA

---

```

1:  $err = \gamma / (e_t - e(X)) + \|PX - X_{gps}\|^2$ 
2:  $UpdateD = 1; \lambda = 0.001;$ 
3: for ( $It = 0; It < It_{max}; It++$ ) {
4: if ( $UpdateD$ ) {
5:    $UpdateD = 0;$ 
6:    $g = 2J^T E(X); H = 2J^T J;$ 
7:    $g_I = \frac{\gamma}{(e_t - e)^2} g + 2P^T (PX - X_{gps}); H = \frac{\gamma}{(e_t - e)^2} H + 2P^T P;$ 
8:    $\tilde{g} = \sqrt{\frac{2\gamma}{(e_t - e)^3}} g;$ 
9: }
10:  $\tilde{H} = H + \lambda diag(H + \tilde{g}\tilde{g}^T)$ 
11:  $solve\tilde{H}(a, b) = (-g, \tilde{g});$ 
12:  $\Delta = a - \frac{\tilde{g}^T a}{1 + \tilde{g}^T b} b;$ 
13: if ( $e(X + \Delta) \geq e_t$ ) {
14:    $\lambda = 10\lambda;$  continue;
15: }
16:  $err' = \frac{\gamma}{e_t - e(X + \Delta)} + \|P(X + \Delta) - X_{gps}\|^2;$ 
17: if ( $err' < err$ ) {
18:    $X = X + \Delta;$ 
19:   if ( $0.9999err < err'$ ) break;
20:    $err = err'; UpdateD = 1; \lambda = \lambda / 10;$ 
21: } else
22:  $\lambda = 10\lambda;$ 
23:}

```

---

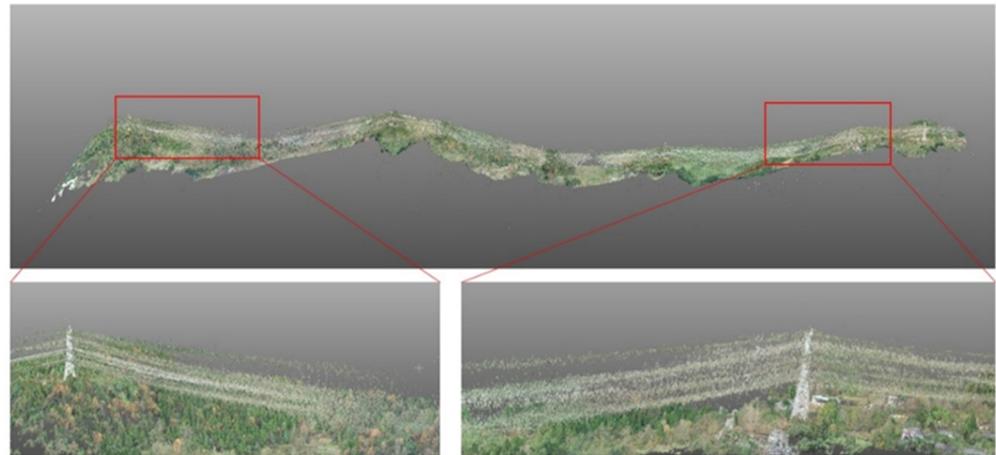
### 2.3. Camera Self-Calibration for Long Corridor UAV Images

The camera self-calibration for long corridor UAV images is realized under the framework of incremental SfM in ColMap. Firstly, the incremental SfM selects two seed images with enough matching feature points, which are of uniform distribution in the images and the intersection angle between the two-image pair should be large enough. Then the relative orientation and 3D coordinates of tie points in the initial image pair are calculated. Secondly, the next best images are selected, which are most fully connected with the existing reconstructed model. The image poses and 3D coordinates of tie points are recovered immediately. Finally, to eliminate accumulated error, the local and global BA optimization is carried out iteratively: (1) when the number of newly added images exceeds a given threshold, local BA optimization is performed for the local-oriented images; (2) when the percentage of registered images grows by a certain threshold, the reconstructed model is optimized by global BA [37]. The feature point extraction, exhaustive matching strategy, criteria for seed image selection, and the strategies of controlling local BA in ColMap are directly used in the proposed method without any changes.

For the long corridor UAV images, the existing incremental SfM framework has the following shortcomings:

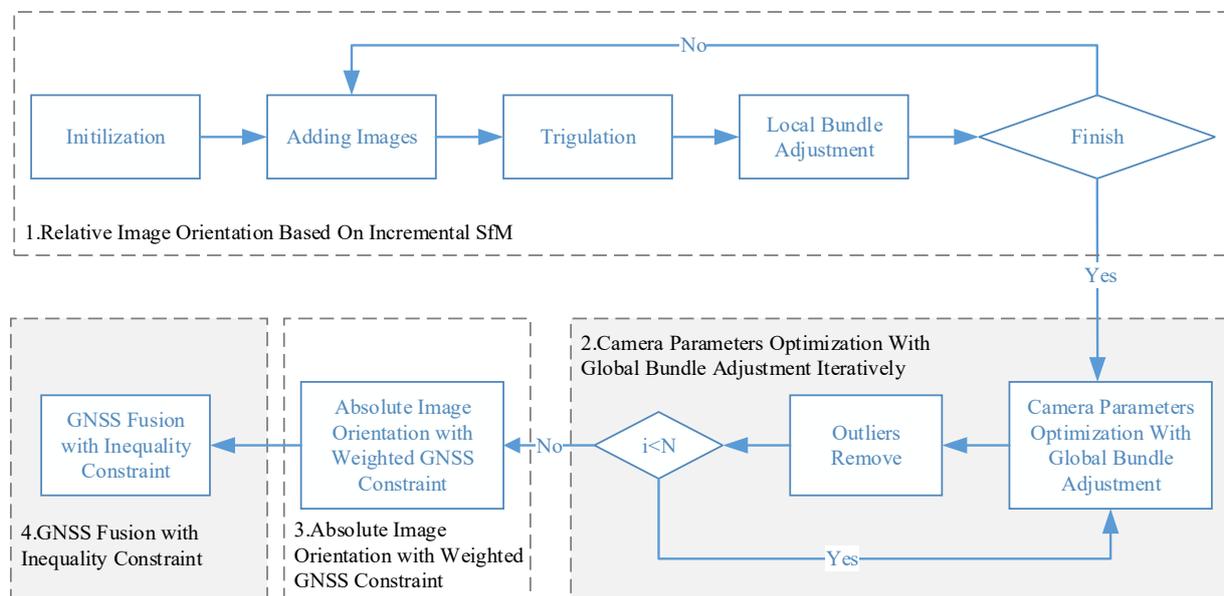
1. From the perspective of camera self-calibration, the next best image selection does not consider whether the scene structure is degraded. If the structure of the seed image is poor and lacks height variation, the camera intrinsic parameters are unstable and may even lead to the failure of the final reconstruction.
2. At present, UAV images often record high-precision GNSS location information, which can alleviate the "bowl effect" of long corridor images. The existing incremental SfM framework of camera self-calibration does not take full advantage of GNSS information for absolute orientation.
3. The inaccurately estimated distortion parameters have an adverse impact on the 3D point clouds generated by dense matching technology. The power lines in UAV

images of high-voltage transmission are usually 1~2 pixels in width. When the distortion parameters are estimated inaccurately, the reconstructed point clouds of power lines are noisy and diverged around. Figure 3 shows the dense point clouds reconstructed by ColMap with inaccurate camera distortion parameters.



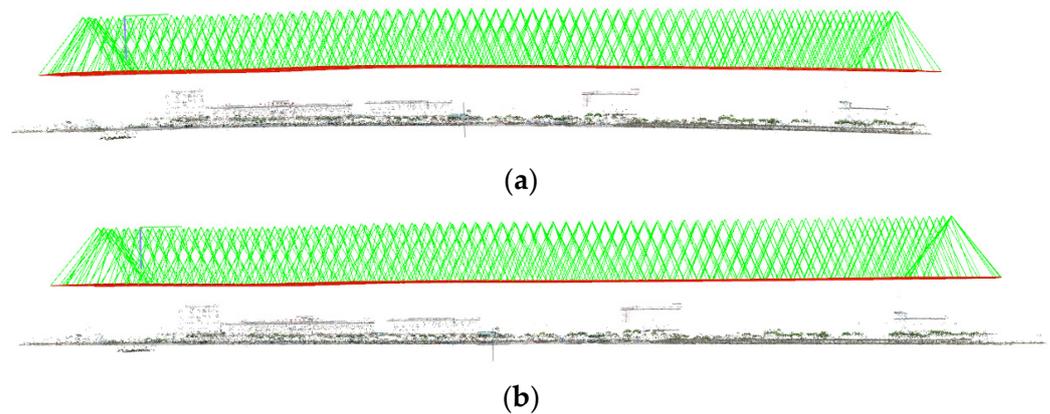
**Figure 3.** The influence of inaccurately estimated distortion parameters with dense matching for power lines.

Therefore, a camera self-calibration method is proposed, which combines the camera parameters initialization and high-precision differential GNSS position information fusion. The workflow is shown in Figure 4, and details of key steps are listed as follows:



**Figure 4.** The camera self-calibration workflow for long corridor UAV images.

1. Image-relative orientation based on incremental SfM framework. Only the local BA is performed to reduce the error accumulation, and the focal length, principal point, and distortion parameters of the image are kept fixed to avoid the problem of the unstable and large variation of distortion parameters and focal length caused by the instability structure of image and scene degradation. The results of image-relative orientation are shown in Figure 5a.



**Figure 5.** The reconstructed models with the proposed camera self-calibration strategy. (a) The result of image relative orientation, and (b) the result of image absolute orientation with the proposed strategy.

2. Global BA and iterative optimization of camera parameters. In the process of iterative global BA and gross error elimination, the optimization strategy of gradually freeing the intrinsic and distortion of camera parameters is adopted, that is, (a) distortion parameters, (b) focal length, and (c) principal point. This strategy can alleviate the correlation between focal length, principal point, and distortion parameters of the camera. According to the experiment, when the number of iterations is bigger than two, the camera intrinsic parameters become stable. In this paper, the global BA is iterated three times. In each iteration, the distortion, focal length, and principal point parameters of the camera are optimized using the strategy of gradually freeing parameters, to provide better initial values for GNSS-constrained BA.
3. Traditional weighted GNSS-constrained absolute orientation. At this time, the GNSS-constrained BA optimizes the focal length, principal point, and distortion parameters as unknowns together. Further, the error equation is shown in Formula (12).

$$e_u = \sum_j \rho_j \left( \left\| \pi(P_c, X_k, \theta) - x_j \right\|_2^2 \right) + \sum_n \rho_n \left( \|w(X_c - X_{gps})\|_2^2 \right) \quad (12)$$

where  $\theta = (f, c_x, c_y, Dist)$ , including the focal length  $f$ , principal point  $c_x, c_y$ , and the distortion parameters  $Dist$ . The distortion parameters  $Dist$  are determined by the selected camera distortion model introduced in Section 2.1. The weight of GNSS  $w$  is set to 10, keeping consistent with [33]. The cost function  $\rho$  is the Cauchy function with stronger noise resistance, as shown in Formula (13).

$$\rho(s) = \log(1 + s) \quad (13)$$

4. GNSS fusion based on IBA. This paper combines IBA to further fuse the GNSS. The main steps are as follows: (a) the camera focal length, principal point, and distortion parameters are used as unknowns to optimize together during camera self-calibration; (b) the initial input parameters  $e(X^*)$  are the sum of the squares' reprojection error of weighted GNSS-constrained BA, and  $X^*$  are the minimum solvers; (c) all the image projection centers and the corresponding GNSS positions are used as constraints for global IBA to solve iteratively. The final reconstructed model with the proposed camera self-calibration strategy is shown in Figure 5b.

In summary, there are several differences between the proposed method with the incremental SfM in ColMap and the usage of IBA in [33]. In ColMap, the local BA is performed on the images that are connected with the most recently registered images, and the global BA is performed according to the growing percentage of the registered images. The camera intrinsic parameters are optimized during local BA and global BA. In the proposed method, only local BA is performed before all images are registered, and

the global BA is iteratively performed after all images are registered. During the local BA, the camera intrinsic parameters are kept fixed. Further, during the global BA, the camera intrinsic parameters are gradually freed to get better initial values. Then the traditional weighted and inequality constrained BA with GNSS is performed for absolute orientation. In [33], the IBA is used in local BA to fuse the low-cost GNSS and image projection centers to refine the k-most recent images. The input-optimized initial parameters are the results of local BA and the 3D GNSS location of the corresponding most recent image. The camera intrinsic parameters are known and kept fixed. In the proposed method, the IBA is used in global BA to fuse high-precision GNSS locations and image projection centers. The camera intrinsic parameters are freed and optimized together with the GNSS constraint. The inputs of IBA in the proposed method are the optimized parameters of the weighted GNSS constraint BA and all the GNSS locations of registered images. The differences between the proposed method and the incremental SfM in ColMap and the IBA in [33] are listed in Tables 1 and 2, respectively.

**Table 1.** The differences between the incremental SfM in ColMap [38] and the proposed method.

Method	Incremental SfM in ColMap	The Proposed Method
Local BA	Camera intrinsic parameters are freed in local BA	Camera intrinsic parameters are fixed in local BA
Global BA	Global BA is performed after growing the registered images by a certain percentage. The camera intrinsic parameters are freed in global BA.	Global BA is performed after registering all images. Iteratively free distortion parameters, focal length, and principal point in global BA.

**Table 2.** The differences between the IBA used in [33] and the proposed method.

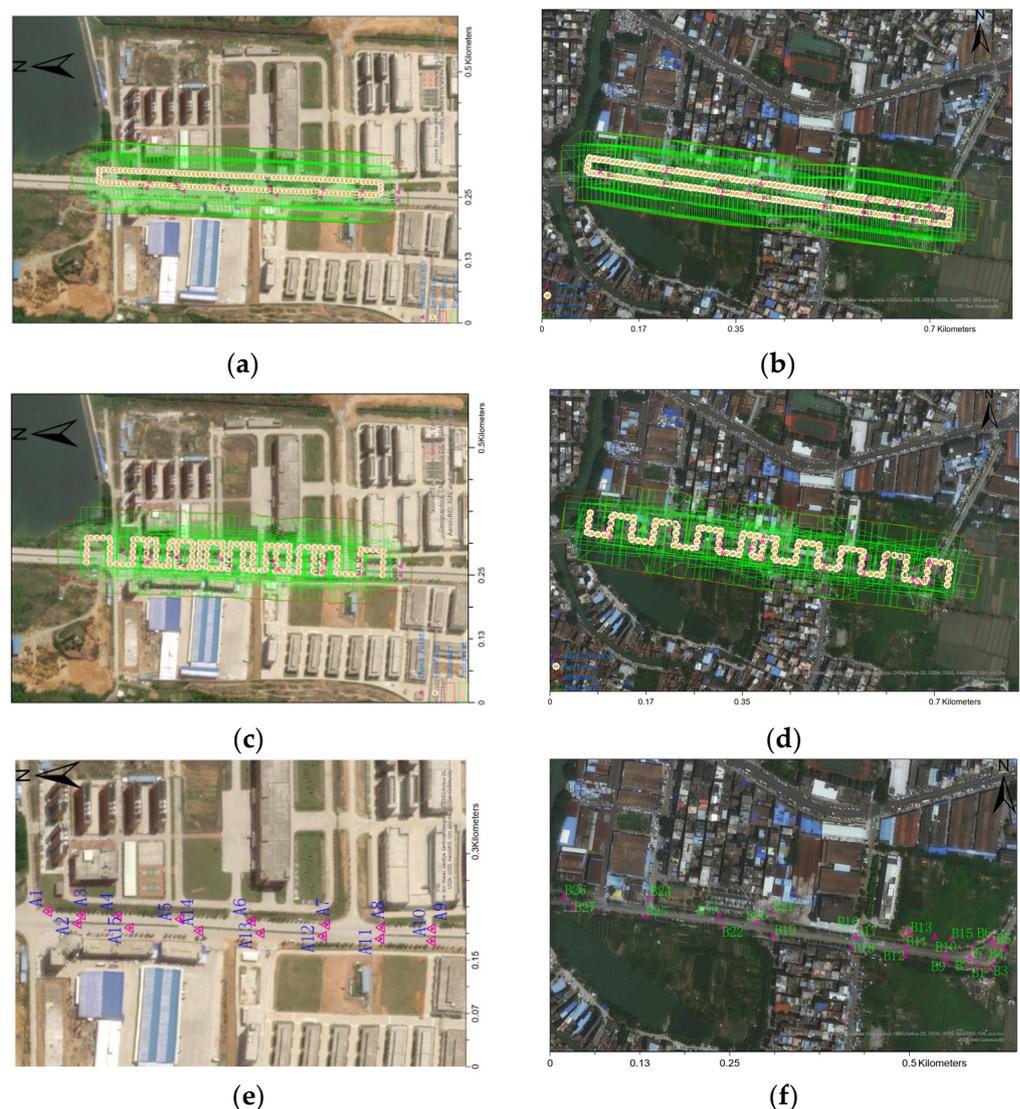
Method	IBA in [33]	The Proposed Method
Applied Stage	IBA is applied in Local BA	IBA is applied in Global BA
Initial Parameters	The input-optimized initial parameters are the results of local BA and the most recent GNSS location.	The input-optimized initial parameters are the results of weighted GNSS constraint BA and all the GNSS locations.
Camera Intrinsic Parameters	Keep fixed	Free

### 3. Results and Discussion

#### 3.1. Test Sites and Datasets

Two datasets of long corridor transmission line UAV images were collected by DJI Phantom 4 RTK UAV, as shown in Figure 6. Figure 6a,b was collected using the rectangle closed-loop trajectory; Figure 6c,d was collected using the S-shaped strip trajectory. The rectangle and S-shaped trajectories were made by a third-party software developed based on DJI Mobile SDK, and during the flight, the standard control algorithm provided by DJI was applied to fly and take photos in autonomous mode along the trajectories. For the rectangle trajectories, the forward and side overlap ratio of images were set to 88% and 75%, respectively, and the flight speed was set to 4 m/s. For the S-shaped trajectories, the forward and side overlap ratio of images were set to 82% and 61%, respectively, and the flight speed was set to 6 m/s. The time interval for taking photos was 3 s for all flights. The camera focal length was kept fixed during image collection. The UAV flight height was set to 70 m, which is relative to the position from where the UAV takes off and the camera takes images vertically downward. The GSD (ground resolution distance) of images was 2.1 cm. The numbers of each image datasets were 140, 166, 165, and 132, respectively. To verify the absolute orientation accuracy of BA, the accurate ground coordinates in the two test sites were collected by Hi-Target iRTK2 GNSS receiver and FindCM CORS. The targets

were marked with red paint manually on the road using the perpendicular lines with a width of about 10 cm, and the ground coordinates inside the right angle were measured. Finally, 15 targets were collected in test site 1 and 27 targets were collected in test site 2. The distribution of targets is listed in Figure 6e,f. The targets in test site 1 were labeled from A1 to A15, and the targets in test site 2 were labeled from B1 to B27. For the experiments of camera self-calibration without GCP constraint, all the targets were regarded as check points to evaluate the accuracy. For the experiments with one GCP constraint, the target of A14 in test site 1 and the target of B20 in test site 2 were regarded as control points and the rest of the targets were regarded as check points for accuracy evaluation. Both A14 and B20 were located in the middle of the long corridors.



**Figure 6.** UAV images of the test sites. (a,c) The rectangle and S-shaped datasets of test site 1, respectively; (b,d) the rectangle and S-shaped datasets of test site 2, respectively. (e,f) the location distribution of GCPs in test site 1 and test site 2.

### 3.2. Analysis of the Influence of Camera Distortion Models

Firstly, the influence of different image acquisition modes and camera distortion models on the accuracy of bundle block adjustment was analyzed. For the hybrid Fourier and Jacobi–Fourier models, the radial and quadratic polynomial distortion parameters were first calculated with all images. Then, we kept these parameters fixed and calculated the parameters of Fourier and Jacobi–Fourier. For other distortion models, all the distortion

parameters were calculated at one time. The mean, SD (standard deviation), and RMSE (root mean square error) were used to evaluate the accuracy of checkpoints.

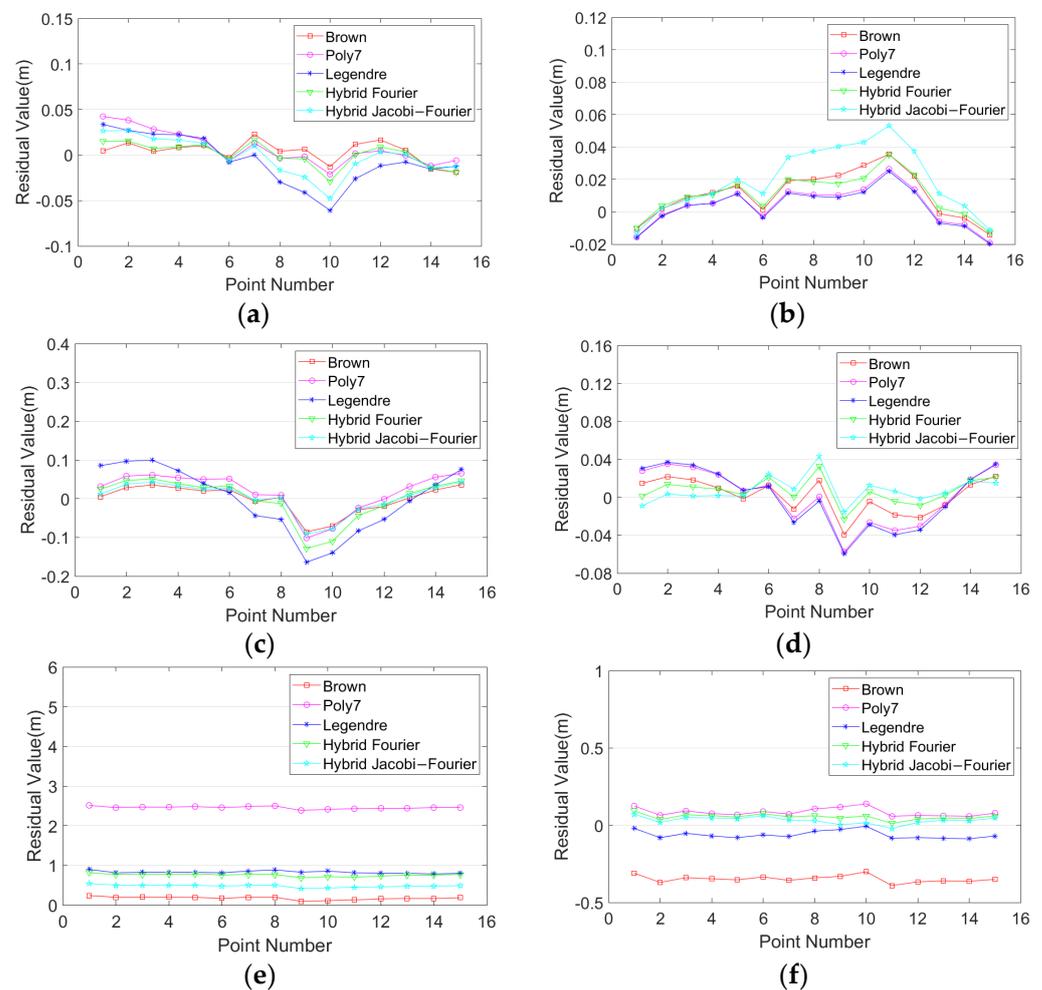
The statistical results are listed in Table 3. Figures 7 and 8 show the residuals of checkpoints for different distortion models. It can be seen that the accuracy of S-shaped image datasets was significantly better than that of rectangle image datasets (except for the vertical accuracy of the Brown model and hybrid Jacobi–Fourier model in test site 2). The main reason is that when the images are collected in the S-shaped method, the angle between the images is always changing, which can reduce the correlation between distortion and other parameters of the camera, and then the horizontal accuracy and vertical accuracy are improved.

**Table 3.** Statistical results of self-calibration for different camera distortion models.

Datasets	Camera Model	Mean(m)			SD(m)			RMSE(m)			
		X	Y	Z	X	Y	Z	X	Y	Z	
Test Site 1	Rectangle	Brown	0.004	−0.001	0.174	0.011	0.036	0.037	0.012	0.036	0.178
		Poly7	0.008	0.018	2.456	0.018	0.050	0.032	0.020	0.053	2.456
		Legendre	−0.006	−0.002	0.829	0.026	0.082	0.032	0.027	0.082	0.830
		Fourier	0.001	0.000	0.755	0.013	0.054	0.032	0.013	0.054	0.755
		Jacobi–Four	−0.001	0.003	0.481	0.020	0.040	0.032	0.020	0.040	0.482
	S-Shaped	Brown	0.011	0.002	−0.348	0.014	0.018	0.022	0.018	0.018	0.349
		Poly7	0.004	0.001	0.084	0.012	0.028	0.025	0.012	0.028	0.088
		Legendre	0.003	−0.000	−0.061	0.012	0.030	0.026	0.012	0.030	0.066
		Fourier	0.010	0.007	0.054	0.013	0.013	0.018	0.016	0.015	0.057
		Jacobi–Four	0.019	0.007	0.031	0.020	0.014	0.022	0.028	0.015	0.038
Test Site 2	Rectangle	Brown	−0.018	0.028	0.753	0.030	0.025	0.036	0.035	0.037	0.754
		Poly7	0.001	0.014	0.240	0.022	0.032	0.028	0.022	0.035	0.242
		Legendre	−0.018	0.027	0.223	0.033	0.027	0.033	0.038	0.038	0.226
		Fourier	0.000	0.004	0.208	0.022	0.035	0.028	0.022	0.035	0.210
		Jacobi–Four	−0.011	0.024	0.104	0.023	0.025	0.029	0.026	0.035	0.108
	S-Shaped	Brown	0.016	−0.006	−1.259	0.012	0.012	0.041	0.020	0.013	1.260
		Poly7	0.004	0.003	0.058	0.016	0.012	0.018	0.017	0.012	0.061
		Legendre	0.003	0.004	0.125	0.017	0.012	0.018	0.017	0.013	0.127
		Fourier	0.005	0.005	0.098	0.014	0.012	0.015	0.013	0.013	0.099
		Jacobi–Four	0.022	0.003	−0.724	0.016	0.012	0.035	0.027	0.013	0.725

For further analysis, we can see the following: (1) For the rectangle dataset of test site 1, the horizontal and vertical accuracy with the Brown model was higher than other distortion models. However, for the other three datasets, the accuracy of the Brown model was the worst. (2) From the comparison of camera self-calibration between the Poly7 model and the Legendre model, it can be seen that the horizontal accuracy of Poly7 was better than the Legendre model for all the datasets. Further, the Poly7 model had better vertical accuracy in the S-shaped dataset of test site 2. However, the Legendre model had better vertical accuracy than the Poly7 model in the other three datasets. The main reason is that the orthogonality of the Legendre model can improve the accuracy of focal length for camera self-calibration, and then improve the vertical accuracy, but meanwhile, it loses the horizontal accuracy. (3) From the comparison of camera self-calibration between the hybrid Fourier model and hybrid Jacobi–Fourier model, it can be seen that the horizontal accuracy of the two models had little difference. However, the vertical accuracy of the

hybrid Jacobi–Fourier model was better than the hybrid Fourier model in general (except for the S-shaped dataset in test site 2).

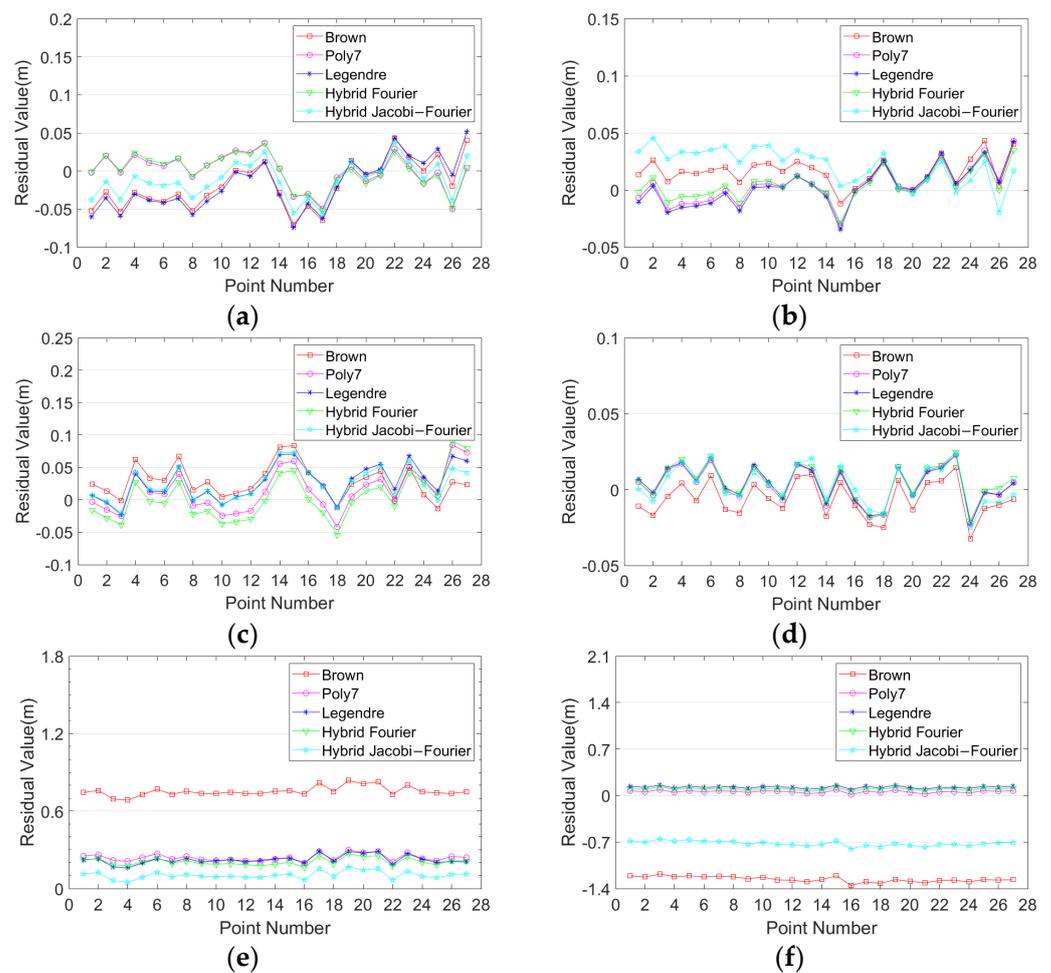


**Figure 7.** The residuals of check points after self-calibration for test site 1. (a,c,e) The residuals of check points of the rectangle dataset in the direction of X, Y, and Z, respectively; (b,d,f) the residuals of check points of the S-shaped dataset in the direction of X, Y, and Z, respectively.

In summary, the bundle block adjustment accuracy in the horizontal direction with the five different models can reach the centimeter level, while the vertical accuracy has great differences for the four datasets. No one distortion model can achieve the best accuracy among all four datasets. Overall, the horizontal and vertical accuracy of bundle block adjustment with mathematical distortion models was better than the physical model in the long corridor structure datasets, and the vertical accuracy of the hybrid Jacobi–Fourier model was generally better than the other three mathematical distortion models.

### 3.3. Analysis of the Performance of Proposed Self-Calibration

To verify the feasibility of the proposed strategy for camera self-calibration, this paper is compared with the scheme of ColMap [38]. Since ColMap does not implement the GNSS-constrained BA, the similarity transforms were applied for the projection center of images and GNSS after the final global BA in ColMap. Then, the weighted BA with GNSS as described in Formula (12) was conducted. Considering that ColMap does not provide any mathematical distortion models, the Brown model was adopted to make the experimental comparison.

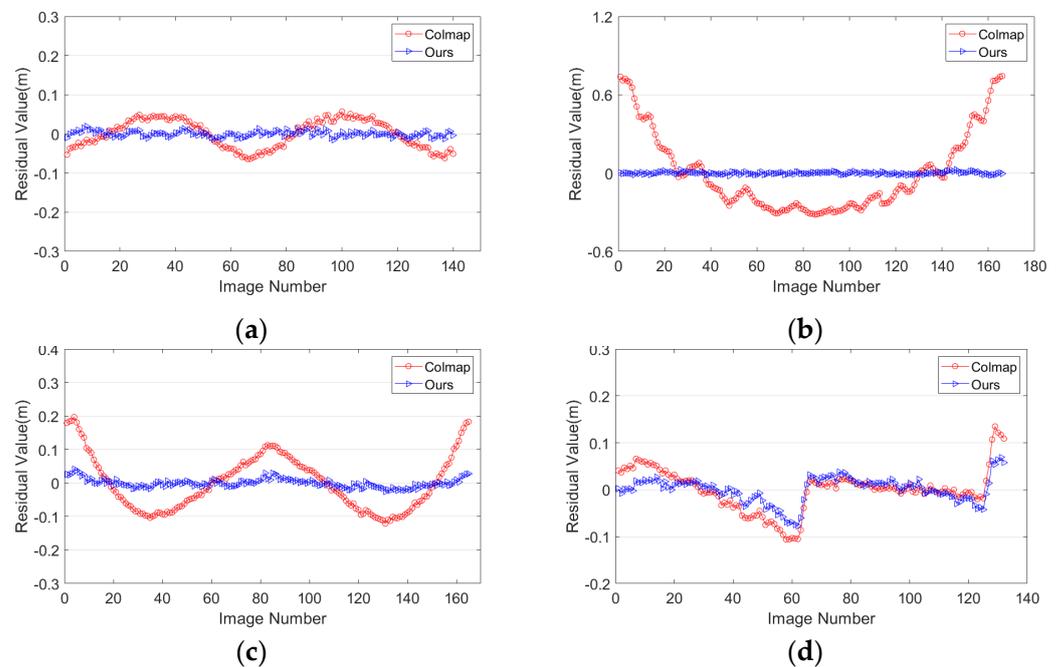


**Figure 8.** The residuals of check points after self-calibration for test site 2. (a,c,e) The residuals of check points of rectangle dataset in the direction of X, Y, and Z, respectively; (b,d,f) the residuals of check points of S-shaped dataset in the direction of X, Y, and Z, respectively.

The statistical results of check points after camera self-calibration are shown in Table 4. The results show that the horizontal accuracy with the proposed strategy is better than that of ColMap’s in the four datasets of the two test sites. In the direction of elevation, the proposed strategy significantly improved the RMSE accuracy than ColMap (except for the rectangle dataset in test site 2). Due to the “bowl effect” in ColMap camera self-calibration, although its mean value was smaller than that of the proposed strategy in the S-shaped dataset of test site 1, it had a large standard deviation, which indicates that there is a large fluctuation in the vertical direction of the reconstructed model with ColMap, as shown in Figure 9b. In the rectangle dataset of test site 2, the elevation of ColMap had a higher accuracy in the mean and RMSE than that of the proposed strategy. The reason is that the initial image pair selected during the incremental SfM framework had a better structure, which led to a smaller variation range of focal length and a higher accuracy. However, it should be noted that the standard deviation of ColMap in this dataset was 0.055 larger than ours, which shows that the ColMap’s fluctuation of elevation error is still large and there is a “bowl effect” in the reconstructed model with a certain bending phenomenon, as shown in Figure 9c. In conclusion, the proposed camera self-calibration strategy had advantages in horizontal accuracy and had better vertical accuracy than ColMap in three of the datasets.

**Table 4.** Statistical results of self-calibration for different calibration methods.

Datasets	Method	Mean(m)			SD(m)			RMSE(m)			
		X	Y	Z	X	Y	Z	X	Y	Z	
Test Site 1	Rectangle	ColMap	0.006	−0.007	0.574	0.015	0.050	0.036	0.016	0.050	0.575
		Ours	0.004	−0.001	0.174	0.011	0.036	0.037	0.012	0.036	0.178
	S-Shaped	ColMap	0.107	0.103	0.055	0.040	0.320	0.404	0.114	0.337	0.407
		Ours	0.011	0.002	−0.348	0.014	0.018	0.022	0.018	0.018	0.349
Test Site 2	Rectangle	ColMap	−0.038	0.033	0.239	0.055	0.026	0.091	0.067	0.042	0.256
		Ours	−0.018	0.028	0.753	0.030	0.025	0.036	0.035	0.037	0.754
	S-Shaped	ColMap	0.012	−0.005	−1.380	0.018	0.011	0.019	0.021	0.012	1.381
		Ours	0.016	−0.006	−1.259	0.012	0.012	0.041	0.020	0.013	1.260



**Figure 9.** The coordinate offsets in the vertical direction between image projection center and differential GNSS location. (a,b) The offsets in the vertical direction with the rectangle dataset and S-shaped dataset in test site 1; (c,d) the offsets in the vertical direction with the rectangle dataset and S-shaped dataset in test site 2.

The DJI Phantom 4 RTK UAV records the differential GNSS location of the image with centimeter positioning accuracy. The relative accuracy between the projection center of the image after camera self-calibration with the proposed strategy and the corresponding GNSS location were analyzed. As the horizontal relative errors between the proposed method and ColMap had no obvious regularity and the vertical relative errors can indicate whether the “bowl effect” appears, the vertical offset distribution of the image projection centers after camera self-calibration and the corresponding GNSS locations in the four datasets are listed in Figure 9. It can be seen that, in the elevation direction, the reconstructed model of ColMap had obvious bending. It is a convex shape that is higher in the middle and lower on both sides in the rectangle dataset of test site 1. Further, it shows a concave shape that is lower in the middle and higher on both sides in both the S-shaped dataset of test site 1 and the rectangle dataset of test site 2. The offsets in the vertical direction of the reconstructed model with the proposed method are small, which significantly alleviates the “bowl effect”. For the S-shaped dataset of test site 2, there are two broken jumps in the

proposed method and ColMap. The main reason is that the large illumination leads to the increase of mismatch feature points, which affects the accuracy of bundle block adjustment. In summary, the vertical relative accuracy of the proposed method in the vertical direction is significantly improved compared with ColMap and the bending of the reconstructed model is reduced.

### 3.4. Comparison with State-Of-The-Art Software

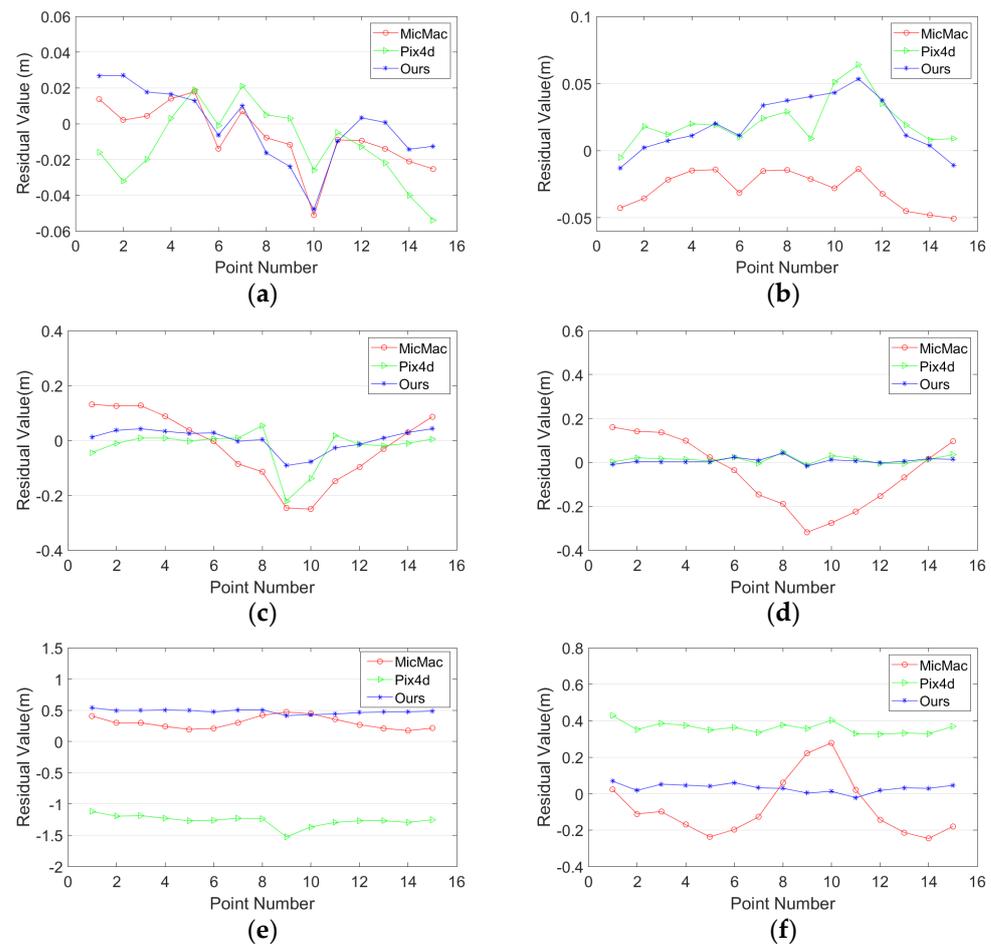
The open-source software MicMac and the commercial software Pix4d Mapper were selected to compare and analyze with the proposed method. Based on the comparative experiments mentioned above with different distortion models, it was found that the overall performance of the hybrid Jacob–Fourier model was the best. Therefore, the hybrid Jacobi–Fourier distortion model was selected for further comparative analysis. The F15P7 distortion model was adopted in MicMac with the strategies proposed in [25]. The distortion model of Pix4d Mapper was unknown. In this section, self-calibration is conducted without and with GCPs.

#### 3.4.1. Bundle Adjustment without GCP

For bundle adjustment without GCP, the experimental results are listed in Table 5. Figures 10 and 11 show the residual of check points with this software after camera self-calibration. From the analysis of horizontal accuracy, the following can be seen: (1) The proposed method had the smallest mean value in the two datasets of test site 1. The mean values of Pix4d in both datasets of test site 2 are the smallest. For the two datasets of test site 1, the standard deviation of MicMac in the Y direction is the largest, reaching 0.1 m, while Pix4d and the proposed method are both less than 0.07 m. Therefore, MicMac performs the worst overall. (2) The RMSE values in the X direction of Pix4d and the proposed method have little difference in the datasets except for the S-shaped dataset of test site 1. However, the RMSE values of Pix4d in the Y direction are 0.03 m, 0.015 m, and 0.010 m larger than the proposed method. Therefore, the horizontal accuracy of the proposed method is generally better than Pix4d.

**Table 5.** Statistical results of self-calibration for different software.

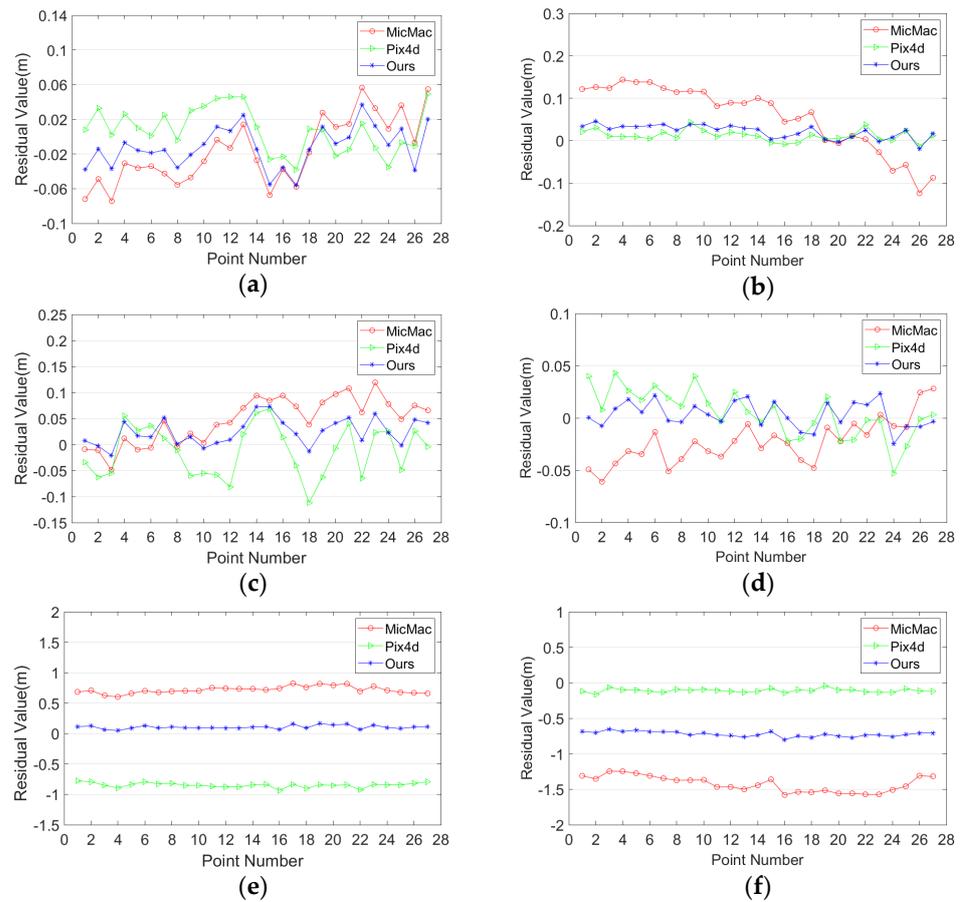
Datasets	Software	Mean(m)			SD(m)			RMSE(m)			
		X	Y	Z	X	Y	X	Y	Z	Z	
Test site 1	Rectangle	MicMac	−0.007	−0.023	0.301	0.017	0.125	0.095	0.019	0.127	0.315
		Pix4d	−0.012	−0.024	−1.266	0.020	0.066	0.088	0.024	0.070	1.269
		Ours	−0.001	0.003	0.481	0.020	0.040	0.032	0.020	0.040	0.482
	S-shaped	MicMac	−0.029	−0.050	−0.075	0.013	0.155	0.157	0.031	0.163	0.174
		Pix4d	0.021	0.013	0.360	0.017	0.016	0.029	0.027	0.021	0.362
		Ours	0.019	0.007	0.031	0.020	0.014	0.022	0.028	0.015	0.038
Test site 2	Rectangle	MicMac	−0.016	0.047	0.716	0.037	0.043	0.056	0.041	0.064	0.718
		Pix4d	0.008	−0.013	−0.847	0.025	0.048	0.037	0.026	0.050	0.848
		Ours	−0.011	0.024	0.104	0.023	0.025	0.029	0.026	0.035	0.108
	S-shaped	MicMac	0.056	−0.023	−1.425	0.076	0.021	0.108	0.095	0.031	1.429
		Pix4d	0.011	0.005	−0.111	0.013	0.023	0.025	0.018	0.023	0.113
		Ours	0.022	0.003	−0.724	0.016	0.012	0.035	0.027	0.013	0.725



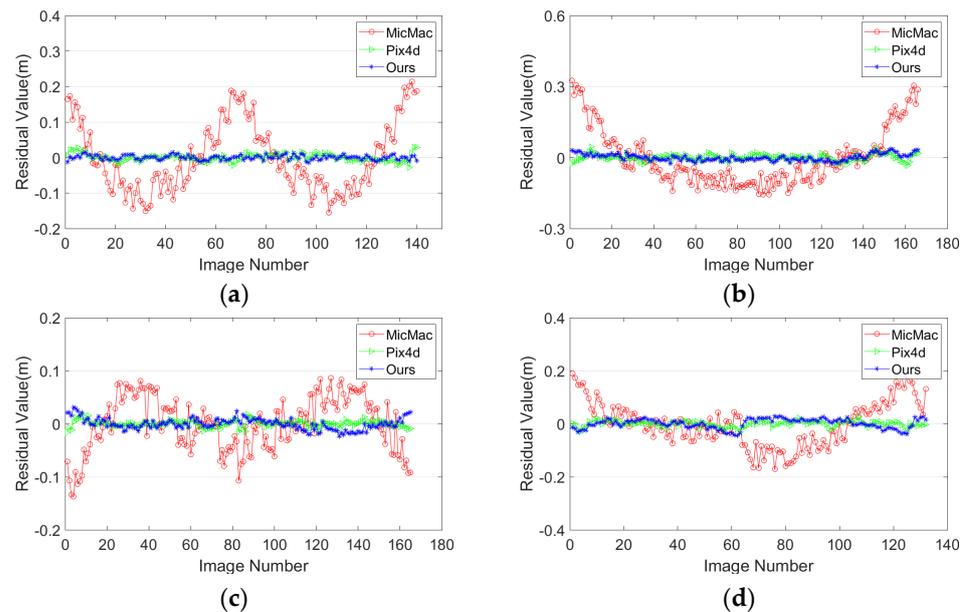
**Figure 10.** The residuals of check points after self-calibration with different software for test site 1. (a,c,e) The residuals of check points of rectangle dataset in the direction of X, Y, and Z, respectively; (b,d,f) the residuals of check points of S-shaped dataset in the direction of X, Y, and Z, respectively.

From the analysis of the vertical accuracy, the following can be seen: (1) The proposed method had the smallest standard deviation in the datasets except for the S-shaped dataset of test site 2. Pix4d had the smallest standard deviation in the S-shaped dataset of test site 2. MicMac had the largest standard deviation in the four datasets and the accuracy fluctuates greatly. (2) For test site 1, the RMSE values of MicMac were the smallest, but it had the largest standard deviation and the reconstructed model had obvious bending, as shown in Figure 12a,b. Pix4d had the smallest RMSE and standard deviation in the S-shaped dataset of test site 2. The possible reason is that the feature points matching of Pix4d is more robust with the large change of illumination. However, the RMSE values of Pix4d in the vertical direction were 0.787 m, 0.324 m, and 0.74 m larger than the proposed method, which indicates that the proposed has better accuracy in the vertical direction. To sum up, compared with MicMac and Pix4d, the proposed method still has certain advantages.

To evaluate the “bowl effect” with different software, the relative errors in the Z direction between the projection centers and the corresponding GNSS locations are shown in Figure 12. Compared with Pix4d and the proposed method, MicMac had the worst performance in the vertical relative accuracy between the projection centers and GNSS locations. The fluctuation range of vertical relative errors was between  $-0.2$  m and  $0.3$  m with MicMac, which is much bigger than Pix4d and the proposed method. There was a “bowl effect” with MicMac in the four datasets except for the S-shaped dataset in test site 2, while the bending of the reconstructed models was significantly reduced with Pix4d and the proposed method. The vertical relative errors of the proposed method were close to Pix4d. The “bowl effect” was alleviated with the proposed method and Pix4d.



**Figure 11.** The residuals of check points after self-calibration with different software for test site 2. (a,c,e) The residuals of check points of rectangle dataset in the direction of X, Y, and Z, respectively; (b,d,f) the residuals of check points of S-shaped dataset in the direction of X, Y, and Z, respectively.



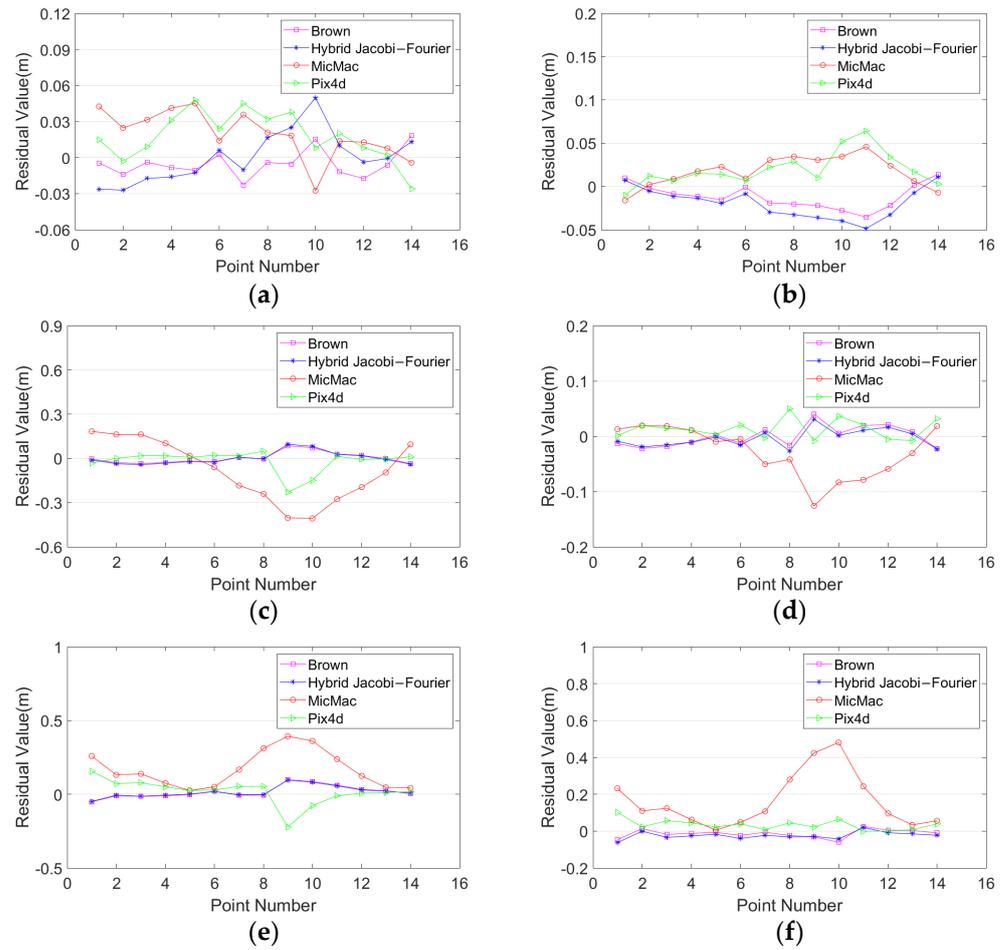
**Figure 12.** The coordinate offsets in the vertical direction between image projection center and differential GNSS location with different software. (a,b) The offsets in the vertical direction with the rectangle dataset and S-shaped dataset in test site 1; (c,d) the offsets in the vertical direction with the rectangle dataset and S-shaped dataset in test site 2.

### 3.4.2. Bundle Adjustment with GCP

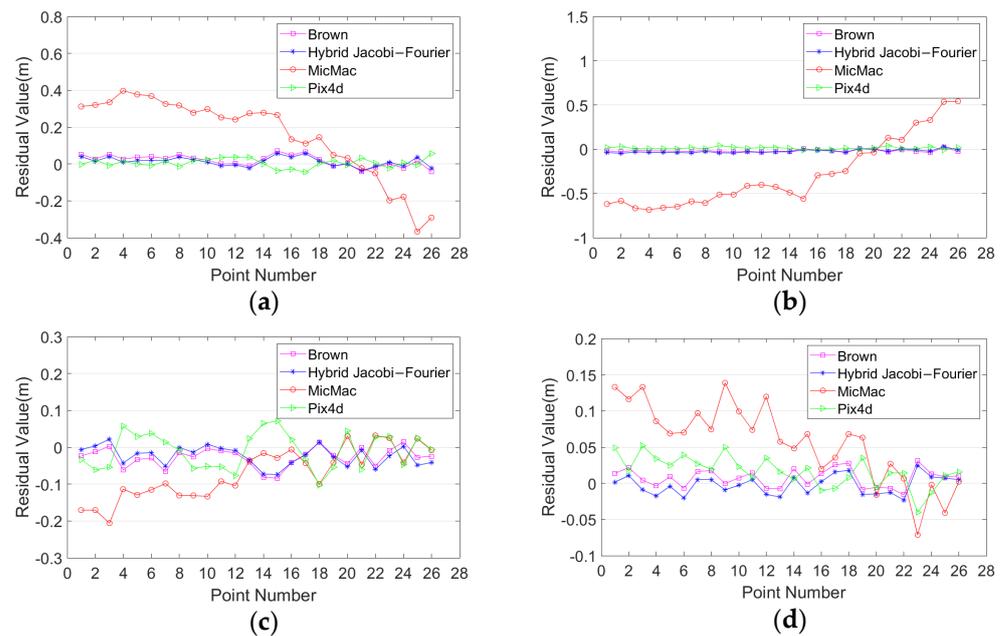
For bundle adjustment with GCP, a single GCP was applied for the camera self-calibration experiment compared with MicMac and Pix4d Mapper. All the optimized parameters of Section 3.4.1 were used as the initial parameters for MicMac and the proposed method with one single GCP-constrained BA. The Brown and hybrid Jacobi–Fourier distortion models were selected with the proposed method. Table 6 lists the experimental results. Figures 13 and 14 show the residuals of check points after camera self-calibration with one single GCP. From the analysis of horizontal accuracy, the accuracy of the Brown and hybrid Jacobi–Fourier distortion models with the proposed strategy is comparable. The mean values in the horizontal direction of X and Y with MicMac were the largest in the datasets except for the S-shaped dataset of test site 1, and the horizontal RMSE values and standard deviation values were also the largest among the four datasets. In the two datasets of test site 1, the mean values and RMSE values of Pix4d in the X and Y horizontal directions were larger than the Brown and hybrid Jacobi–Fourier model. Further, in the two datasets of the test site 2, the mean values and RMSE values of Pix4d in the horizontal direction of X were smaller than Brown and hybrid Jacobi–Fourier model with the proposed strategy, but in the Y direction, they were larger than the two distortion models with the proposed strategy. In general, the horizontal accuracy of the proposed method is relatively close to Pix4d. The horizontal RMSE value of the two distortion models with the proposed method was better than 0.04 m, while MicMac was less than 0.5 m and Pix4d was less than 0.08 m.

**Table 6.** Statistical results of self-calibration for different software with one single ground control point constraint.

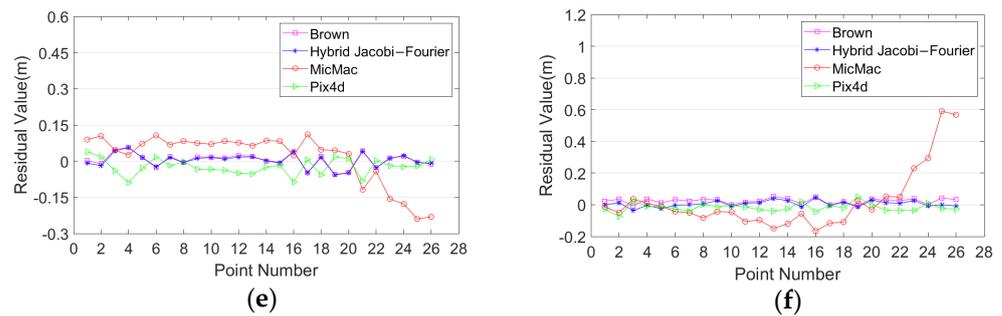
Datasets	Software	Mean(m)			SD(m)			RMSE(m)			
		X	Y	Z	X	Y	Z	X	Y	Z	
Test site 1	Rectangle	Brown	−0.005	0.003	0.016	0.011	0.036	0.040	0.012	0.036	0.043
		Jacobi–Four	0.000	0.000	0.016	0.021	0.041	0.038	0.021	0.041	0.042
		MicMac	0.020	−0.081	0.169	0.019	0.200	0.119	0.027	0.216	0.207
		Pix4d	0.018	−0.019	0.018	0.020	0.073	0.083	0.027	0.076	0.085
	S-shaped	Brown	−0.011	−0.001	−0.014	0.014	0.018	0.022	0.018	0.018	0.026
		Jacobi–Four	−0.019	−0.004	−0.024	0.017	0.016	0.019	0.026	0.016	0.030
		MicMac	0.017	−0.029	0.164	0.017	0.044	0.142	0.024	0.053	0.217
		Pix4d	0.020	0.013	0.034	0.019	0.017	0.027	0.027	0.022	0.043
Test site 2	Rectangle	Brown	0.018	−0.027	0.005	0.030	0.026	0.029	0.035	0.038	0.029
		Jacobi–Four	0.012	−0.024	0.003	0.024	0.026	0.028	0.027	0.035	0.028
		MicMac	0.154	−0.071	0.016	0.214	0.067	0.105	0.264	0.098	0.106
		Pix4d	0.005	−0.010	−0.023	0.023	0.048	0.033	0.024	0.049	0.040
	S-shaped	Brown	−0.018	0.007	0.024	0.014	0.013	0.017	0.023	0.015	0.029
		Jacobi–Four	−0.023	−0.002	0.005	0.018	0.013	0.019	0.029	0.013	0.019
		MicMac	−0.283	0.057	0.022	0.381	0.053	0.190	0.475	0.078	0.191
		Pix4d	0.012	0.017	−0.017	0.013	0.021	0.024	0.017	0.027	0.030



**Figure 13.** The residuals of check points after self-calibration with one single GCP for test site 1. (a,c,e) The residuals of check points of rectangle dataset in the direction of X, Y, and Z, respectively; (b,d,f) the residuals of check points of S-shaped dataset in the direction of X, Y, and Z, respectively.



**Figure 14.** Cont.



**Figure 14.** The residuals of check points after self-calibration with one single GCP for test site 2. (a,c,e) The residuals of check points of rectangle dataset in the direction of X, Y, and Z, respectively; (b,d,f) the residuals of check points of S-shaped dataset in the direction of X, Y, and Z, respectively.

From the analysis of elevation, the two distortion models with the proposed strategy achieved the best accuracy, while MicMac achieved the worse. The RMSE in the vertical direction of MicMac was greater than 0.1 m, while Pix4d was less than 0.1 m and the two distortion models with the proposed strategy were both less than 0.05 m. The vertical accuracy of the two distortion models was relatively close, and the difference between the first three datasets was only at the millimeter level, while the vertical RMSE of the hybrid Jacobi–Fourier model was 0.01 m better than the Brown in the S-shaped dataset of test site 2. Therefore, in the case of a single GCP constraint, the accuracy of the Brown model is comparable to the hybrid Jacobi–Fourier model, and the overall accuracy of camera self-calibration with the proposed strategy is better than that of MicMac and Pix4d. Additionally, compared with Table 5 in the vertical direction, MicMac had the smallest RMSE value in the rectangle dataset of test site 1 while Pix4d had the smallest RMSE value in the S-shaped dataset of test site 2. However, the vertical RMSE of Brown and hybrid Jacobi–Fourier models with the proposed strategy was better than MicMac and Pix4d after adding a single GCP. The reason is that the horizontal accuracy is ensured and the relative errors between image projection centers and GNSS locations are small with the experiments of direct georeferencing in Sections 3.3 and 3.4.1, which indicates that the proposed method can significantly reduce the bending of the reconstructed model and the image structures have litter distortion. The “bowl effect” is alleviated. The only problem is that the vertical accuracy is unstable because the focal length is highly correlated with other distortion parameters. With one GCP constraint, the correlation between focal length and other distortion parameters is reduced and the focal length can be accurately estimated. In this case, vertical accuracy can be ensured.

#### 4. Conclusions

The UAV images collected on a linear axis and fixed height are critical configurations for camera self-calibration, which may lead to the “bowl effect”. To solve such a tough problem, traditional methods rely on more than three GPCs, while the proposed method relies on only one GCP. The proposed new camera self-calibration method for long corridor UAV images in high transmission lines combines the initialization of the camera calibration parameters and the fusion of high-precision differential GNSS position information for long corridor UAV images in high transmission lines. Based on the comprehensive analysis of the physical and mathematical models of camera distortion, the new camera self-calibration method was designed, which takes full consideration of the initialization of the camera intrinsic parameters in long corridor UAV images and the fusion of differential GNSS with inequality constrained BA.

The UAV images of two test sites with two different acquisition modes were applied for camera self-calibration experiments. The experimental results show that the proposed camera self-calibration method can significantly alleviate the “bowl effect” for long corridor UAV images, reduce the bending of the reconstructed model, and improve the absolute accuracy. Compared to the accuracy using the physical distortion model without any GCPs,

the mathematical distortion models achieve better horizontal and vertical accuracy in the weak structure datasets. Among them, the vertical accuracy of the hybrid Jacobi–Fourier distortion model is generally better than the other mathematical models. Furthermore, with only one single GCP constraint, the proposed method with Brown and hybrid Jacobi–Fourier distortion models achieved the best accuracy compared with open-source and commercial software. Compared with the open-source software MicMac, the RMSEs in the directions of X, Y, and Z improved the GSD value on average approximately 8.36, 4.02, and 7.07 times, respectively, in the four datasets with the Brown model, and improved the GSD value on average approximately 8.18, 4.05, and 7.12 times, respectively, with hybrid Jacobi–Fourier model. Compared with the commercial software Pix4d, the RMSEs in the directions of X, Y, and Z improved the GSD value on average approximately 0.08, 0.80, and 0.85 times, respectively in the four datasets with the Brown model. Although, in the direction of X, the RMSE lost 0.01 times the GSD value with the hybrid Jacobi–Fourier model, while the RMSEs in the directions of Y and Z improved 0.82 and 0.94 times the GSD value on average, respectively. Considering that different distortion models perform differently with different scenes, this study focuses on how to select the appropriate distortion model according to the characteristics and uncertainty of the scene in future work.

**Author Contributions:** W.H. and W.J. conceived and designed the experiments; W.H. and S.J. performed the experiments; W.J. collected the data; S.J. and W.H. wrote the manuscript; W.J. revised the manuscript. 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 (Grant No. 42001413), Nature Science Foundation of Hubei Province (Grant No. 2020CFB324), the Open Fund of State Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University (Grant No. 20E03).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data available on request due to restrictions, e.g., privacy or ethical. The data presented in this study are available on request from the corresponding author. The data are not publicly available due to [the policy to protect the coordinates of pylon].

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

## References

1. Xiang, T.-Z.; Xia, G.-S.; Zhang, L. Mini-unmanned aerial vehicle-based remote sensing: Techniques, applications, and prospects. *IEEE Geosci. Remote Sens. Mag.* **2019**, *7*, 29–63. [\[CrossRef\]](#)
2. Jiang, S.; Jiang, W.; Huang, W.; Yang, L. UAV-Based Oblique Photogrammetry for Outdoor Data Acquisition and Offsite Visual Inspection of Transmission Line. *Remote Sens.* **2017**, *9*, 278. [\[CrossRef\]](#)
3. Jiang, S.; Jiang, W. UAV-based Oblique Photogrammetry for 3D Reconstruction of Transmission Line: Practices and Applications. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, Enschede, The Netherlands, 10–14 June 2019; pp. 401–406.
4. Zhang, H.; Yang, W.; Yu, H.; Zhang, H.; Xia, G.-S. Detecting power lines in UAV images with convolutional features and structured constraints. *Remote Sens.* **2019**, *11*, 1342. [\[CrossRef\]](#)
5. Zhang, Y.; Yuan, X.; Li, W.; Chen, S. Automatic Power Line Inspection Using UAV Images. *Remote Sens.* **2017**, *9*, 824. [\[CrossRef\]](#)
6. Huang, W.; Jiang, S.; Jiang, W. A Model-Driven Method for Pylon Reconstruction from Oblique UAV Images. *Sensors* **2020**, *20*, 824. [\[CrossRef\]](#)
7. Triggs, B. Autocalibration from Planar Scenes. In Proceedings of the European Conference on Computer Vision, Berlin, Germany, 2–6 June 1998; pp. 89–105.
8. Zhang, Z. A Flexible New Technique for Camera Calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1330–1334. [\[CrossRef\]](#)
9. Oniga, E.; Pfeifer, N.; Loghin, A.-M. 3D Calibration Test-Field for Digital Cameras Mounted on Unmanned Aerial Systems (UAS). *Remote Sens.* **2018**, *10*, 2017. [\[CrossRef\]](#)
10. Duane, C.B. Close-Range Camera Calibration. *Photogramm. Eng.* **1971**, *37*, 855–866.
11. Fraser, C.S. Digital camera self-calibration. *ISPRS J. Photogramm. Remote Sens.* **1997**, *52*, 149–159. [\[CrossRef\]](#)
12. Luhmann, T.; Robson, S.; Kyle, S.; Harley, I.A. *Close Range Photogrammetry: Principles, Techniques and Applications*; Whittles Publishing: Dunbeath, Caithness, Scotland, 2006; Volume 3.

13. Fitzgibbon, A.W. Simultaneous Linear Estimation of Multiple View Geometry and Lens Distortion. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, 8–14 December 2001; p. I.
14. Kukelova, Z.; Pajdla, T. A Minimal Solution to the Autocalibration of Radial Distortion. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007; pp. 1–7.
15. Kukelova, Z.; Pajdla, T. A Minimal Solution to Radial Distortion Autocalibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 2410–2422. [[CrossRef](#)]
16. Jiang, F.; Kuang, Y.; Solem, J.E.; Åström, K. A Minimal Solution to Relative Pose with Unknown Focal Length and Radial Distortion. In Proceedings of the Asian Conference on Computer Vision, Singapore, 1–5 November 2014; pp. 443–456.
17. Kukelova, Z.; Heller, J.; Bujnak, M.; Fitzgibbon, A.; Pajdla, T. Efficient Solution to the Epipolar Geometry for Radially Distorted Cameras. In Proceedings of the 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 11–18 December 2015; pp. 2309–2317.
18. Ebner, H. Self calibrating block adjustment. *Bildmess. Und Luftbildwesen* **1976**, *44*, 128–139.
19. Gruen, A. Accuracy, reliability and statistics in close-range photogrammetry. In Proceedings of the Inter-Congress Symposium of ISP Commission V, Stockholm, Sweden, 14–17 August 1978.
20. Tang, R.; Fritsch, D.; Cramer, M.; Schneider, W. A Flexible Mathematical Method for Camera Calibration in Digital Aerial Photogrammetry. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 1069–1077. [[CrossRef](#)]
21. Tang, R.; Fritsch, D.; Cramer, M. New rigorous and flexible Fourier self-calibration models for airborne camera calibration. *ISPRS J. Photogramm. Remote Sens.* **2012**, *71*, 76–85. [[CrossRef](#)]
22. Babapour, H.; Mokhtarzade, M.; Valadan Zoej, M.J. Self-calibration of digital aerial camera using combined orthogonal models. *ISPRS J. Photogramm. Remote Sens.* **2016**, *117*, 29–39. [[CrossRef](#)]
23. Wu, C. Critical Configurations for Radial Distortion Self-Calibration. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 28 June 2014; pp. 25–32.
24. Zhou, Y.; Rupnik, E.; Meynard, C.; Thom, C.; Pierrot-Deseilligny, M. Simulation and Analysis of Photogrammetric UAV Image Blocks—Influence of Camera Calibration Error. *Remote Sens.* **2019**, *12*, 22. [[CrossRef](#)]
25. Tournadre, V.; Pierrot-Deseilligny, M.; Faure, P.H. UAV Linear Photogrammetry. In Proceedings of the International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, La Grande Motte, France, 28 September–3 October 2015; p. 327.
26. Polic, M.; Steidl, S.; Albl, C.; Kukelova, Z.; Pajdla, T. Uncertainty based camera model selection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–18 June 2020; pp. 5991–6000.
27. Griffiths, D.; Burningham, H. Comparison of pre- and self-calibrated camera calibration models for UAS-derived nadir imagery for a SfM application. *Prog. Phys. Geogr. Earth Environ.* **2018**, *43*, 215–235. [[CrossRef](#)]
28. Jaud, M.; Passot, S.; Le Bivic, R.; Delacourt, C.; Grandjean, P.; Le Dantec, N. Assessing the accuracy of high resolution digital surface models computed by PhotoScan<sup>®</sup> and MicMac<sup>®</sup> in sub-optimal survey conditions. *Remote Sens.* **2016**, *8*, 465. [[CrossRef](#)]
29. Salach, A.; Bakula, K.; Pilarska, M.; Ostrowski, W.; Górski, K.; Kurczyński, Z. Accuracy assessment of point clouds from LiDAR and dense image matching acquired using the UAV platform for DTM creation. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 342. [[CrossRef](#)]
30. Jaud, M.; Passot, S.; Allemand, P.; Le Dantec, N.; Grandjean, P.; Delacourt, C. Suggestions to limit geometric distortions in the reconstruction of linear coastal landforms by SfM photogrammetry with PhotoScan<sup>®</sup> and MicMac<sup>®</sup> for UAV surveys with restricted GCPs pattern. *Drones* **2019**, *3*, 2. [[CrossRef](#)]
31. Nahon, A.; Molina, P.; Blázquez, M.; Simeon, J.; Capo, S.; Ferrero, C. Corridor mapping of sandy coastal foredunes with UAS photogrammetry and mobile laser scanning. *Remote Sens.* **2019**, *11*, 1352. [[CrossRef](#)]
32. Ferrer-González, E.; Agüera-Vega, F.; Carvajal-Ramírez, F.; Martínez-Carricondo, P. UAV Photogrammetry Accuracy Assessment for Corridor Mapping Based on the Number and Distribution of Ground Control Points. *Remote Sens.* **2020**, *12*, 2447. [[CrossRef](#)]
33. Maxime, L. Incremental Fusion of Structure-from-Motion and GPS Using Constrained Bundle Adjustments. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2489–2495. [[CrossRef](#)]
34. Gopaul, N.S.; Wang, J.; Hu, B. Camera auto-calibration in GPS/INS/stereo camera integrated kinematic positioning and navigation system. *J. Glob. Position. Syst.* **2016**, *14*, 3. [[CrossRef](#)]
35. Snow, W.L.; Childers, B.A.; Shortis, M.R. The calibration of video cameras for quantitative measurements. In Proceedings of the 39th International Instrumentation Symposium, Albuquerque, NM, USA, 2–6 May 1993; pp. 103–130.
36. Wang, J.; Wang, W.; Ma, Z. Hybrid-model based Camera Distortion Iterative Calibration Method. *Bull. Surv. Mapp.* **2019**, *4*, 103–106.
37. Jiang, S.; Jiang, C.; Jiang, W. Efficient structure from motion for large-scale UAV images: A review and a comparison of SfM tools. *ISPRS J. Photogramm. Remote Sens.* **2020**, *167*, 230–251. [[CrossRef](#)]
38. Schonberger, J.L.; Frahm, J.M. Structure-from-Motion Revisited. In Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 4104–4113.