

Article

An Efficient Closed Form Solution to the Absolute Orientation Problem for Camera with Unknown Focal Length

Kai Guo , Hu Ye, Zinian Zhao and Junhao Gu

Northwest Institute of Nuclear Technology, Xi'an 710024, China; yxc0228@163.com (H.Y.); zhao_9711@163.com (Z.Z.); gujunhao13@163.com (J.G.)

* Correspondence: guokai@mail.nwpu.edu.cn; Tel.: +86-185-102-37517

Abstract: In this paper we propose an efficient closed form solution to the absolute orientation problem for cameras with an unknown focal length, from two 2D–3D point correspondences and the camera position. The problem can be decomposed into two simple sub-problems and can be solved with angle constraints. A polynomial equation of one variable is solved to determine the focal length, and then a geometric approach is used to determine the absolute orientation. The geometric derivations are easy to understand and significantly improve performance. Rewriting the camera model with the known camera position leads to a simpler and more efficient closed form solution, and this gives a single solution, without the multi-solution phenomena of perspective-three-point (P3P) solvers. Experimental results demonstrated that our proposed method has a better performance in terms of numerical stability, noise sensitivity, and computational speed, with synthetic data and real images.



Citation: Guo, K.; Ye, H.; Zhao, Z.; Gu, J. An Efficient Closed Form Solution to the Absolute Orientation Problem for Camera with Unknown Focal Length. *Sensors* **2021**, *21*, 6480. <https://doi.org/10.3390/s21196480>

Academic Editors: Liang Zhang, Mohammed Bennamoun and Mingtao Feng

Received: 13 August 2021

Accepted: 26 September 2021

Published: 28 September 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/>).

Keywords: absolute orientation; camera position; angle constraint; single solution; unknown focal length; perspective-three-point

1. Introduction

Many methods have been proposed to estimate absolute camera pose, i.e., the position and orientation, such as the perspective- n -point (PnP) solver [1–9], which uses n known 2D–3D point correspondences. Pose estimation is one of the key steps in computer vision [2,10,11], photogrammetry [3,11,12], augmented reality (AR) [4,13–15], structure from motion (SfM) [4,14,16], multi-view 3D reconstruction [17,18], and simultaneous localization and mapping (SLAM) [4,8,13,19]. The absolute pose of a fully uncalibrated camera pose contains six unknown parameters, and each 2D–3D point correspondence gives two constraints [20], which means that the P3P is the minimal subset to determine the camera pose if the position and orientation are both unknown [10,21–24]. Many P3P solvers have been proposed, and all the solvers have up to four possible solutions [12,25,26]. In general, disambiguating the multi-solution phenomena can be done by using a fourth point. We can see that, although the P3P needs minimal 2D–3D point correspondences, all P3P solvers have some disadvantages: a fully calibrated camera is needed and multi-solution phenomena exists. These disadvantages thus prevent their application when the intrinsic camera parameters change online or are unknown. Hence, for pose estimation, many methods have been proposed to work with a partially calibrated camera and more 2D–3D point correspondences [27]. Some methods, namely the PnPf solvers, work well with cases of unknown focal length [28–30]. Four or more 2D–3D point correspondences are needed for all PnP solvers. The P4Pf is the minimal subset, and different methods have been proposed to focus on the planar case [31], the non-planar case [27], or both [32]. Compared to the P3P solvers, only one more parameter, i.e., focal length, must be obtained, and they are iterative algorithms or need to solve quadratic or quadric polynomial equations of several variables. Hence, some methods have been proposed to work with unknown focal length and unknown radial distortion (namely, the PnPfr solvers [33,34]), while some

work with unknown focal length and unknown aspect ratio [35], or unknown focal length and unknown principal point [27]. When $n \geq 6$, the pose estimation can be linearly estimated, known as direct linear transform (DLT) [18,32], and all the parameters of a fully uncalibrated camera can be obtained.

Note that more parameters can be estimated with more 3D control points. However, in some cases, not enough 3D control points can be obtained because accurate 3D control points are expensive to acquire and maintain. This requires us to use as few points as possible to estimate the pose with a partially calibrated camera, and there are two ways to reduce the number of the 3D control points in existing PnP solvers. The first way is to use some prior knowledge of the intrinsic camera parameters. For most modern digital cameras, the aspect ratio of the pixels, the skew, and the principal point are known and do not change [32,33]; hence, these parameters can be assumed as prior knowledge, which means we can use fewer 3D control points to estimate the remaining unknown parameters. With this assumption, only the focal length is unknown of the intrinsic camera parameters, and it will be shown that, in our experiments and practical application, this assumption works well, even though it is not always strictly met.

In addition, since modern digital cameras can be equipped with various positioning and orientation sensors, the second method is to measure some pose parameters in advance, as prior knowledge. Some methods focus on the pose problem with the known vertical direction. This can be obtained directly using orientation sensors, such as gyroscopes, accelerometers, or inertial measurement units (IMUs) [3,20,36–41]. The vertical direction can give knowledge of the orientation of roll and pitch, which means only four pose parameters are left to be estimated [13,15,17,42–45]. These methods can use two 3D points for pose problem and give two solutions. Some methods solve the pose problem with three 2D–3D point correspondences and the vertical direction. In this case, six parameters (one orientation parameter, three position parameters, radial distortion, and focal length) can be determined with a single solution.

In this paper, the idea is also to measure some pose parameters in advance, as prior knowledge, but not the orientation parameters. Pose parameters include the orientation and position. However, to the best of our knowledge, almost all recent research has focused on the known orientation parameters, and very few works focused on the known position parameters. Moreover, in some cases, the camera position and 3D control point positions can be obtained accurately as prior knowledge using a positioning device (e.g., RTK, total station). In a missile testing range, for example, altitude measurement based on fixed cameras is an important test. These cameras are fixed and for absolute pose problem, some 3D control points in the world frame must be exactly known. Hence, in this paper, we focus on the known position parameters [46] to solve the pose problem, and we give an efficient closed form solution to the absolute orientation problem with unknown focal length from two 2D–3D point correspondences. Since each point correspondence can give us two constraints [3], this is the minimum number of point correspondences needed to estimate the absolute orientation and focal length in this case. Here, the problem can be decomposed into two sub-problems and can be solved with angle constraints. Rewriting the camera model with the known camera position leads to a simpler and more efficient method for pose estimation, and it gives a single solution, without the multi-solution phenomena of existing P3P solvers.

The rest of this paper is organized as follows. In Section 2, we propose our method to efficiently estimate the focal length and the absolute orientation. In Section 3, we present a thorough analysis of our proposed method with synthetic data and real images, compared to some other existing PnP solvers. In Section 4, we present the discussion. In Section 5, we present the conclusions.

2. Materials and Methods

In this paper, we propose an efficient closed form solution to the absolute orientation problem for cameras with unknown focal length from two 2D–3D point correspondences

and the camera position. The standard pinhole camera model [18] is used, as shown in Figure 1. In our problem, we assume that the skew is zero, the aspect ratio of the pixels is one, and the principal point is the center of the image, which is true for most modern digital cameras and can yield good results, even when they are not exactly satisfied; as will be shown in the experiments [3,33]. In this paper, the camera position $O_c(X_{Oc}, Y_{Oc}, Z_{Oc})$ is known, which can be obtained by positioning sensors [45,47] or measured by the total station [48].

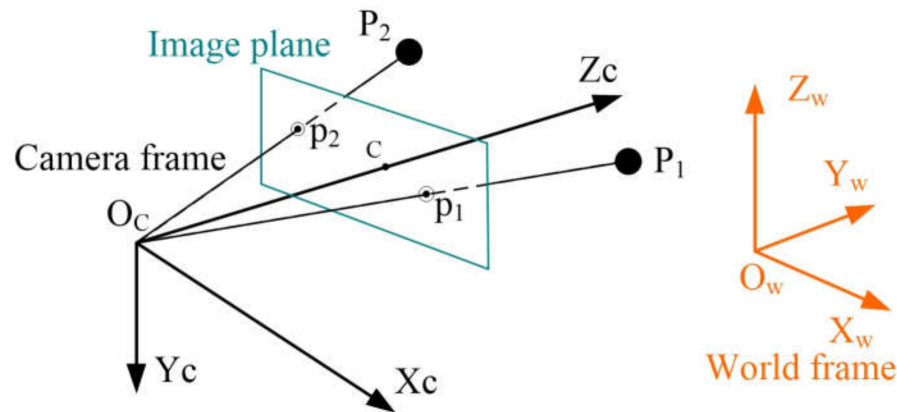


Figure 1. Standard pinhole camera model with two 3D control points. Here C is the principal point.

In Figure 1, 3D points $P_i(X_{wi}, Y_{wi}, Z_{wi})$, $i = 1, 2$ in the world frame O_XYZ_w are projected onto 2D image points $p_i(u_i, v_i)$ on the camera image plane. This can be written as

$$\lambda_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = M \begin{bmatrix} X_{wi} \\ Y_{wi} \\ Z_{wi} \\ 1 \end{bmatrix} \quad (1)$$

In this equation, M is a 3×4 camera projection matrix and λ_i is an unknown scale factor. From the standard pinhole camera model, M can be written as

$$M = K[R|t] \quad (2)$$

Here, K is a 3×3 camera calibration matrix that contains the focal length information. R and t , which contain all the pose information, are respectively a 3×3 rotation matrix and a 3×1 translation vector from the world frame to the camera frame. Our problem is to estimate R , t , and the focal length f from two 2D–3D point correspondences. Next, we propose our method to estimate the focal length and absolute orientation with angle constraints.

2.1. Closed Form Solution to the Focal Length

In this paper, we assume

$$K = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

Then we can obtain the vector $\vec{O_c p_i}$ in the camera frame

$$\vec{O_c p_i} = [u_i, v_i, f] \quad (4)$$

From Figure 1, the angle constraint now can be used to estimate the focal length, as illustrated in Figure 2.

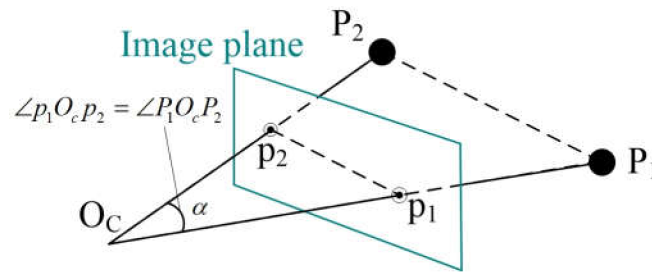


Figure 2. Angle constraint for the focal length estimation.

With the positions of 3D point P_1, P_2 and the camera position O_c in the world frame, we can obtain the vector $\vec{O_c P_i}$

$$\vec{O_c P_i} = P_i - O_c \quad (5)$$

Then, $\angle P_1 O_c P_2$ can be computed as α

$$\alpha = \arccos \frac{\vec{O_c P_1} \cdot \vec{O_c P_2}}{\|\vec{O_c P_1}\| \cdot \|\vec{O_c P_2}\|} \quad (6)$$

In the camera frame, from Equation (4) $\angle p_1 O_c p_2$ can be computed and from Figure 2, we can see $\angle p_1 O_c p_2 = \angle P_1 O_c P_2$, which leads to the equation

$$\cos \alpha = \frac{\vec{O_c p_1} \cdot \vec{O_c p_2}}{\|\vec{O_c p_1}\| \cdot \|\vec{O_c p_2}\|} = \frac{u_1 u_2 + v_1 v_2 + f^2}{\sqrt{u_1^2 + v_1^2 + f^2} \cdot \sqrt{u_2^2 + v_2^2 + f^2}} \quad (7)$$

We let $f^2 = a$, $u_1 u_2 + v_1 v_2 = b$, $u_1^2 + v_1^2 = c$ and $u_2^2 + v_2^2 = d$. From Equation (7), a quadratic equation with one variable, i.e., a , can be given

$$(1 - \cos^2 \alpha) a^2 + (2b - c \cdot \cos^2 \alpha - d \cdot \cos^2 \alpha) a + b^2 - cd \cdot \cos^2 \alpha = 0 \quad (8)$$

Two possible solutions to a can be obtained. Then up to four possible solutions to the focal length can be given from Equation (8). Note that $a > 0$, $f > 0$, and $\cos \alpha > 0$, then a single closed form solution can be given.

2.2. Pose Estimation with Angle Constraint

In this paper, we first place the camera with an original known pose in the world frame, which means the transformation between the camera frame and the world frame is known. Then the pose estimation is obtained through rotating the camera and world frame to make the camera position O_c , 2D image point p_i and 3D point P_i collinear. The process is illustrated in Figure 3.

In the original state, the camera pose is known in the original world frame O_ZYZ_w ; however, the 2D image point p_i and 3D point P_i have no correspondence, as shown in Figure 3 (left). The main work is to rotate the original camera frame O_ZYZ_c and world frame O_ZYZ_w to make the camera position O_c , 2D image point p_i , and 3D point P_i collinear in the final state, as shown in Figure 3 (right).

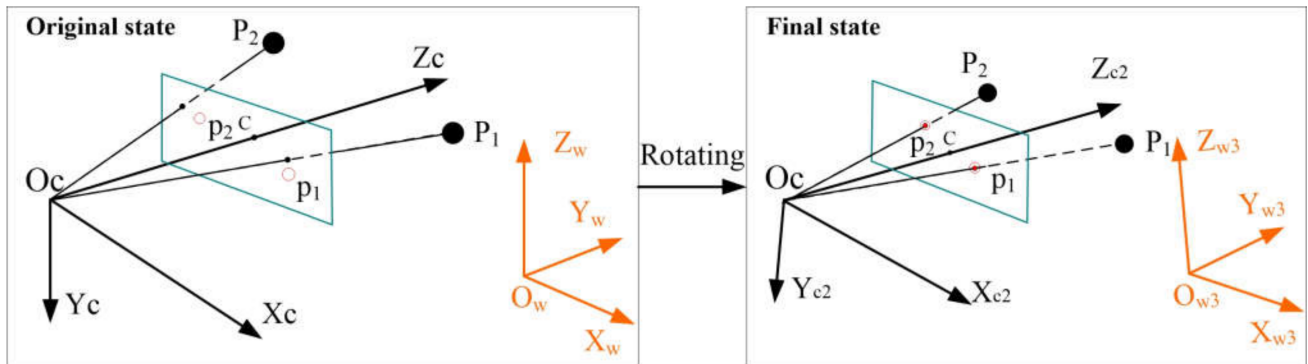


Figure 3. Rotating for pose estimation.

Now we formulate the absolute orientation estimation problem as follows:

(1) Finish the 2D–3D point correspondence between point p_1 and point P_1 . In the original camera frame O_YZZ_c , the X_c -axis and Z_c -axis are parallel with the X -axis and Y -axis of the original world frame O_YZZ_w in the same direction, and the Y_c -axis is parallel with the Z -axis in the opposite direction. Then the position of point P_1 in the camera frame O_YZZ_c , which is named P_1^c , can be obtained using the formula

$$P_1^c = R_{ox} \cdot [P_1 - O_c] \quad (9)$$

Here,

$$R_{ox} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 90^\circ & -\sin 90^\circ \\ 0 & \sin 90^\circ & \cos 90^\circ \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad (10)$$

The position of point p_1 in the camera frame O_YZZ_c , meanwhile, which is named p_1^c , can be obtained using the formula

$$p_1^c = [u_1 \quad v_1 \quad f]^T \quad (11)$$

In the camera frame, we rotate the camera around the Y_c -axis to make the projections of $\vec{O_c P_1^c}$ and $\vec{O_c p_1^c}$ onto the plane $Y_c = 0$ collinear. The rotation angle A_{Y_c} can be obtained using the formula

$$A_{Y_c} = \arccos \left(\frac{\vec{O_c P_1^c} \cdot \vec{O_c p_1^c}}{\|\vec{O_c P_1^c}\| \|\vec{O_c p_1^c}\|} \right)_{Y_c=0} \quad (12)$$

After the first rotation, a new camera frame O_YZZ_c1 is obtained and in this frame, the position of point P_1 , named P_1^{c1} , can be written as

$$P_1^{c1} = R_{cY_c} \cdot P_1^c \quad (13)$$

Here,

$$R_{cY_c} = \begin{bmatrix} \cos A_{Y_c} & 0 & -\sin A_{Y_c} \\ 0 & 1 & 0 \\ \sin A_{Y_c} & 0 & \cos A_{Y_c} \end{bmatrix} \quad (14)$$

The position of point p_1 in the new camera frame, named p_1^{c1} , is unchanged, which means $p_1^{c1} = p_1^c$.

Next, we rotate the camera around the X_{c1} -axis to make $\vec{O_c P_1}$ and $\vec{O_c p_1}$ collinear. Now we obtain another camera frame O_YZZ_c2 , as shown in Figure 4.

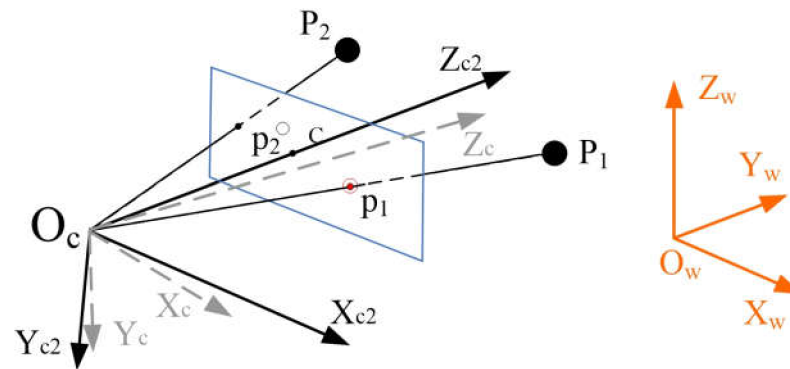


Figure 4. Camera frame O_ZYZ_c2 . Now the 2D–3D point correspondence between point p_1 and point P_1 is finished.

The rotation angle $A_{X_{c1}}$ can be obtained using the formula

$$A_{X_{c1}} = \arccos \left(\frac{\vec{O_c P_1^{c1}} \cdot \vec{O_c p_1^{c1}}}{\|\vec{O_c P_1^{c1}}\| \|\vec{O_c p_1^{c1}}\|} \right)_{X_{c1}=0} \quad (15)$$

The 2D–3D point correspondence between point p_1 and point P_1 is completed as shown in Figure 4.

(2) Finish the 2D–3D point correspondence between point p_2 and point P_2 . When the point correspondence between point p_2 and point P_2 is finished and the point correspondence between point p_1 and point P_1 is unchanged, the camera absolute orientation is obtained.

Now the position of point p_2 in the original world frame O_ZYZ_w , named p_2^w , can be computed with

$$p_2^w = (R_{cX_{c1}} \cdot R_{cY_c} \cdot R_{ox})^{-1} \cdot \begin{bmatrix} u_2 \\ v_2 \\ f \end{bmatrix} + O_c \quad (16)$$

In this equation,

$$R_{cX_{c1}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos A_{X_{c1}} & -\sin A_{X_{c1}} \\ 0 & \sin A_{X_{c1}} & \cos A_{X_{c1}} \end{bmatrix} \quad (17)$$

To maintain the point correspondence between point p_1 and point P_1 , we rotate the original world frame around the line $O_c P_1$. We thus define a new world frame, O_ZYZ_w1 , whose origin O_{w1} is camera position O_c , and

$$\begin{aligned} \vec{O_{w1} X_{w1}} &= \frac{\vec{O_{w1} P_1}}{\|\vec{O_{w1} P_1}\|} \\ \vec{O_{w1} Z_{w1}} &= \frac{\vec{O_{w1} P_1} \times \vec{O_{w1} P_2}}{\|\vec{O_{w1} P_1} \times \vec{O_{w1} P_2}\|} \\ \vec{O_{w1} Y_{w1}} &= \vec{O_{w1} Z_{w1}} \times \vec{O_{w1} X_{w1}} \end{aligned} \quad (18)$$

The new world frame O_ZYZ_w1 is illustrated in Figure 5.

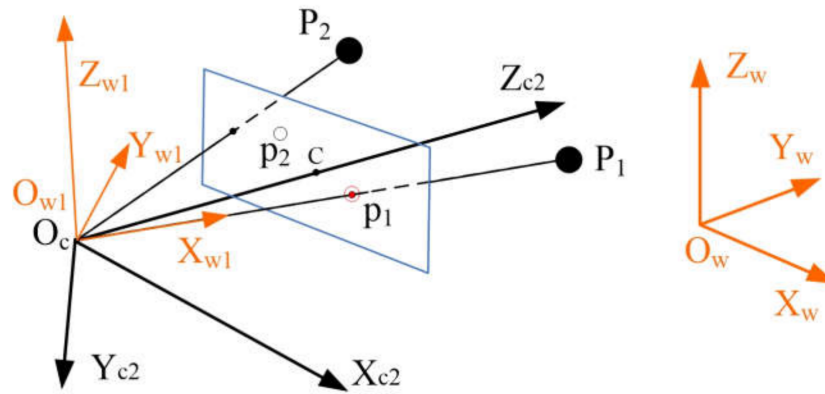


Figure 5. New world frame O_ZYZ_w1 . The $Xw1$ -axis is collinear with the line $OcP1$.

In the new world frame O_ZYZ_w1 the positions of point $P2$ and $p2$ can be given with

$$\begin{aligned} P_2^{w1} &= R_{w1}(P_2 - O_c) \\ p_2^{w1} &= R_{w1}(p_2^w - O_c) \end{aligned} \quad (19)$$

Here,

$$R_{w1} = \begin{bmatrix} \vec{O_{w1}X_{w1}} & \vec{O_{w1}Y_{w1}} & \vec{O_{w1}Z_{w1}} \end{bmatrix}^T \quad (20)$$

We rotate the world frame O_ZYZ_w1 , point $P1$, and $P2$ around the $Xw1$ -axis. With this rotation the relative pose between the world frame and point P_i is unchanged, while the relative pose between the world frame and the camera frame is changed.

To make $\vec{O_cP_2}$ and $\vec{O_cp_2}$ collinear, we rotate the world frame O_ZYZ_w1 around the $Xw1$ -axis with an angle

$$A_{xw1} = \arccos \left(\frac{\vec{O_{w1}P_2^{w1}} \cdot \vec{O_{w1}p_2^{w1}}}{\|\vec{O_{w1}P_2^{w1}}\| \|\vec{O_{w1}p_2^{w1}}\|} \right) \quad (21)$$

After this rotation, another world frame O_ZYZ_w2 is obtained and the rotation matrix between the world frame O_ZYZ_w1 and the world frame O_ZYZ_w2 is written as

$$R_{w2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos A_{xw1} & -\sin A_{xw1} \\ 0 & \sin A_{xw1} & \cos A_{xw1} \end{bmatrix} \quad (22)$$

In addition, the original world frame O_ZYZ_w is changed to a new world frame O_ZYZ_w3 . Finally, the two 2D–3D point correspondences are finished as shown in Figure 6.

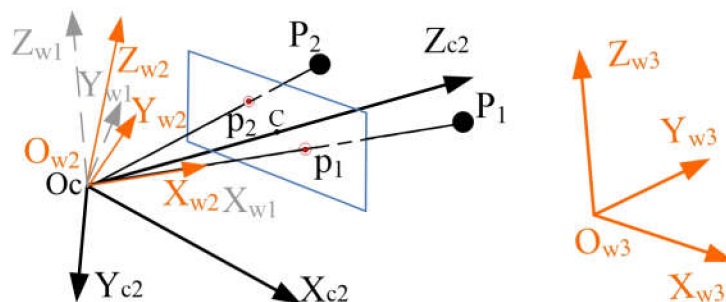


Figure 6. Two 2D–3D point correspondences in the final state. Now the absolute pose estimation is finished.

(3) Estimate the absolute orientation. Several frames are involved in our proposed method, and now the transformations between each are known, except for the transformation between the world frame O_XYZ_w3 and the camera frame O_XYZ_c2 , which is the very pose information that needs to be estimated in this paper. The transformations are shown in Figure 7.

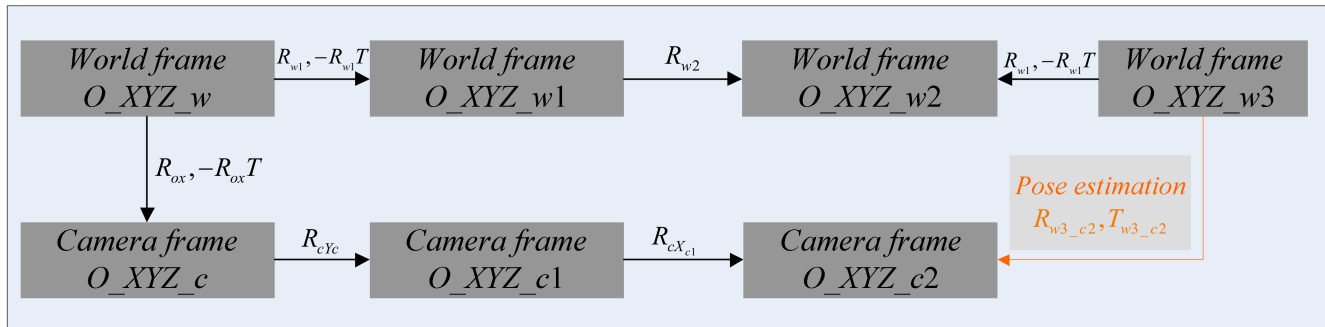


Figure 7. Transformations of all the frames. The transformation (yellow) between the world frame O_XYZ_w3 and the camera frame O_XYZ_c2 is unknown and needs to be estimated, while the other transformations have been computed.

Based on Figure 7, we can finally transform point P_i^{w3} in the world frame O_XYZ_w3 into point P_i^{c2} in the camera frame O_XYZ_c2 using

$$\begin{aligned} P_i^{c2} &= R_{w3_c2} \cdot P_i^{w3} + T_{w3_c2} \\ R_{w3_c2} &= R_{cxc1} \cdot R_{cyc} \cdot R_{ox} \cdot R_{w1}^{-1} \cdot R_{w2}^{-1} \cdot R_{w1} \\ T_{w3_c2} &= -R_{w3_c2} \cdot O_c \end{aligned} \quad (23)$$

The absolute orientation estimation with unknown focal length is finished.

3. Experiments and Results

We first tested the robustness to camera position noise of our proposed method with synthetic data.

Then we thoroughly tested our proposed method with synthetic data, including numerical stability, noise sensitivity, and computational speed, compared to other existing PnP solvers: the GP4Pf [28] and Kneip's method [10]. The two existing PnP solvers both give up to four possible solutions, while we used one more point to give a single solution.

Lastly, our proposed method was tested with real images to show its performance in a practical application.

3.1. Synthetic Data

In this paper, the synthetic data consisted of three thousand 2D–3D point correspondences. Here, these 3D points were randomly distributed in a box of $[-20, 20] \times [-20, 20] \times [180, 220]$ in the world frame. Then they were projected onto 2D points in the image plane using a virtual perspective camera, whose position was fixed at $O_c = [1, 1, 1]^T$ and the angles in degree of the orientation were kept at $[roll, pitch, yaw] = [5, 5, 5]$. For the intrinsic parameters of the virtual perspective camera, the focal length was set to 50 mm and the image resolution was set to 1280×800 pixels.

For each trail, two 2D–3D point correspondences were randomly selected from the synthetic data for our proposed method, while three 2D–3D point correspondences were randomly selected from the synthetic data for Kneip's method, and four 2D–3D point correspondences were randomly selected from the synthetic data for the GP4Pf. Moreover, one further 2D–3D point correspondence was selected for Kneip's method and GP4Pf to disambiguate the multi-solution phenomena.

3.1.1. Robustness to Camera Position Noise

Our proposed method uses the camera position as the prior knowledge, which is different from the existing methods. Therefore, the camera position is important, and it is necessary to analyze the effect of error in the camera position on the estimation of the absolute orientation and the focal length.

The camera position is usually obtained by RTK or total station. In general, the measuring precision of RTK is better than 3 cm and the measuring precision of total station is better than 0.5 cm. Therefore, zero-mean Gaussian noise was added to the camera position and the noise deviation level varied from 0 to 3 cm. Next, 50,000 independent trails with two 2D–3D point correspondences of synthetic data were performed at each noise level. Then the average error of the absolute orientation and focal length were reported, as shown in Figure 8.

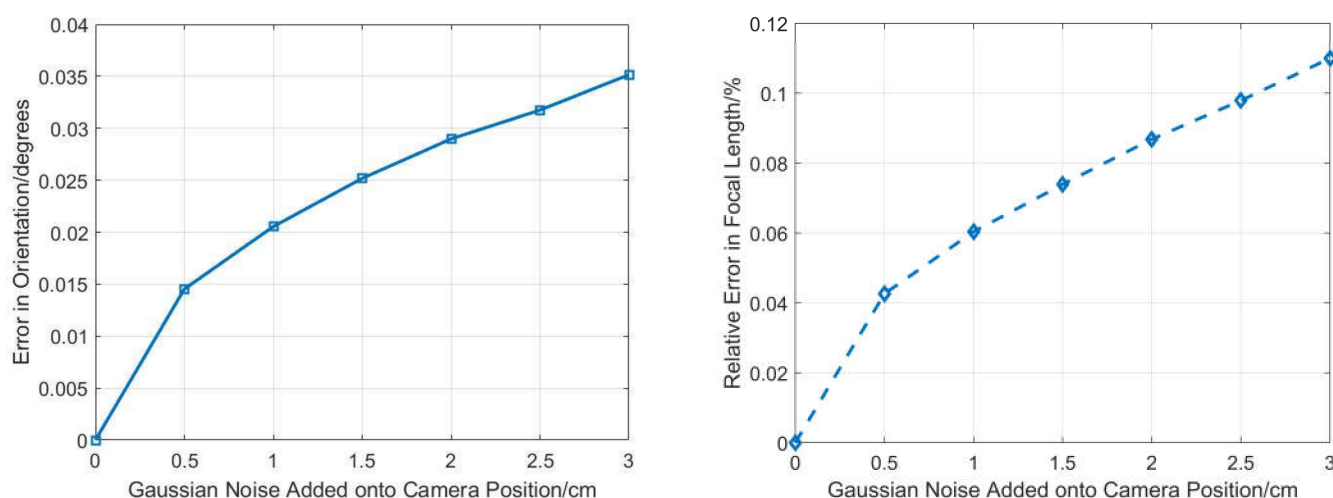


Figure 8. Robustness to camera position noise for orientation (left) and focal length (right).

From Figure 8, we can see the orientation error and focal length error increase with the increase of camera position error. However, the max errors in orientation and focal length when the camera position error is 3 cm were both low, which means our proposed method has good robustness to camera position noise and still yields good results, even though camera position error existed.

3.1.2. Numerical Stability

In this section, 50,000 trails were performed independently and there was no noise added to the 2D–3D point correspondences. The \log_{10} value of the relative error between the ground truth and the focal length, estimated using our proposed method and GP4Pf, respectively, is shown in Figure 9 (left). The \log_{10} value of the error in orientation between the ground truth and the estimated value using our proposed method and Kneip's method, respectively, is shown in Figure 9 (right).

From Figure 9 (left), the distribution of the \log_{10} value of the relative focal length error can be observed. Clearly, our proposed method has much higher numerical stability than the GP4Pf.

From Figure 9 (right), the distribution of \log_{10} value of error in orientation can be observed. Obviously, our proposed method has much higher numerical stability than Kneip's method.

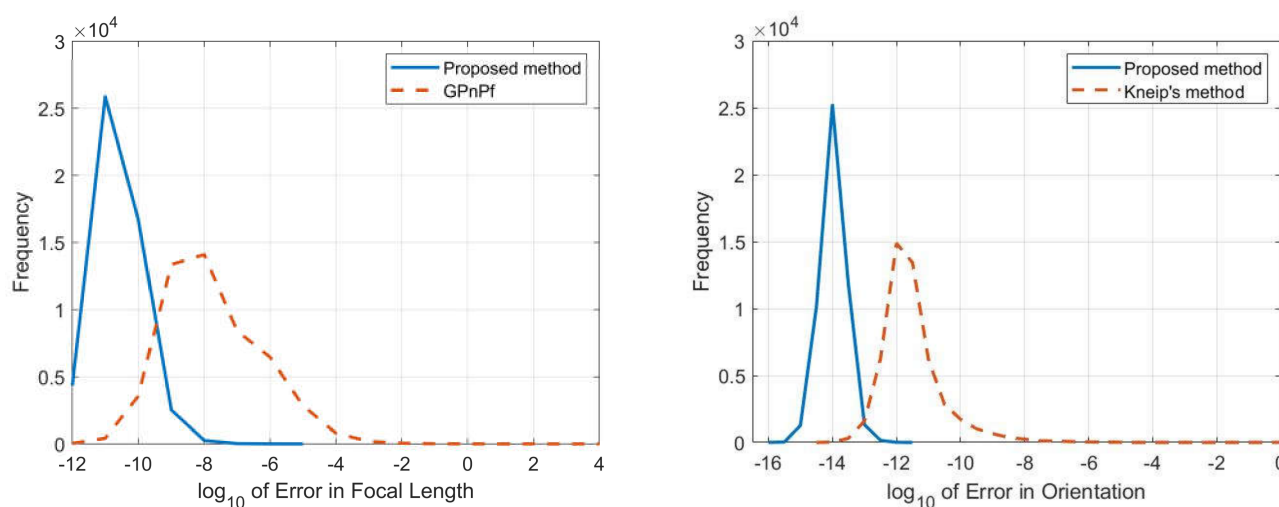


Figure 9. Relative error in focal length (left) and error in orientation (right) for our proposed method (blue) and the other methods (yellow).

3.1.3. Noise Sensitivity

Zero-mean Gaussian noise was added to the 2D image points and the noise deviation level varied from 0 to 2 pixels. Then, 50,000 independent trails were performed at each noise level. The average error of the rotation, translation, focal length, and reprojection error were reported, as shown in Figure 10.

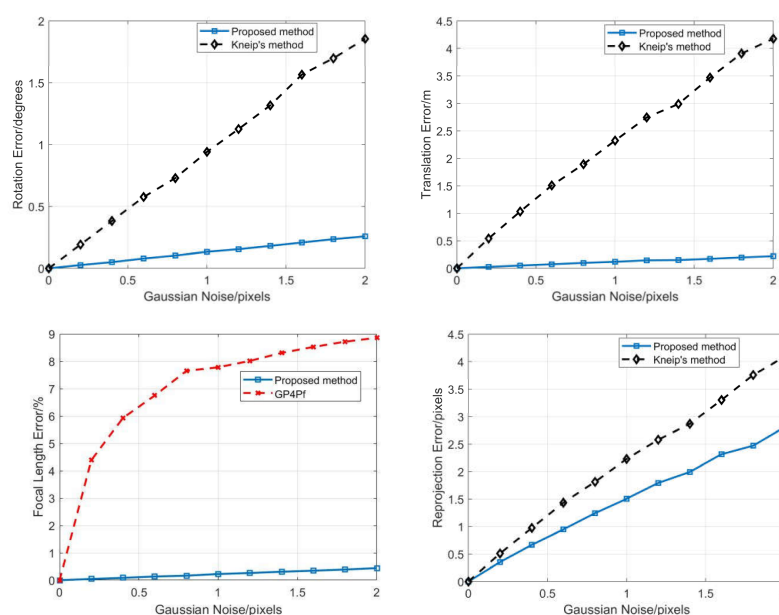


Figure 10. Average error of rotation (top left), translation (top right), focal length (bottom left) and reprojection (bottom right) for our proposed method (blue), Kneip's method (black), and GP4Pf (red).

From Figure 10, in terms of the rotation and translation error, our proposed method performed much better than Kneip's method, while it was slightly better in terms of reprojection error. In terms of the relative focal length error, our proposed method performed much better than the GP4Pf. Moreover, as the noise increases, the performance superiority of our proposed method over the other methods became more obvious.

3.1.4. Computational Time

In this section, to analyze the computational time, 50,000 trials were executed independently on a 3.3 GHz 4-core laptop, and there was no noise added to the 2D–3D point correspondences. In each trial, note that one more point was needed to disambiguate multi-solution phenomena for Kneip’s method and the GP4Pf. The average computational time is reported in Table 1.

Table 1. Computational time.

Method	Proposed Method	Kneip’s Method	GP4Pf
Computational time	0.543 ms	0.556 ms	2.683 ms

We note that our proposed method performed much faster than the GP4Pf, while it was slightly faster than Kneip’s method.

3.2. Real Images

When we generated the synthetic data, the focal length and absolute orientation of the virtual perspective camera were ground truth. Therefore, we could make direct comparisons, leading to direct results. However, in the real-image experiments, we fixed a high-speed camera with a zoom lens on a tripod, and set the focal length to roughly 50 mm. This meant that the ground truth of the focal length and absolute orientation could not be directly and accurately measured by direct physical measurement. Although many methods have been proposed to estimate the focal length and absolute orientation, these are just measured values, not the ground truth.

Although the focal length and absolute orientation cannot be directly and accurately measured by direct physical measurement, the spatial position of the points can be directly and accurately measured by direct physical measurement (total station). The world frame can be established by total station in the lab, and the measurement accuracy of total station is generally better than 0.5 cm. Therefore, in this paper we took the spatial position of a point measured by total station as the ground truth, to test the performance of our proposed method. Certainly, the point position is not estimated directly by our proposed method, but the purpose of the focal length and absolute orientation estimation in our method is 3D measurement, such as point position and 3D reconstruction. The absolute position of a point is generally measured by binocular vision, based on two cameras, after intrinsic and extrinsic camera parameter estimation, including the focal length and camera pose. When the intrinsic and extrinsic camera parameters are known, the least square method can be used to estimate the point position, and then the relative position error can be given, which is very simple. We can see that the key step of the point position estimation is the intrinsic and extrinsic camera parameter estimation, i.e., the focal length and absolute orientation in this paper. Therefore, the accuracy of the absolute orientation and focal length estimation directly affects the relative position error of points, and in turn, the relative position error can reflect the accuracy of the absolute orientation and focal length estimation with our proposed method. Moreover, the relative position error can be measured in our lab, since the ground truth of a point position can be given by the total station, and the measured value can be given using binocular vision with our proposed method.

In addition, the ground truth of a point position is known, and then we can obtain the reprojection, based on the standard pinhole camera model [18], with the focal length and absolute orientation measured by our proposed method. The reprojection is the measured value of the imaging position and the ground truth can be obtained by corner detection from the real images. Therefore, the reprojection error is affected by the focal length and absolute orientation estimation, and in turn, the reprojection error can reflect the accuracy of the focal length and absolute orientation estimation with our proposed method.

Therefore, indirect analysis and comparison, for testing the performance of our method with real images, are practicable. Moreover, in this paper we use relative position and

reprojection error to reflect the error of the focal length and absolute orientation estimation when the focal length and absolute orientation cannot be directly and accurately measured using direct physical measurement in the lab. The experiments and results with real images are as follows.

In this section, real images were captured using two cameras, and then we tested our proposed method with them. Some control points were placed in these two camera fields of view, as shown in Figure 11.

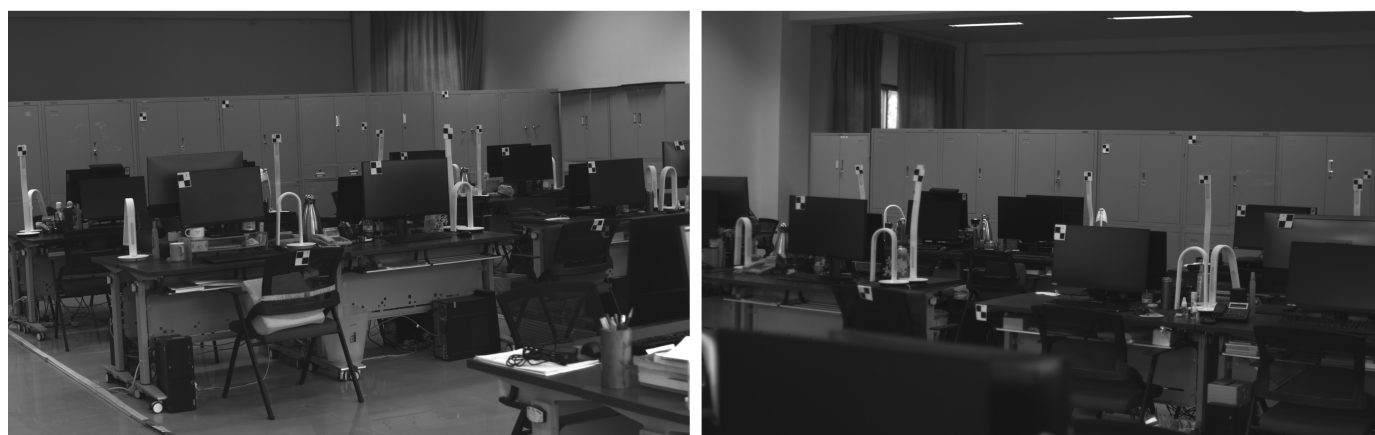


Figure 11. Real images from two cameras. Some control points were placed and measured using a total station.

These control points and the camera positions were measured as the ground truth using a total station (NTS-330R, measuring precision better than 0.5 cm). Since we did not know the ground truth of the camera pose in the real scenarios, the accuracy of the focal length and absolute orientation was not compared directly. In this paper, the accuracy of the absolute pose and focal length estimation is, thus, demonstrated by measuring the relative position and reprojection error of these known control points.

Then two 2D–3D point correspondences for our proposed method, three 2D–3D point correspondences for Kneip’s method, and four 2D–3D point correspondences for the GP4Pf were selected from these known control points to estimate the camera pose and focal length. Finally, we measured the relative position and reprojection of the rest of the control points using binocular vision and reported the average relative positional error between the ground truth and the measured values; the average reprojection error between the position in the real image and the measured value in Table 2.

Table 2. Relative position error and reprojection error for real images.

Method	Proposed Method	Kneip’s Method	GP4Pf
Relative position error/%	0.39	0.47	1.37
Reprojection error/pixel	0.36	0.56	0.78

From Table 2, according to the relative position error and reprojection error, we can observe that our proposed method performed better than Kneip’s method and GP4Pf, which shows our proposed method can work well in real scenarios.

At the beginning of Section 2, we assumed that the skew was zero, the aspect ratio of the pixels was one, and the principal point was the center of the image for our proposed method. Since we do not know the ground truth of the skew and the aspect ratio in real scenarios, the error of these assumptions cannot be directly discussed. However, the relative position and reprojection error in real images can indirectly show that our method can obtain good results under these assumptions. Actually, the relative position error directly reflects the total error introduced by our algorithm model and these assumptions. The

relative position error was 0.39%, which is low and can meet the actual application requirements. We can see the relative positional error includes the error of these assumptions and, therefore, the error of these assumptions was less than 0.39%, which shows that these assumptions can yield good results in a real scenario experiment, even though they are not strictly true.

4. Discussion

Orientation and focal length estimation is one of the key steps in computer vision, photogrammetry, SLAM, and SfM. In this paper we propose an efficient closed form solution to the absolute orientation problem with unknown focal length and two 2D–3D point correspondences. The problem can be decomposed into two sub-problems and can be solved with angle constraints. A quadratic equation of one variable is solved to determine the focal length, and then a geometric approach is used to determine the absolute orientation, which is different from the existing orientation estimation solvers.

4.1. Differences and Advantages

In this paper, our core contribution is to use fewer 3D control points, for both absolute orientation and focal length estimation. With the development of measurement technology and the reduction in cost, more and more devices are being used to obtain partial pose parameters as prior knowledge, which is the reason why we performed our work with a known camera position. Our proposed method only needs two 3D control points and can estimate both pose and focal length. In contrast, the existing P3P solvers need three 3D points and can only estimate camera pose.

Our proposed method uses partial pose parameters and, hence, can use fewer 3D control points. These partial pose parameters, i.e., camera position, are measured with high precision using RTK or total station (e.g., NTS-330R in Section 3), which is a reason why our proposed method performs better in terms of numerical stability and noise sensitivity.

The P3P solvers in previous studies used an iterative algorithm or needed to solve systems of quadratic or quartic polynomial equations; however, our proposed method only uses a geometric approach with angle constraints. This is another reason why our proposed method performs better in terms of numerical stability, noise sensitivity, and computational speed. In addition, the existing P3P solvers all have up to four possible solutions and need an extra point to give a single solution, which is also a main reason why our proposed method has a faster computational speed.

Our proposed method uses the camera position as the prior knowledge, which is different from the existing methods. Therefore, the camera position is important and we have analyzed the effect of error in the camera position on the estimation of the absolute orientation and of the focal length, as shown in Section 3.1.1. In geometric derivation, the camera position error contributes error to the angle in Equation (6) when we estimate the focal length. However, the camera position error is low, because of high-accuracy measurement using RTK or total station, which means that the error of angle in Equation (6) is very low. This is the reason why our proposed method still yields good results even though camera positional error exists.

As shown in Section 3, because of the lower noise sensitivity in rotation and translation error, our proposed method gives better result in terms of the reprojection error. It should be noted that the Harris algorithm [49] was used for feature point extraction in real images, and its precision is below 0.2 pixels. Hence the reprojection error in real images matches that in the synthetic data of a 0.2 pixel noise. In addition, an ideal focal length was used for the synthetic data and a focal length directly written on the lens, which has a small error, was used for real images. This is a reason why the reprojection error with synthetic data was slightly smaller than that in the real images. Finally, the higher precision in focal length and absolute orientation estimation led our proposed method to have results, in terms of the relative position error in binocular vision.

In brief, our proposed method has the following advantages: (1) Only two 3D points are needed to estimate the absolute orientation and focal length; (2) It gives a single solution and has no multi-solution phenomenon; (3) It performs better, in terms of numerical stability, noise sensitivity, computational speed, and robustness to camera position noise; and (4) It obtains better results, both with synthetic data and real images.

4.2. Future Work

Our proposed method has to use a positioning device (e.g., RTK, total station) to obtain the camera position and, as described in Section 1, some existing methods use the known vertical direction to obtain some orientation information using IMUs. Those methods can all use fewer 3D points to estimate camera pose than the existing P3P solvers. This may inspire us to use both camera position and vertical direction for pose and partial intrinsic parameter estimation in the future. This idea may lead to a faster and more efficient method.

Another work that will be completed in the future is to use a camera with a positioning device in practice, such as SfM and 3D reconstruction with the RANSAC algorithm [50]. The superior computational efficiency of our proposed method is particularly suitable as a RANSAC outlier rejection step.

5. Conclusions

We have proposed an efficient closed-form solution to the absolute orientation problem for a camera with unknown focal length from two 2D–3D point correspondences and the camera position. In the original state, the camera frame and the two 2D image points are known, and the world frame and the two 3D control points are also known. However, the 2D–3D point correspondences are unknown in the original state. Our main process is to rotate the original camera frame and world frame to make the camera position, 2D image point, and 3D control point collinear, and then obtain two 2D–3D point correspondences geometrically in the final state. Finally, the absolute orientation can be estimated based on the known camera frame, the known world frame in the original state, and the rotation angles. Before this, the focal length is estimated using angle constraint.

By decomposing the problem into two sub-problems and solving them with angle constraints, only two 2D–3D point correspondences are needed to estimate the focal length and absolute orientation, and a single solution can be given with our method. The geometric derivations are easy to understand and significantly improve the performance. Experimental results show that our proposed method works well with synthetic data and real scenarios. It is particularly suitable for estimating the focal length and orientation of a zooming digital camera with fixed position or with a positioning device mounted on it.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the manuscript.

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

References

1. Terzakis, G.; Lourakis, M. A consistently fast and globally optimal solution to the perspective-n-point problem. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 478–494.
2. Lourakis, M.; Terzakis, G. A globally optimal method for the PnP problem with MRP rotation parameterization. In Proceedings of the International Conference on Pattern Recognition, Milan, Italy, 10–15 January 2021; pp. 3058–3063.
3. Bujnák, M. Algebraic Solutions to Absolute Pose Problems. Ph. D. Thesis, Czech Technical University, Prague, Czech Republic, 2012.
4. Zhou, L.; Kaess, M. An efficient and accurate algorithm for the perspective-n-point problem. In Proceedings of the International Conference on Intelligent Robots and Systems, Macau, China, 3–8 November 2019; pp. 6245–6252.
5. Lepetit, V.; Moreno-Noguer, F.; Fua, P. Epnp: An accurate o(n) solution to the pnp problem. *Int. J. Comput. Vis.* **2009**, *81*, 155. [\[CrossRef\]](#)
6. Zheng, Y.; Kuang, Y.; Sugimoto, S.; Astrom, K.; Okutomi, M. Revisiting the pnp problem: A fast, general and optimal solution. In Proceedings of the IEEE International Conference on Computer Vision, Sydney, Australia, 1–8 December 2013; pp. 2344–2351.
7. Hamel, T.; Samson, C. Riccati observers for the nonstationary PnP problem. *IEEE Trans. Autom. Control.* **2017**, *63*, 726–741. [\[CrossRef\]](#)
8. Youyang, F.; Qing, W.; Yuan, Y.; Chao, Y. Robust improvement solution to perspective-n-point problem. *Int. J. Adv. Robot. Syst.* **2019**, *16*, 1729881419885700. [\[CrossRef\]](#)
9. Ferraz, L.; Binefa, X.; Moreno-Noguer, F. Very fast solution to the PnP problem with algebraic outlier rejection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 501–508.
10. Kneip, L.; Scaramuzza, D.; Siegwart, R. A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, 20–25 June 2011; pp. 2969–2976.
11. Li, J.; Hu, Q.; Zhong, R.; Ai, M. Exterior orientation revisited: A robust method based on lq-norm. *Photogramm. Eng. Remote. Sens.* **2017**, *83*, 47–56. [\[CrossRef\]](#)
12. Gao, X.S.; Hou, X.R.; Tang, J.; Cheng, H.F. Complete solution classification for the perspective-three-point problem. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *25*, 930–943.
13. Sweeney, C.; Flynn, J.; Nuernberger, B.; Turk, M.; Höllerer, T. Efficient computation of absolute pose for gravity-aware augmented reality. In Proceedings of the IEEE International Symposium on Mixed and Augmented Reality, Fukuoka, Japan, 29 September–3 October 2015; pp. 19–24.
14. Cao, M.W.; Jia, W.; Zhao, Y.; Li, S.J.; Liu, X.P. Fast and robust absolute camera pose estimation with known focal length. *Neural Comput. Appl.* **2018**, *29*, 1383–1398. [\[CrossRef\]](#)
15. Kotake, D.; Satoh, K.; Uchiyama, S.; Yamamoto, H. A hybrid and linear registration method utilizing inclination constraint. In Proceedings of the IEEE and ACM International Symposium on Mixed and Augmented Reality, Vienna, Austria, 5–8 October 2005; pp. 140–149.
16. Camposeco, F.; Cohen, A.; Pollefeys, M.; Sattler, T. Hybrid camera pose estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 136–144.
17. Chang, Y.J.; Chen, T. Multi-view 3D reconstruction for scenes under the refractive plane with known vertical direction. In Proceedings of the International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 351–358.
18. Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2003.
19. Zhou, L.; Ye, J.; Kaess, M. A stable algebraic camera pose estimation for minimal configurations of 2D/3D point and line correspondences. In Proceedings of the Asian Conference on Computer Vision, Perth, Australia, 2–6 December 2018; pp. 273–288.
20. Kukelova, Z.; Bujnak, M.; Pajdla, T. Closed-form solutions to minimal absolute pose problems with known vertical direction. In Proceedings of the Asian Conference on Computer Vision, Queenstown, New Zealand, 8–12 November 2010; pp. 216–229.
21. Nistér, D.; Stewénius, H. A minimal solution to the generalised 3-point pose problem. *J. Math. Imaging Vis.* **2007**, *27*, 67–79. [\[CrossRef\]](#)
22. Masselli, A.; Zell, A. A new geometric approach for faster solving the perspective-three-point problem. In Proceedings of the International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014; pp. 2119–2124.
23. Ke, T.; Roumeliotis, S.I. An efficient algebraic solution to the perspective-three-point problem. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7225–7233.
24. Wolfe, W.; Mathis, D.; Sklair, C.; Magee, M. The perspective view of three points. *IEEE Trans. Pattern Anal. Mach. Intell.* **1991**, *13*, 66–73. [\[CrossRef\]](#)
25. Wang, P.; Xu, G.; Wang, Z.; Cheng, Y. An efficient solution to the perspective-three-point pose problem. *Comput. Vis. Image Underst.* **2018**, *166*, 81–87. [\[CrossRef\]](#)
26. DeMenthon, D.; Davis, L.S. Exact and approximate solutions of the perspective-three-point problem. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 1100–1105. [\[CrossRef\]](#)
27. Triggs, B. Camera pose and calibration from 4 or 5 known 3d points. In Proceedings of the Seventh IEEE International Conference on Computer Vision, Corfu, Greece, 20–25 September 1999; Volume 1, pp. 278–284.
28. Zheng, Y.; Sugimoto, S.; Sato, I.; Okutomi, M. A general and simple method for camera pose and focal length determination. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 430–437.

29. Wu, C. P3.5p: Pose estimation with unknown focal length. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 2440–2448.
30. Kanaeva, E.; Gurevich, L.; Vakhitov, A. Camera pose and focal length estimation using regularized distance constraints. In Proceedings of the British Machine Vision Conference, Swansea, UK, 7–10 September 2015; p. 162.
31. Abidi, M.A.; Chandra, T. A new efficient and direct solution for pose estimation using quadrangular targets: Algorithm and evaluation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1995**, *17*, 534–538. [[CrossRef](#)]
32. Bujnak, M.; Kukelova, Z.; Pajdla, T. New efficient solution to the absolute pose problem for camera with unknown focal length and radial distortion. In Proceedings of the Asian Conference on Computer Vision, Queenstown, New Zealand, 8–12 November 2010; pp. 11–24.
33. Josephson, K.; Byrod, M. Pose estimation with radial distortion and unknown focal length. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, USA, 20–25 June 2009; pp. 2419–2426.
34. Kukelova, Z.; Bujnak, M.; Pajdla, T. Real-time solution to the absolute pose problem with unknown radial distortion and focal length. In Proceedings of the IEEE International Conference on Computer Vision, Sydney, Australia, 1–8 December 2013; pp. 2816–2823.
35. Guo, Y. A Novel Solution to the P4P Problem for an Uncalibrated Camera. *J. Math. Imaging Vis.* **2013**, *45*, 186–198. [[CrossRef](#)]
36. Kalantari, M.; Hashemi, A.; Jung, F.; Guédon, J.P. A new solution to the relative orientation problem using only 3 points and the vertical direction. *J. Math. Imaging Vis.* **2011**, *39*, 259–268. [[CrossRef](#)]
37. Emanuele, G.; Pietro, M.; Pugliese, P. Camera and inertial sensor fusion for the PnP problem: Algorithms and experimental results. *Mach. Vis. Appl.* **2021**, *32*, 90.
38. D’Alfonso, L.; Garone, E.; Muraca, P.; Pugliese, P. On the use of IMUs in the PnP Problem. In Proceedings of the International Conference on Robotics and Automation, Hong Kong, China, 31 May–5 June 2014; pp. 914–919.
39. D’Alfonso, L.; Garone, E.; Muraca, P.; Pugliese, P. On the use of the inclinometers in the PnP Problem. In Proceedings of the European Control Conference, Zurich, Switzerland, 17–19 July 2013; pp. 4112–4117.
40. Merckel, L.; Nishida, T. Solution of the perspective-three-point problem. In Proceedings of the International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, Kyoto, Japan, 26–29 June 2007; pp. 324–333.
41. Sweeney, C.; Flynn, J.; Turk, M. Solving for relative pose with a partially known rotation is a quadratic eigenvalue problem. In Proceedings of the International Conference on 3D Vision, Kyoto, Japan, 8–11 December 2014; pp. 483–490.
42. Albl, C.; Kukelova, Z.; Pajdla, T. Rolling shutter absolute pose problem with known vertical direction. In Proceedings of the Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 3355–3363.
43. Hee Lee, G.; Pollefeys, M.; Fraundorfer, F. Relative pose estimation for a multi-camera system with known vertical direction. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 540–547.
44. D’Alfonso, L.; Garone, E.; Muraca, P.; Pugliese, P. P3P and P2P Problems with known camera and object vertical directions. In Proceedings of the Mediterranean Conference on Control and Automation, Crete, Greece, 25–28 June 2013; pp. 444–451.
45. Horanyi, N.; Kato, Z. Generalized pose estimation from line correspondences with known vertical direction. In Proceedings of the International Conference on 3D Vision, Qingdao, China, 10–12 October 2017; pp. 244–253.
46. Merckel, L.; Nishida, T. Evaluation of a method to solve the perspective-two-point problem using a three-axis orientation sensor. In Proceedings of the IEEE International Conference on Computer and Information Technology, Khulna, Bangladesh, 25–27 December 2008; pp. 862–867.
47. Aratani, S.; Uchiyama, S.; Satoh, K.; Endo, T. Position and Orientation Measurement Method and Apparatus. U.S. Patent 7,698,094, 13 April 2010.
48. Guo, K.; Ye, H.; Gu, J.; Chen, H. A Novel Method for Intrinsic and Extrinsic Parameters Estimation by Solving Perspective-Three-Point Problem with Known Camera Position. *Appl. Sci.* **2021**, *11*, 6014. [[CrossRef](#)]
49. Harris, C.; Stephens, M. A combined corner and edge detector. In Proceedings of the Alvey Vision Conference, Manchester, UK, 31 August–2 September 1988; pp. 147–151.
50. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **1981**, *24*, 381–395. [[CrossRef](#)]