Geometry-Based Camera Calibration Using Closed-Form Solution of Principal Line

Jen-Hui Chuang, *Senior Member, IEEE*, Chih-Hui Ho, Ardian Umam, Hsin-Yi Chen, Jenq-Neng Hwang, *Fellow, IEEE*, and Tai-An Chen

***Abstract*— Camera calibration is a crucial prerequisite in many applications of computer vision. In this paper, a new geometry-based camera calibration technique is proposed, which resolves two main issues associated with the widely used Zhang's method: (i) the lack of guidelines to avoid outliers in the computation and (ii) the assumption of fixed camera focal length. The proposed approach is based on the closed-form solution of principal lines with their intersection being the principal point while each principal line can concisely represent relative orientation/position (up to one degree of freedom for both) between a special pair of coordinate systems of image plane and calibration pattern. With such analytically tractable image features, computations associated with the calibration are greatly simplified, while the guidelines in (i) can be established intuitively. Experimental results for synthetic and real data show that the proposed approach does compare favorably with Zhang's method, in terms of correctness, robustness, and flexibility, and addresses issues (i) and (ii) satisfactorily.**

***Index Terms*— Camera calibration, geometry-based method, closed-form solution, vanishing point, principal lines, principal point, non-fixed focal length, outlier removal.**

I. Introduction

CAMERA calibration is a crucial step in many 3D vision applications, such as robotic navigation [1], depth map estimation [2], and 3D reconstruction [3]. Camera calibration establishes the geometric relation between 3D world coordinate system (**WCS**) and the 2D image plane of camera by finding extrinsic and intrinsic camera parameters. The extrinsic parameters, which define the translation and orientation of the camera with respect to the world frame, transform 3D **WCS** into 3D camera coordinate system (**CCS**). On the other hand, intrinsic parameters, including principal point, focal length and skewness factor, transform 3D **CCS** to 2D image plane of the camera.

Camera calibration can be roughly classified into two categories: photogrammetric [4], [5] and self-calibration [6], [7]. The former methods perform calibration based on sufficient measurements of 3D points with known correspondence in the scene and assume some calibration objects/templates are available. However, in a large-scale camera network, it is hard to acquire this kind of measurement or available information for each camera. Therefore, many methods have been proposed to self-calibrate the camera automatically based on certain assumptions on the online camera scenes [8], [9]. On the other hand, both types of method can also accomplish camera calibration through vanishing points-based methods [10], [11] or pure rotation approaches [11].

Zhang's method [12] is considered as the most widely used photogrammetric calibration method due to its low cost and flexibility, which only needs to use a printed pattern of checkerboard pasted on a flat surface, and captured by the camera with at least two different orientations. However, two main issues are associated with such an approach: (i) The checkerboard patterns are usually placed randomly and used all together, without a systematic procedure to screen out ill-posed patterns, and (ii) all intrinsic parameters are assumed to be fixed throughout the pattern capturing process.For Issue (i), inconsistent or unreasonable calibration results may be generated from different sets of checkerboard patterns for the same camera, as the two (intrinsic and extrinsic) apparently independent sets of parameters are simultaneously calculated via purely algebraic formulations. Moreover, the complexity of such formulations, which are not established for the original intrinsic parameters but for their nonlinear transformations, also greatly decrease the feasibility of the development of more general formulations to resolve Issue (ii) which may occur in practice.[1]

Tan *et al.* [13] partly address issue (i) by first replacing physical checkerboard patterns with virtual ones displayed on a screen to minimize localization error of point features (the corner points) resulting from a blurry image due to hand motion. Then, by conveniently using different sets of virtual patterns in the experiments, appropriate poses of these virtual patterns are suggested: the selected point features should distribute uniformly across the image captured by the camera.

Manuscript received November 8, 2019; revised November 16, 2020; accepted December 15, 2020. Date of publication January 8, 2021; date of current version February 5, 2021. This work was supported in part by the Ministry of Science and Technology, Taiwan, R.O.C., under Contract MOST 109-2634-F-009-015 and in part by the Ministry of Economic Affairs, Taiwan, R.O.C., under Contract MOEA 109-EC-17-A-02-S5-008. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Junsong Yuan. *(Corresponding author: Jen-Hui Chuang.)*

Jen-Hui Chuang and Hsin-Yi Chen are with the Department of Computer Science, National Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: jchuang@cs.nctu.edu.tw).

Chih-Hui Ho is with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA.

Ardian Umam is with the School of Electrical Engineering and Informatics, Bandung Institute of Technology, Bandung 40132, Indonesia.

Jenq-Neng Hwang is with the Department of Electrical and Computer Engineering, University of Washington, Seattle, WA 98115 USA.

Tai-An Chen is with E.Sun Financial Holding Company Ltd., Hsinchu 300, Taiwan.

Digital Object Identifier 10.1109/TIP.2020.3048684

[1]For example, one may take pictures of a stationary calibration pattern from different distances while adjusting zooming levels for proper image size.

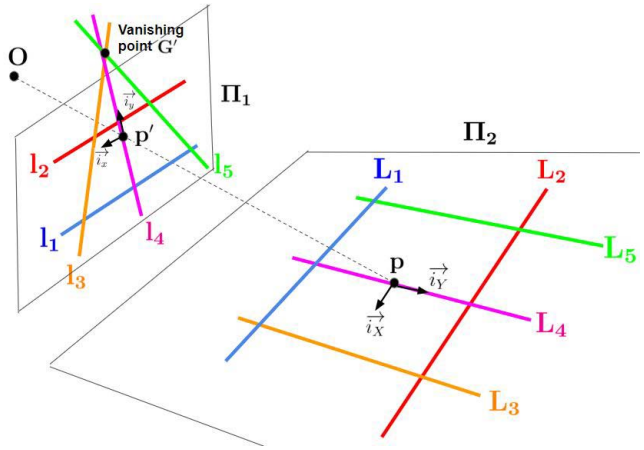

Fig. 1. Special geometric relationship of **WCS** and **IPCS** for 1D localization of principal point with principal line.

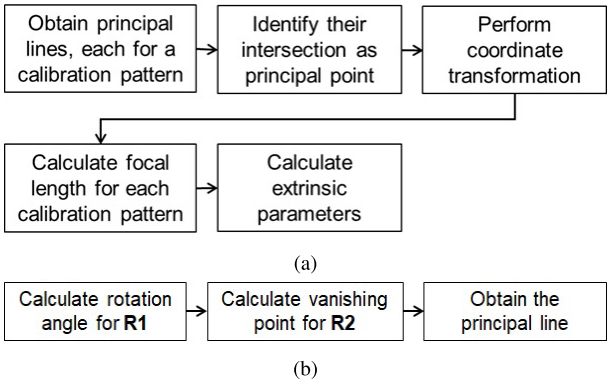

Fig. 2. (a) Flowchart of the proposed calibration process, (b) the detailed flowchart for getting the principal line.

Nonetheless, such conclusion may not be decisive as different suggestions for appropriate poses of calibration object/pattern also exist [14], [15].

To partly address Issue (ii), only the principal point is assumed fixed and estimated in [16] and [17], with the skewness factor ignored and the focal length not assumed to be fixed. The estimation is based on the establishment of a coordinate system of the image plane (**IPCS**), which has a special geometric relationship to a corresponding **WCS**, with the X-Y plane (***the calibration plane***) of the latter containing the calibration pattern. Specifically, their relationship can be described by the following rules of geometry, as depicted in Figure 1, wherein Π_1 and Π_2 are the image plane and the calibration plane, respectively, and the projection center O is colinear with the line connecting the origins of **IPCS** and **WCS**:

R1. $\overrightarrow{i_X}$ is parallel to intersection of Π_1 and Π_2, and $\overrightarrow{i_X} /\!/ \overrightarrow{i_x}$.

R2. Image of $\overrightarrow{i_Y}$ (colinear with $\overrightarrow{i_y}$) is perpendicular to $\overrightarrow{i_x}$.

For **R1**, it is not hard to see that only lines parallel to $\overrightarrow{i_X}$, e.g., L_1 and L_2, will have their images parallel to $\overrightarrow{i_x}$. On the other hand, the line containing $\overrightarrow{i_y}$ in **R2**, e.g., l_4 in Figure 1, will pass through the vanishing point of images of lines perpendicular to $\overrightarrow{i_X}$, e.g., L_3,L_4 and L_5, and correspond to the axis of symmetry of them. Moreover, it is shown in [16] and [17] that such image line feature, called ***principal line*** in this paper, will also pass through the principal point p, i.e., the intersection of the optical axis and the image plane. Therefore, the camera principal point can be identified as the intersection of principal lines obtained for a set of calibration planes of different poses.

In this paper, a new geometry-based camera calibration is proposed, with its flowchart shown in Figure 2. Our analytically tractable technique is based on **R1** and **R2** described in Figure 1 and has the following desirable features:

F1. Efficiency —While the principal line of each calibration plane is obtained empirically by analyzing a sequence of planar image patterns in [16] and [17] before **R1** and **R2** are achieved, equation of the principal line is obtained in closed form in this paper for each calibration pattern.

F2. Completeness —By assuming the circular symmetry[2] of the imaging system, the proposed approach can derive all intrinsic parameters, while extrinsic ones can be found readily for each calibration plane if needed.

F3. Robustness/accuracy —Based on the geometry associated with each corresponding principal line, effective way of identifying ill-posed calibration planes is developed so that the robustness and accuracy of the calibration can be greatly improved by discarding such outliers, resolving Issue (i) of Zhang's method [12].

F4. Flexibility —Without assuming a fixed camera focal length (FL), the proposed approach can find different values of FL used in capturing calibration patterns of different pose, and successfully address Issue (ii) of Zhang's method.

The rest of this paper will be organized as follows. In the next section, a closed-form solution of the principal line is first derived from the homography matrix,

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}, \tag{1}$$

which is obtained from corresponding point features on a calibration plane and the image plane. A set of such line features can then be used to determine the principal point and, subsequently, the rest camera parameters. Possible ways of identifying ill-posed calibration planes are also described. In Section III, experimental results on both synthetic and real data are provided to demonstrate the superiority of proposed approach in robustness, accuracy, and flexibility, as well as the capability of performance improvements via the identification of inappropriate calibration planes. Finally, some concluding remarks will be given in Section IV.

II. Derivation of Closed-Form Solutions of Camera Parameters

In this section, camera parameters are derived analytically via the establishment of the special geometric relation of **IPCS** and **WCS** described in **F2**. In Section II-A, the derivation of principal line with a single image of the calibration pattern is elaborated, which includes the establishment of $\overrightarrow{i_X}$ and $\overrightarrow{i_x}$ of **R1** via rotation, followed by finding the principal line

[2]Such assumption is quite reasonable nowadays for a variety of cameras, as one can see later in Sec. III.

(and $\overrightarrow{i_y}$ of **R2**) using the vanishing point. In Section II-B, closed-form solutions of intrinsic parameters are derived, which include the derivation of the principal point from a set of principal lines, followed by the derivation of the camera focal length by using the principal point to establish the special **WCS-IPCS** pair described in Figure 1. Finally, extrinsic parameters associated with each calibration pattern can be obtained.

A. Deriving Closed-Form Solution of the Principal Line

In this section, in order to simplify the derivation of principal line on the image plane, the reorientation of a unit square on the calibration plane is considered. Specifically, the rotation of the square which results in a trapezoidal shape of its image is derived in closed form. Subsequently, the direction of the two bases of the trapezoid is identified as the direction of $\overrightarrow{i_X}$ and $\overrightarrow{i_x}$ of **R1**, while the principal line is identified as the line orthogonal to the bases and passing through the intersection of the two legs of the trapezoid.

*1) Finding the Direction of $\overrightarrow{i_X}$ and $\overrightarrow{i_x}$ of **R1**:* Assume a square $ABCD$ in **WCS**, as shown in Figure 3 (a), is captured by a camera, with $A'B'C'D'$ being its image in the **IPCS**. Since $ABCD$ and $A'B'C'D'$ are planar surfaces, a homography matrix **H** can be used to represent their relationship. The goal of this subsection is to derive the angle θ in Figure 3 (b) such that $\overrightarrow{A'B'} \,//\, \overrightarrow{C'D'} \,//\, \overrightarrow{i_X} \,//\, \overrightarrow{i_x}$, as shown in Figure 3 (c). Assume **R** is the rotation matrix associated with angle θ, and by rotating rectangle $ABCD$ with $A = [0\ 0\ 1]^T$, $B = [1\ 0\ 1]^T$, $C = [0\ 1\ 1]^T$ and $D = [1\ 1\ 1]^T$ with respect to A using R, i.e.,

$$R = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \triangleq \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

$$A = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad B = \begin{bmatrix} c \\ -s \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} s \\ c \\ 1 \end{bmatrix}, \quad D = \begin{bmatrix} c+s \\ c-s \\ 1 \end{bmatrix},$$

followed by multiplying with H to transform the rotated rectangle from **WCS** to **IPCS**, we have

$$A' = \begin{bmatrix} \dfrac{h_3}{h_9} \\ \dfrac{h_6}{h_9} \\ 1 \end{bmatrix}, \quad B' = \begin{bmatrix} \dfrac{h_1\ c - h_2\ s + h_3}{h_7\ c - h_8\ s + h_9} \\ \dfrac{h_4\ c - h_5\ s + h_6}{h_7\ c - h_8\ s + h_9} \\ 1 \end{bmatrix},$$

$$C' = \begin{bmatrix} \dfrac{h_1\ s + h_2\ c + h_3}{h_7\ s + h_8\ c + h_9} \\ \dfrac{h_4\ s + h_5\ c + h_6}{h_7\ s + h_8\ c + h_9} \\ 1 \end{bmatrix},$$

$$D' = \begin{bmatrix} \dfrac{h_1(c+s) + h_2(c-s) + h_3}{h_7(c+s) + h_8(c-s) + h_9} \\ \dfrac{h_4(c+s) + h_5(c-s) + h_6}{h_7(c+s) + h_8(c-s) + h_9} \\ 1 \end{bmatrix},$$

in homogeneous coordinates. By defining $k_1 = h_9(h_7\ c - h_8\ s + h_9)$ and $k_2 = (h_7\ s + h_8\ c + h_9)(h_7(c+s) + h_8(c-s) + h_9)$,

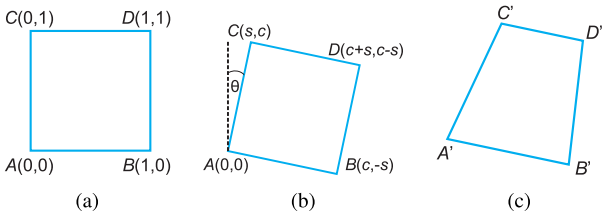

Fig. 3. Real world plane projection of a rectangle (a) to the corresponding camera image (c), based on rotation given in (b).

$\overrightarrow{A'B'}$ and $\overrightarrow{C'D'}$ can be represented as

$$\overrightarrow{A'B'} = \begin{bmatrix} \dfrac{(h_3\ h_7 - h_1\ h_9)c + (h_2\ h_9 - h_3\ h_8)s}{k_1} \\ \dfrac{(h_6\ h_7 - h_4\ h_9)c + (h_5\ h_9 - h_6\ h_8)s}{k_1} \end{bmatrix} = \begin{bmatrix} \dfrac{Uc + Vs}{k_1} \\ \dfrac{Xc + Ys}{k_1} \end{bmatrix} \tag{2}$$

and

$$\overrightarrow{C'D'} = \begin{bmatrix} \dfrac{(h_3\ h_7 - h_1\ h_9)c + (h_2\ h_9 - h_3\ h_8)s + (h_2\ h_7 - h_1\ h_8)}{k_2} \\ \dfrac{(h_6\ h_7 - h_4\ h_9)c + (h_5\ h_9 - h_6\ h_8)s + (h_5\ h_7 - h_4\ h_8)}{k_2} \end{bmatrix} = \begin{bmatrix} \dfrac{Uc + Vs + W}{k_2} \\ \dfrac{Xc + Ys + Z}{k_2} \end{bmatrix},$$

with

$$U \triangleq (h_3h_7 - h_1h_9), \quad V \triangleq (h_2h_9 - h_3h_8),$$
$$W \triangleq (h_2h_7 - h_1h_8), \quad X \triangleq (h_6h_7 - h_4h_9),$$
$$Y \triangleq (h_5h_9 - h_6h_8), \quad Z \triangleq (h_5h_7 - h_4h_8).$$

Since $\overrightarrow{A'B'}$ is parallel to $\overrightarrow{C'D'}$, we have

$$(Uc + Vs)(Xc + Ys + Z) = (Xc + Ys)(Uc + Vs + W)$$

and

$$\tan(\theta) = \frac{s}{c} = \frac{UZ - XW}{YW - VZ} = \frac{h_7}{h_8} \tag{3}$$

Thus, we can obtain the desired rotation angle,

$$\theta = tan^{-1}(\frac{s}{c}) = tan^{-1}(\frac{h_7}{h_8}). \tag{4}$$

*2) Finding the Principal Line (and $\overrightarrow{i_y}$ of **R2**):* After the rotation shown in Figure 3(b), the intersection between lines $\overrightarrow{AC}$ $(cx - sy = 0)$ and $\overrightarrow{BD}$ $(cx - sy - (c^2 + s^2) = 0)$ of the square $ABCD$, denoted by G, can be calculated by using the cross product in homogeneous coordinates,

$$G = \begin{bmatrix} c \\ -s \\ 0 \end{bmatrix} \times \begin{bmatrix} c \\ -s \\ -(c^2 + s^2) \end{bmatrix} = \begin{bmatrix} s \\ c \\ 0 \end{bmatrix} \tag{5}$$

Then, the intersection point in **IPCS**, denoted by G', which is the vanishing point in **IPCS**, can be obtained by

$$G' = HG = \begin{bmatrix} \dfrac{h_1\ s + h_2\ c}{h_7\ s + h_8\ c} & \dfrac{h_4\ s + h_5\ c}{h_7\ s + h_8\ c} & 1 \end{bmatrix}^T. \tag{6}$$

The principal line can be calculated by finding a line that is perpendicular to $\overrightarrow{A'B'}$ and passing through the vanishing point G'. By substituting $k_1 = 1$[3] and $h_9 = 1$, (2) becomes

$$\overrightarrow{A'B'} = \begin{bmatrix} (h_3\ h_7 - h_1)c + (h_2 - h_3\ h_8)s \\ (h_6\ h_7 - h_4)c + (h_5 - h_6\ h_8)s \end{bmatrix}. \tag{7}$$

Followed by substituting (4) to (7), we have

$$\overrightarrow{A'B'} = \begin{bmatrix} -h_1\ c + h_2\ s \\ -h_4\ c + h_5\ s \end{bmatrix} = \begin{bmatrix} -h_1\ h_8 + h_2\ h_7 \\ -h_4\ h_8 + h_5\ h_7 \end{bmatrix}.$$

The principal line can then be expressed by

$$a'u' + b'v' + c' = 0, \tag{8}$$

where $a' = (-h_1\ h_8 + h_2\ h_7)$, $b' = (-h_4\ h_8 + h_5\ h_7)$ and c' can be solved by plugging (6) into (8). (See Appendix A.)

Thus, a principal line is derived using the homography matrix alone. By repeating the abovementioned procedure for all images of calibration pattern, multiple principal lines can be obtained and used to find the principal point, as discussed next.

B. Deriving Closed-Form Solution of Intrinsic Parameters

1) Derivation of the Principal Point: Given n principal lines from section II-A from different calibration poses, with each lines being represented by $a'_i u' + b'_i v' + c'_i = 0$ (8), the principal point (u'_0, v'_0) in **IPCS** can now be estimated as their intersection via the least squares solution, or

$$\begin{bmatrix} u'_0 \\ v'_0 \end{bmatrix} = (D^{\mathsf{T}} D)^{-1} D^{\mathsf{T}} C, \tag{9}$$

where

$$D = \begin{bmatrix} a'_1\ a'_2\ \dots\ a'_n \\ b'_1\ b'_2\ \dots\ b'_n \end{bmatrix}^T, C = \begin{bmatrix} -c'_1\ -c'_2\ \dots\ -c'_n \end{bmatrix}^T.$$

2) Derivation of the Focal Length: Considering the original **IPCS** and **WCS** coordinate systems, the relationship between point p on the calibration plane ($Z = 0$) and its image p' can be expressed as

$$\begin{aligned} p' &= M_{int} M_{ext} p = M_{int} \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix} p \\ &= \begin{bmatrix} f & 0 & u'_0 & 0 \\ 0 & f & v'_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_X \\ r_{21} & r_{22} & r_{23} & t_Y \\ r_{31} & r_{32} & r_{33} & t_Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 0 \\ 1 \end{bmatrix}, \end{aligned} \tag{10}$$

or with 2D coordinates

$$\begin{aligned} p' &= \begin{bmatrix} f & 0 & u'_0 \\ 0 & f & v'_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_X \\ r_{21} & r_{22} & t_Y \\ r_{31} & r_{32} & t_Z \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \\ &\triangleq \tilde{M}_{int} \tilde{M}_{ext} p \\ &= \begin{bmatrix} fr_{11} + u'_0\ r_{31} & fr_{12} + u'_0 r_{32} & ft_X + u'_0 t_Z \\ fr_{21} + u'_0\ r_{31} & fr_{22} + v'_0 r_{32} & ft_Y + v'_0 t_Z \\ r_{31} & r_{32} & t_Z \end{bmatrix} p \\ &= Hp, \end{aligned} \tag{11}$$

where H is the 3×3 homography matrix in (1).

In this section, the formulation of focal length estimation is greatly simplified by transforming **IPCS** and **WCS** into the special geometry depicted in Figure 1, with the optical axis passing through their origins. Note that principal point $p'_0 = \left[u'_0, v'_0, 1\right]^{\mathrm{T}}$ in **IPCS** is derived in Section II-B1, whereas its corresponding point $p_0 = \left[u_0, v_0, 1\right]^{\mathrm{T}}$ in **WCS** can be obtained using (11). Thus, the coordinate transformation is performed by *shifting* these origins accordingly, followed by *rotating* both **IPCS** and **WCS** according to the principal line (8) and its counterpart in **WCS**, respectively. Therefore, the following formulation can be established from (11):

$$\begin{aligned} p'_{sr} &= H_{1r} H_{1s} H p \\ &= H_{1r} H_{1s} H (H_{2r} H_{2s})^{-1} (H_{2r} H_{2s}) p \\ &= H_{1r} H_{1s} H (H_{2r} H_{2s})^{-1} p_{sr} \\ &\triangleq H^{new} p_{sr}, \end{aligned} \tag{12}$$

where p'_{sr} and p_{sr} are points of the transformed new **IPCS** and **WCS**, respectively, with

$$H_{1r} = \begin{bmatrix} a' & b' & 0 \\ -b' & a' & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad H_{1s} = \begin{bmatrix} 1 & 0 & -u'_0 \\ 0 & 1 & -v'_0 \\ 0 & 0 & 1 \end{bmatrix}$$

while H_{2r} and H_{2s} are similar to H_{1r} and H_{1s}, but with a', b', u'_0, and v'_0 replaced by a, b, u_0, and v_0, respectively. Note that $\left[a'\ b'\right]$ and $\left[a\ b\right]$ are coefficients associated with (8) and its counterpart in **WCS**, respectively, with $\left[a\ b\ c\right]^T = H^T \left[a'\ b'\ c'\right]^T$.

On the other hand, for the new **IPCS** and **WCS**, it is easy to see that we will have rotation matrix

$$\begin{aligned} \mathbf{R^{new}} &= R_Z(\alpha) R_Y(\beta) R_X(\gamma) \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\gamma) & -sin(\gamma) \\ 0 & sin(\gamma) & cos(\gamma) \end{bmatrix}, \end{aligned} \tag{13}$$

or $\alpha = \beta = 0$, for their relative orientation, and

$$\mathbf{T^{new}} = \begin{bmatrix} 0 & 0 & t_z^{new} \end{bmatrix} \tag{14}$$

for their relative location. By comparing (13) and (14) with (11), we have, up to a scaling factor s,[4]

$$\begin{aligned} H^{new} &= \begin{bmatrix} h_{11}^{new} & h_{12}^{new} & h_{13}^{new} \\ h_{21}^{new} & h_{22}^{new} & h_{23}^{new} \\ h_{31}^{new} & h_{32}^{new} & h_{33}^{new} \end{bmatrix} \\ &= s \begin{bmatrix} f & 0 & 0 \\ 0 & fcos(\gamma) & 0 \\ 0 & sin(\gamma) & t_Z^{new} \end{bmatrix}, \end{aligned} \tag{15}$$

and

$$\gamma = \cos^{-1}\left(\frac{h_{22}^{new}}{h_{11}^{new}}\right) \tag{16}$$

$$t_z^{new} = \frac{h_{33}^{new}}{s} \tag{17}$$

$$f = \frac{h_{11}^{new}}{s} \tag{18}$$

with $s = \frac{h_{32}^{new}}{sin(\gamma)}$.

[3]This is obtained by substituting (4) to k_1

[4]Equivalently, but less directly, homography matrix similar to that in (15) can be obtained by finding coordinates of point features for the new **WCS-IPCS** before such matrix can be estimated.

C. Derivation of Extrinsic Parameters

In the previous subsection, analytic expressions of two extrinsic parameters are derived in (16) and (17). As mentioned before, we have $t_x = t_y = \alpha = \beta = 0$ for the other four parameters for the new **WCS-IPCS** pair. In fact, only five parameters (15) are needed to completely specify the relative position and orientation of the two planes. In particular, the relative position of the new **WCS-IPCS**, can be represented by the distance (t_z^{new}) between the two origins, with $t_x = t_y = 0$, while their relative orientation can be represented by: (a) the azimuth angle of the principal line, i.e., $tan^{-1}(\frac{b'}{a'})$ and (b) the elevation angle (γ) between the two planes. Such concise formulation of extrinsic parameters is more intuitive and useful. For example, (b) can be used to screen out calibration patterns of bad poses, while a set of good but redundant patterns can be identified using (a), as will be discussed in the next subsection.

On the other hand, the set of extrinsic parameters similar to the found in Zhang's method for the original **IPCS** and **WCS** can also be obtained if needed. Following the notation in (11) and plugging the solution of principal point (9) and focal length (18) to $\tilde{M}_{int}$, $\tilde{M}_{ext}$ can be solved as

$$\tilde{M}_{ext} = \tilde{M}_{int}^{-1} \frac{H}{s}. \tag{19}$$

However, there are minor differences between $\tilde{M}_{ext}$ and M_{ext} as r_{13}, r_{23} and r_{33} are not defined in $\tilde{M}_{ext}$ (see $\tilde{M}_{ext}$ in (10)). To solve M_{ext} in (10), we first rewrite rotation matrix $\mathbf{R} = [\mathbf{r_1}\ \mathbf{r_2}\ \mathbf{r_3}]$, with $\mathbf{r_i}$ being a column vector. By using the property of rotation matrix $\mathbf{R}$, $\mathbf{r_3}$ can be calculated as $\mathbf{r_3} = \mathbf{r_1} \times \mathbf{r_2}$. Thus, the extrinsic parameters M_{ext} for the original **IPCS** and **WCS** pair can be solved.

D. Identification of Ill-Posed Calibration Patterns

While the proposed approach can generate similar calibration results as with Zhang's method, we will show that our results are more robust and accurate (in the next section) and can be obtained with similar time complexity (in Appendix B). On the other hand, as the geometry of the complete set of derived principal lines, essentially reveal the whole situation of the estimation of principal point, i.e., how far off they are deviated from intersecting at a single point, it is possible to identify an ill-posed calibration pattern by examining its corresponding principal line.

While several methods can be employed to remove the above outlier, e.g., in [18]–[23], one simple approach adopted in our experiments is to remove principal lines with excessive root-mean-square error (RMSE) in the estimation of principal point, as the method can be implemented easily by simply setting the maximum RMSE allowed for each principal line. Note that these outliers may result from ill-posed calibration pattern as well as large digitization errors for distant patterns or motion blurring. (Note that γ in (16), whose value cannot be estimated accurately for ill-posed real pattern, is used to identify and remove ill-posed patterns for our experiments with synthetic data.)

On the other hand, a set of good calibration patterns, i.e., those with small RMSEs, may still correspond to a poor estimation of principal point. Such situation occurs when nearly parallel (1D) principal lines are used to estimate the principal point in the (2D) image plane. Nonetheless, such situation can be identified easily, even visually, by examining the angular extent of the complete set of derived principal lines.[5]

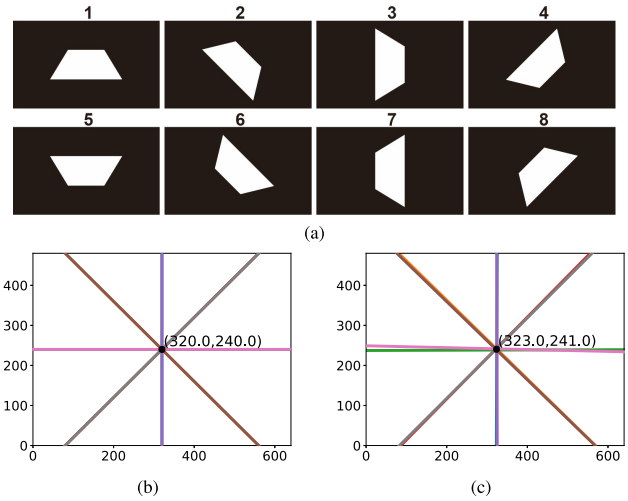

Fig. 4. (a) Eight images obtained for $\gamma = \Delta\alpha = 45°$ (see text). (b) Eight (but merged into four) principal lines obtained for (a). (c) Principal lines obtained with additive noises.

III. Experimental Results and Discussions

In this section, two sets of experimental results will be provided. First, synthetic data are used to: (i) validate the correctness of analytic expressions derived in Section II, and (ii) compare the accuracy of estimated parameters result with Zhang's, with and without noises added to point features in the image plane, as ground truth (GT) are available. Then, calibration results for real data are provided for more comprehensive demonstration of the proposed method. Possible ways of improving the calibration results by screening out inappropriate calibration patterns are also provided for both cases.

A. Performance Evaluation Using Synthetic Data

In this sub-section, correctness of analytic expressions derived in Section II-B1 for the principal point will first be verified using synthetic data described in the following. It is assumed for simplicity, unless specified otherwise, that the camera optical axis is passing through the origin of **WCS** whose X-Y plane, e.g., the calibration plane shown in Figure 1, has a fixed rotation angle γ with respect to $\overrightarrow{iX}$ and a fixed rotation angle β with respect to $\overrightarrow{iY}$, with additional calibration planes obtained by rotating the plane, each time by $\Delta\alpha$, with respect to $\overrightarrow{iZ}$. Figure 4(a) shows a set of images obtained for $\gamma = \Delta\alpha = 45°$ and $\beta = 0°$ with the principal point located exactly at the image center, wherein four corners of a square calibration pattern are used to derive elements of $\mathbf{H}$ in (1) for each (virtual) 640×480 image.

Multiple sets of images under different extrinsic parameters (γ, β, t_x, t_y and t_z), perturbation (with/without noise) and

[5]Please refer to Figure 9 in the next section for an example of such situation.

TABLE I
COMPARISON OF FULL CAMERA CALIBRATION (FIXED FOCAL LENGTH)

Set	Given params.*		Estimation errors**								
			Ours				Zhang's				
	γ, β	t_x, t_y, t_z	ΔPP	ΔFL	ΔR	ΔT	ΔPP	ΔFL_x	ΔFL_y	ΔR	ΔT
1	45, 0	0, 0, 35	**3.9**	**3.13**	**1.44**	**0.79**	10.2	36.34	36.68	1.67	4.23
2	45, 5	2, 3, 35	**4.3**	**3.08**	**1.45**	**0.82**	27.9	16.28	18.60	3.98	2.99
3	Set 2 *w/* 4 bad poses		**5.9**	**13.10**	**1.41**	**2.64**	24.6	17.48	19.44	3.46	3.50
4	Set 3 *w/o* bad poses		**6.1**	**4.29**	**1.50**	**0.93**	N/A	N/A	N/A	N/A	N/A

* Ground truth principal point: $(320, 240)$; ground truth focal length: 400.
** All values are the average of 20 repetitions.

different focal lengths (FL) are then collected for evaluating the performance of the proposed method and Zhang's method. The evaluation metrics for PP, FL, rotation R [24] and translation T are based on the difference between the estimation and the ground truth, and can be defined as

$$\Delta PP = ||PP_{GT} - PP_{est}||_2, \tag{20}$$

$$\Delta FL = |FL_{GT} - FL_{est}|, \tag{21}$$

$$\Delta R = cos^{-1}(\frac{\mathbf{Tr}(R_{GT}R_{est}^T) - 1}{2}), \tag{22}$$

$$\Delta T = ||T_{GT} - T_{est}||_2. \tag{23}$$

1) Synthetic Data Without Noise: Under the noise-free condition, Figure 4(b) shows the resultant eight (perfect) principal lines obtained in closed-form using (8), which intersect exactly at the principal point. One can see that due to the symmetry in their orientations, only four (pairs of) principal lines can be seen from the illustration. Perfect estimation results for focal length (FL) and other extrinsic parameters are obtained for both the proposed and Zhang's method for this simple setting. Note that real numbers are used to represent all numerical values during the associated computations, while the input images and the final results shown in Figure 4 are illustrated with virtual 640×480 images.

2) Synthetic Data With Noise and Variation of Pose: To investigate the robustness of the proposed approach, noises are added to the point features used to find $\mathbf{H}$ in (1), which are the only source of interference that may affect the correctness of each principal line. Figure 4(c) illustrates calibration results similar to those shown in Figure 4(b), with noises uniformly distributed between ± 1.0 pixels added to x and y coordinates of corner points shown in Figure 4(a) to simulate point feature localization errors resulted from image digitization.

For more systematic error analysis, and also taking into account the influence from different poses of the calibration plane, similar simulations are performed for two different noise levels with γ ranging from 5° to 85° (with $\Delta\gamma = 5°$ and $\Delta\alpha = 45°$), and repeated 20 times for each pose of the calibration plane. Because of the simplicity of the proposed approach, a total of $17 \times 20 = 340$ principal points are estimated in 4.5 seconds for each noise level, including the analysis of $340 \times 8 = 2720$ images for generating the same number of principal lines.

Figures 5(a) and (b) show means and standard deviations of estimation error (in image pixels), respectively, for the above simulation. It readily observable that larger noises will result in less accurate calibration results which are also less robust.

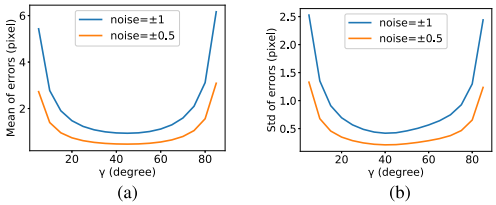

Fig. 5. (a) Mean values and (b) standard deviations of estimation errors obtained with the proposed approach for locating the principal point 20 times with two different noise levels, and for different poses of calibration plane.

Aside from the statistical comparison, we found that better results are generated for poses of calibration plane away from two degenerated conditions of Figure 1, i.e., $\Pi_1 \parallel \Pi_2$ and $\Pi_1 \perp \Pi_2$. Moreover, according to the results in Figure 5, it is reasonable to suggest that: (i) the best values of γ are around 45°. Based on such observation, γ is selected to be approximately 45° for the following experiments. Furthermore, as the principal point is derived via least square solution (9) from all the principal lines, it is also suggested that: (ii) it is better to distribute α uniformly between 0° and 180°.

3) Full Camera Calibration for Fixed Focal Length: As for a more complete error analysis for the estimation of all parameters of a camera with fixed focal length, more general **WCS-IPCS** configurations (also with $\Delta\alpha = 45°$) are considered, which correspond to the first two datasets listed in Table I, with additive noises of ± 1.0 pixels. One can see that smaller estimation errors are achieved for all parameters by the proposed approach, possibly due to the simplicity of the formulations established in Section II, compared with those given in [12]. (Note that ΔFL is computed for the average value of FLs each obtained for a single input image for the proposed approach, and for the estimated values of f_x and f_y, respectively, for Zhang's method.)

As for the robustness of camera calibration, it is possible for the proposed geometric-based approach to improve the parameter estimation by screening out calibration pattern with bad poses. For example, the 3rd dataset in Table I is obtained by replacing four input patterns of the 2nd dataset with four unfavorable ones ($\gamma < 20°$), resulting in less accurate estimates for most parameters. Nonetheless, by removing such patterns via the evaluation of (16), the overall calibration results are greatly improved, as shown in the last row of Table I.

4) Full Camera Calibration for Varied Focal Length: As one of the key feature (**F4**) mentioned in Section I, the proposed method can calibrate cameras with non-fixed focal length (FL), while Zhang's method is not designed to

TABLE II

COMPARISON OF FULL CAMERA CALIBRATION (VARIED FOCAL LENGTH)

Set	Noise	Estimation errors*								
		Ours[6]				Zhang's				
		ΔPP	ΔFL	ΔR	ΔT	ΔPP	ΔFL_x	ΔFL_y	ΔR	ΔT
5	0.0	**0.0**	**0.00**	**0.00**	**0.00**	4.9	99.92	99.31	1.61	3.85
6	± 1.0	**3.8**	**4.28**	**1.33**	**0.82**	11.5	59.96	59.43	2.07	3.75

* Ground truth focal length: $(400, 400, 400, 400, 440, 440, 440, 440)$.

* All values are the average of 20 repetitions.

TABLE III

EFFECTS OF FL VARIANCE AND IMAGE COUNT PER FL

Set	FL	Ours		Zhang's		
		ΔPP	ΔFL	ΔPP	ΔFL_x	ΔFL_y
6	400/440	3.82	4.28	11.53	59.96	59.43
6A	400/480	3.86	4.26	6.70	90.26	89.64
6B	400/520	3.93	4.23	2.43	66.88	67.05
6C	all*	6.51	9.02	14.04	48.42	45.85

* Ground truth FL: $(400, 440, 480, 520)$.

cope with such situation. Table II shows the calibration results obtained with the proposed approach, as well as those from Zhang's method. It is readily observable that even under the noise-free condition, significant estimation errors can already be observed in the latter, although perfect estimations are achieved with the proposed approach. As for the noisy case, both methods generate estimation errors similar to those shown in Table I, except for the much worse estimation of focal length obtained with Zhang's method. In either case, the major difference of calibration performance is in the estimation of focal length, which is greatly constrained by the ability to cope with varied focal length during the image acquisition process.

Table III shows additional experimental results to further evaluate the capability of the proposed approach with respect to focal length variance and the image count for each focal length, wherein datasets 6A and 6B are similar to dataset 6 but with additional 10% and 20% increases, respectively, in the differences in focal length, while dataset 6C only consists of four images with each image having a different focal length. It is readily observable that satisfactory calibration results similar those obtained in Table II can be obtained with the proposed approach with (i) larger variance in the focal length and (ii) less calibration patterns for each focal length, while this is not true for Zhang's method. Such results are well expected from the property of a principal line: an axis of symmetry (passing through the principal point), which is totally independent to the focal length.

The above concept can be observed clearly from the geometry of principal lines of calibration results listed in Table III, as shown in Figure 6, wherein each color of a principal line corresponds to a different focal length. As all principal lines are supposed to pass through the principal point, the estimation of the latter is essentially unaffected by the selection of the former, as long as the degenerated, near collinear principal line configuration mentioned in Sec. II-D does not happen. In general, the above geometric observation will always be

[6]There are actually very small computation errors for the noise free case, which correspond to $\Delta PP = 7.81E-5$, $max\Delta FL = 2.44E-4$, $\Delta R = 3.50E-3$, and $\Delta T = 1.28E-5$.

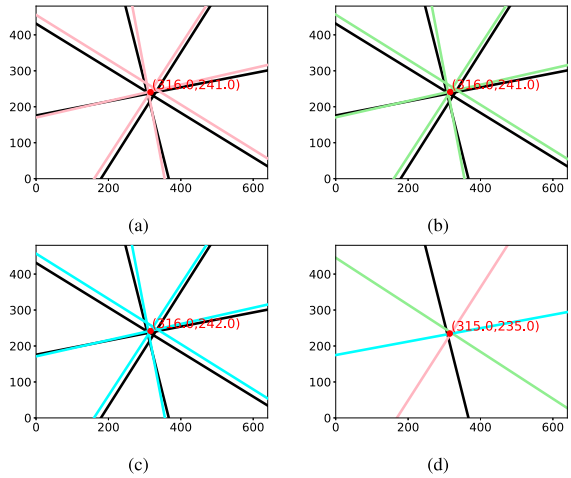

Fig. 6. Principal lines obtained for calibration results listed in Table III, wherein each color corresponds to a different focal length (with PLs from FL = 400 shown in black). (a) Set 6, (b) Set 6A, (c) Set 6B, and (d) Set 6C.

true, as one can also see next in the experiments using real data, if the principal point does not change with focal length. (Intuitively, if such change does happen, estimation of the principal point, which is not fixed anyway, will be impaired with an imperfect intersection of principal lines.)

Incidentally, one may notice that good estimations of principal point are also obtained in Table III for Zhang's method. This is because the test data are obtained by pure rotation of the calibration pattern, resulting in symmetric images of the checkerboard pattern (similar to that shown in Fig. 4 (a)), each has the axis of symmetry passing through the principal points. However, such special case is generally not possible for real data, as one will see at the end of the next subsection wherein poor estimations of principal point are obtained with Zhang's method.

B. Performance Evaluation Using Real Data

For performance evaluation of the proposed calibration method under realistic conditions, 7×8 checkerboard images are employed as calibration patterns.[7] Similar to the experiments considered in Section III-A, we demonstrate the flexibility of the proposed work over Zhang's method by comparing their performance under two different experiment setups, i.e., (a) using fixed focal length and (b) using different focal lengths in the image acquisition process. For setup (a), another approach of outlier removal described in Section II-D, which is based on the RMSE of the estimation of principal point in (9), will be employed to filter out inappropriate calibration patterns to further improve the estimation result.

1) Real Data With Fixed Focal Length: For camera calibration considered in this section, the focal length of the camera is fixed while images of calibration pattern are captured. As suggested in Section III-A2, a good set of images should have $\gamma \cong 45°$ and α should be nearly uniformly distributed between 0° and 180°. Figure 7(a) shows eight images thus obtained with a Logitech C920HD Pro camera with a image

[7]These images are obtained by taking pictures, from different directions, of the checkerboard pattern displayed on a 22-inch flat panel display.

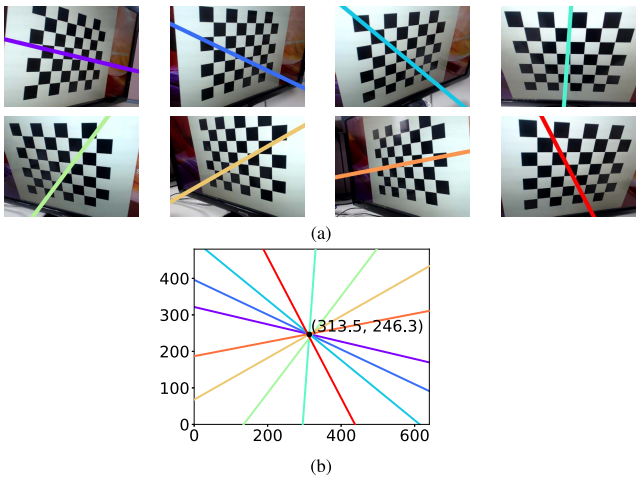

Fig. 7. (a) Eight calibration images with α nearly uniformly distributed between 0° and 180°. (b) Eight principal lines, each determined from an calibration image in (a).

TABLE IV

COMPARISON OF CAMERA CALIBRATION (FIXED FOCAL LENGTH)

Set	Ours		Zhang's	
	PP/RMSE	FL_{avg}/STD	PP	FL
7	(313.5, 246.3)/2.77	617.1/7.8	(317.6, 248.9)	616.8
8	(**310.8**, **240.2**)/59.8	**610.2**/29.7	(342.2, 235.8)	631.8
9	(**313.2**, **237.4**)/56.4	**611.5**/26.3	(350.1, 230.9)	634.8
10	(**307.1**, **235.7**)/30.7	**619.6**/30.7	(335.1, 237.0)	628.2

TABLE V

CALIBRATION RESULTS IMPROVED FROM SETS 8-10 IN TABLE IV VIA OUTLIERS REMOVAL

Set	Ours	
	PP/RMSE	FL_{avg}/STD
11	(312.2, 246.8)/**10.5**	608.5/**23.3**
12	(311.7, 247.7)/**10.2**	607.2/**20.9**
13	(310.1, 244.6)/**8.5**	614.2/**22.5**

resolution of 640 × 480,[8] while Figure 7(b) shows a total of 8 principal lines, with (313.5, 246.3) being estimated as the location of the principal point. The estimated principal point, which is near the center of the image, along with the estimated focal length (FL) are shown as the 1st set (set 7) of data of Table IV. Under this near ideal circumstances, both our method and Zhang's method produce similar results. (Note that results of the estimation of extrinsic parameters are omitted for brevity.)

To evaluate the sensitivity of calibration to unfavorable (ill-posed) calibration patterns, results of three more datasets (Sets 8-10) are also included in Table IV by replacing some good patterns in Set 7 with unfavorable ones, which can be visualized easily by comparing Figure 7(b) and Figures 8(a)-(c). The adverse influence of such replacements are readily observable from the significant deviations from the near ideal calibration results for Zhang's method, while the proposed approach seems to be more robust as the estimated values of PP and FL are much less affected.

On the other hand, it is possible for the proposed approach to remove the above problematic calibration patterns, similar to that performed in Sec. III-A3 for synthetic data. Specifically,

[8]The camera-display distance is manually adjusted to about 45cm so the checkerboard pattern will cover a large portion of the image.

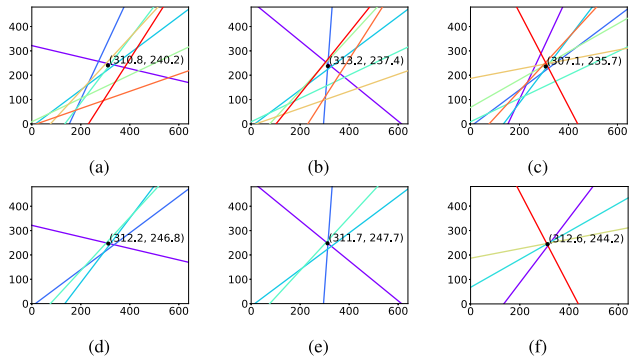

Fig. 8. Selected sets of principal lines (some due to ill-posed calibration patterns): (a) Set 8, (b) Set 9, and (c) Set 10, and their counterparts after performing simple, RMSE-based outlier removal: (d) Set 11, (e) Set 12, and (f) Set 13.

Sets 11, 12 and 13 in Table V are obtained by simply screening out possible outliers whose RMSEs are greater than 15 in Sets 8-10. One can see that most estimations are improved, with both RMSE and STD reduced significantly. Figures 8(d)-(f) show such outlier removal results, i.e., the remaining principal lines, for the original ones shown in Figures 8(a)-(c), respectively.

Beside ill-posed calibration patterns, a set of individually good patterns which violates guideline (ii) mentioned in Sec. III-A2 may also result in unreliable estimation results, as shown in Figures 9(a) and (b) for an example of such calibration patterns and their corresponding principal lines, respectively, wherein extraneous estimation error of PP can be found in the vertical direction. Nonetheless, such condition can be easily detected by examining the angular extent of the principal lines, and additional calibration images can be retaken to generate principal lines of different orientation, and more reliable calibration results.

2) Camera Calibration With Varied Focal Length: In this subsection, calibration of cameras with varied focal length (FL) is considered, which is a more challenging but corresponds to fairly common situation in real world scenarios.[9] Table VI shows the calibration results for three sets of calibration patterns which are captured by a Canon EOS 5D Mark III camera with an image resolution of 3840×2560.[10] As guidelines mentioned in Sec. III-A2 for selecting good poses of the calibration pattern are closely followed, satisfactory calibration results are obtained with both methods for the first two datasets (Sets 14 and 15), each established for a fixed (but different) FL.

On the other hand, consider the last dataset (Set 16) in Table VI, which corresponds to a mixture of calibration patterns from Set 14 and Set 15, with half of them from the 1st half of the former and the rest from the 2nd half of the latter, as shown in Figure 10(a). Figures 10(b) and (c) show the principal lines calculated using all images in (a) and only using images with FL = 39 mm, respectively. It is readily observable from Table VI that our approach

[9]For example, one may use different zooming levels, e.g., in our experiments, to take pictures of a stationary calibration pattern from different distances/angles while adjusting for proper image size.

[10]The camera-display distance is manually adjusted to between 15cm and 30cm so the checkerboard pattern will cover a large portion of the image.

TABLE VI
COMPARISON OF CAMERA CALIBRATION (REAL DATA, VARIED FOCAL LENGTH)

Set	FL (mm)	Ours		Zhang's	
		PP	FL	PP	FL
14	39	(1917.6, 1270.5)	3851.7, 3787.6, 3817.9, 3836.3, 3837.9, 3835.6	(1923.0, 1274.5)	3822.3
15	50	(1920.6, 1263.8)	4770.0, 4741.9, 4727.2, 4669.1, 4721.7, 4707.1	(1919.1, 1275.6)	4712.5
16	Mixed	(**1916.8**, **1266.0**)	**3791.5**, **3823.2**, **3841.7**, **4734.7**, **4710.2**, **4766.1**	(1897.0, 1416.8)	3953.0

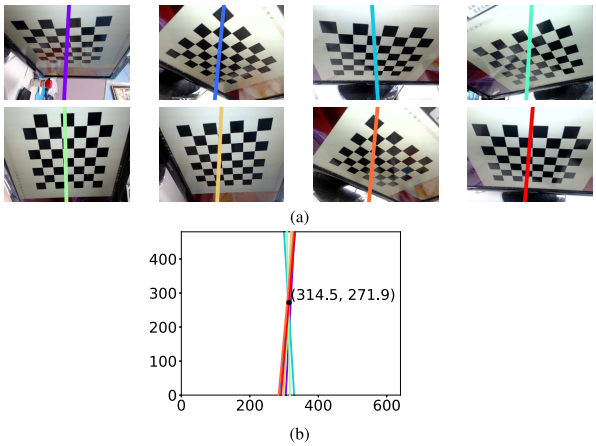

Fig. 9. (a) Eight calibration patterns, each has a reasonably good pose. (b) The eight nearly parallel principal lines.

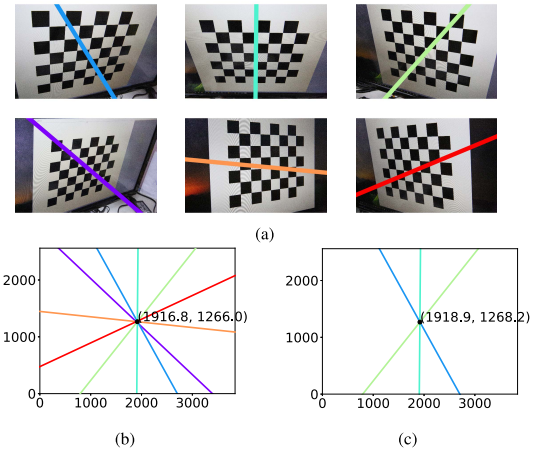

Fig. 10. (a) Calibration patterns captured with FL = 39 mm (first row) and FL = 50 mm (second row). (b) Principal lines obtained for all images in (a). (c) Principal lines obtained for first row of images in (a).

still performs satisfactorily and generates results similar to those for Sets 14 and 15. However, significantly worse results are obtained with Zhang's method which assumed fixed FL. In particular, the estimated FL (3953.0) is quite different from the two FLs obtained for the fixed cases (3822.3 and 4712.5). Moreover, the estimation of principal point is also impaired quite seriously under such situation, i.e., with a deviation of more than 20 pixels (140 pixels) in the horizontal (vertical) direction. (Similar problems can be found for the estimation of extrinsic parameters and are omitted for brevity.)

IV. CONCLUSION

In this paper, we have made a major attempt to establish a new camera calibration procedure based on a geometric perspective. The proposed approach resolves two main issues associated with the widely used Zhang's method, which include the lack of clear hints of appropriate pattern poses and the limitation of its applicability imposed by the assumption of fixed focal length. The main contribution of this work is to provide a closed-form solution to the calibration of extrinsic and intrinsic parameters based on the analytically tractable principal lines, with intersection of such lines identified as the principal point. Each of the lines can then conveniently represent the relative 3D orientation and position (each only has one degree of freedom) between the image plane and the corresponding calibration plane for a special **WCS-IPCS** pair. Consequently, computations associated with the calibration can be greatly simplified, while useful guidelines to avoid outliers in the computation can be established intuitively. Experimental results for both synthetic and real data clearly validate the correctness and robustness of the proposed approach, with both compared favorably with Zhang's method, especially in terms of the possibilities to screen out problematic calibration patterns as well as the ability to cope with the situation of varied focal length.

Accordingly, one major direction of our future research is to establish a more systematic way of identifying aforementioned problematic calibration patterns. In addition, possible improvements of the proposed calibration approach include developments/employments of (i) more accurate feature extraction schemes, as in [25], [26], (ii) better fitting/optimization methods for obtaining camera parameters, as in [27], [28], and (iii) additional ways of outlier removal.

APPENDIX A
DERIVING THE EXPRESSION OF CONSTANT c' IN (8)

In (8), the principal line is expressed as

$$a'u' + b'v' + c' = 0,$$

where $a' = (-h_1\ h_8 + h_2\ h_7)$, $b' = (-h_4\ h_8 + h_5\ h_7)$ and c' is a constant. The value of c' can be solved by plugging the vanishing point G' (6) in (8), or

$$\begin{aligned} c' &= -((-h_1\ h_8 + h_2\ h_7)\frac{h_1\ s + h_2\ c}{h_7\ s + h_8\ c} \\ &\quad + (-h_4\ h_8 + h_5\ h_7)\frac{h_4\ s + h_5\ c}{h_7\ s + h_8\ c}) \\ &= -(\frac{(h_2^2 + h_5^2 - h_1^2 - h_4^2)h_7 s + (h_1\ h_2 + h_4\ h_5)(h_7\ s - h_8\ c)}{h_7\ s + h_8\ c}) \\ &= -(\frac{(h_2^2 + h_5^2 - h_1^2 - h_4^2)h_7 s + (h_1\ h_2 + h_4\ h_5)(h_7\ s - h_8 \frac{h_8}{h_7} s)}{h_7 s + h_8 \frac{h_8}{h_7} s} \\ &= -(\frac{(h_2^2 + h_5^2 - h_1^2 - h_4^2)h_7\ h_8 + (h_1\ h_2 + h_4\ h_5)(h_7^2 - h_8^2)}{h_7^2 + h_8^2}), \end{aligned}$$

where $c \triangleq \cos(\theta)$, $s \triangleq \sin(\theta)$, as also defined in Sec. II-A1, and θ can be found with (4).

APPENDIX B
TIME COMPLEXITY ANALYSIS

In this section, time complexity of our calibration method is compared with Zhang's. Major computations of the proposed approach, as shown in Figure 2, involve consecutive derivations of (i) principal lines, (ii) principal point, (iii) other intrinsic parameters, and extrinsic parameters with respect to (iv) the special (aligned) **WCS-IPCS** (see Figure 1), or (v) the original **WCS-IPCS**. (Note that unlike Zhang's method, (iv) and (v) are optional, as the pose of an arbitrarily selected checkerboard pattern is often unimportant.)

For (i), each principal line can be found with (8), which has time complexity $O(1)$, after the corresponding homography matrix is estimated through the direct linear transformation (DLT) algorithm [29]. Assume there are K feature points, e.g., $K = 6 \times 7$ for a 7×8 checkerboard, time complexity of homography matrix estimation is $O(2K9^2)$ due to the singular-value decomposition (SVD) operation on the $2K \times 9$ matrix [30]. If there are N images of the checkerboard, time complexity of N principal lines calculation is $O(N2K9^2 + N) = O(N2K9^2)$. Once N principal lines are calculated, the principal point in (ii) can be calculated using (9). Again, the least square solution can be obtained with SVD decomposition of an $N \times 2$ matrix, with time complexity $O(N2^2)$.

As for (iii), since the calculation is based on the special geometry depicted in Figure 1, H^{new} needs to be calculated by (12), which includes four matrix multiplications and one matrix inversion of a 3×3 matrix, and the time complexity is $5O(3^3)$. Apparently, time complexity for calculating the scaling factor s and focal length f, i.e., in (18), is $O(1)$. Therefore, time complexity for calculating the focal length for each pose is $5O(3^3) + O(1) = O(3^3)$ and the total time complexity for finding (iii) is $O(N3^3)$.

Next, if we want to calculate extrinsic parameters with respect to the aligned **WCS-IPCS**, wherein $\alpha = \beta = 0$ and $T^{new} = \begin{bmatrix} 0 & 0 & t_z^{new} \end{bmatrix}$, we only need to calculate γ and t_z^{new} using (16) and (17), respectively, for each pose with time complexity $O(1)$. Thus, the total time complexity for (iv) is $O(N)$. However, if we want to calculate extrinsic parameters with respect to the original **WCS-IPCS**, time complexities for obtaining $\tilde{M}_{int}^{-1}$ and $\tilde{M}_{ext}$ in (19) are both $O(3^3)$, while time complexity for computing $r_3 = r_1 \times r_2$ is $O(3^2)$; therefore, the total time complexity for (v) is $O(N3^3)$. Table VII summaries the above analysis of time complexity of our method.

On the other hand, major steps of Zhang's method include: (i) derivation of initial closed-form solutions for (a) nonlinearly transformed intrinsic parameters and (b) extrinsic parameters, and (ii) nonlinear optimization of all parameters based on the maximum-likelihood criterion. For (i-a), homography matrices are calculated to form a system of linear equations $Vb = 0$, with b representing nonlinearly transformed intrinsic parameters [12]; therefore, $O(N2K9^2)$ is added for time complexity estimation. Then, least square solotion of $Vb = 0$ can be obtained via SVD decomposition of a $2N \times 6$ matrix, which has time complexity $O(2N6^2)$. Once b is estimated, the initial intrinsic parameters can be calculated in closed-form with time complexity $O(1)$, or by Cholesky decomposition [29] with time complexity $O(m^3)$, where m is the dimension of the 3×3 matrix formed by b. Similar to step (v) of our method, the initial extrinsic parameters in (i-b) can also be calculated by multiplying the scaling factor, the inverse of intrinsic matrix, and the homography matrix with the constraint $\mathbf{r_3} = \mathbf{r_1} \times \mathbf{r_2}$. Since each pose estimation includes two matrix multiplications and one matrix inversion of a 3×3 matrix, time complexity is equal to $O(N3^3)$. Thus, the total time complexity of (i-a) and (i-b) of Zhang's method is $O(N2K9^2 + 2N6^2 + 1 + N3^3) = O(N2K9^2)$.

In Zhang's method, after the estimation of initial intrinsic/extrinsic parameters, all parameters are refined in (ii) by nonlinear optimization method, i.e., Levenberg-Marquardt Method (LMM). According to [31], the global complexity bound of LMM is $O(\epsilon^{-2})$ where $0 < \epsilon < 1$ gives the stop condition for the iterative minimization of the solution ϕ, such that $||\nabla\phi(x)|| \leq \epsilon$. Table VIII summaries the above analysis of time complexity of Zhang's method.

According to Table VII and Table VIII, estimation of homography matrices corresponds to the most time consuming step in both algorithms. However, as Zhang's method needs to perform iterative computations to find a globally optimized solution, time complexity of our method is lower than or equal to Zhang's method.

TABLE VII
TIME COMPLEXITY OF OUR METHOD

Step	Description	Input	Output	Time Complexity
(i)	principal lines	N images, each with K feature points	N principal lines	$O(2NK9^2)$
(ii)	principal point	N principal lines	A principal point	$O(N2^2)$
(iii)	other intrinsic parameters	N principal lines, a principal point, and N homography matrices H	N focal lengths, scaling factors, and new homography matrices H^{new}	$O(N3^3)$
(iv)*	extrinsic parameters w.r.t. aligned **WCS-IPCS** (see Figure 1)	N new homography matrices H^{new}	N translation vectors $[0, 0, t_z^{new}]$ and N rotation angles $[\gamma, 0, 0]$	$O(N)$
(v)*	extrinsic parameters w.r.t. original **WCS-IPCS**	Outputs of (ii) and (iii), and N homography matrices H	N translation vectors $[t_x, t_y, t_z]$ and N rotation angles $[\gamma, \beta, \alpha]$	$O(N3^3)$
Total time complexity:				$O(2NK9^2)$

* Derivations of (iv) or (v) are optional.

TABLE VIII
TIME COMPLEXITY OF ZHANG'S METHOD

Step	Description	Input	Output	Time Complexity
(i-a)	derivation of initial intrinsic parameters	N images, each with K feature points	5 intrinsic parameters	$O(2NK9^2)$
(i-b)	derivation of initial extrinsic parameters	5 intrinsic parameters and N homography matrices H	$N \times 6$ extrinsic parameters	$O(N3^3)$
(ii)	nonlinear optimization of all parameters based on the maximum-likelihood criterion	all initial parameters, $N \times K$ image points, and $0 < \epsilon < 1$	5 intrinsic parameters and $N \times 6$ extrinsic parameters	$O(\epsilon^{-2})$
Total time complexity:				$O(2NK9^2) + O(\epsilon^{-2})$

REFERENCES

[1] A. Das and S. L. Waslander, "Calibration of a dynamic camera cluster for multi-camera visual SLAM," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 4637–4642.

[2] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2005, pp. 807–814.

[3] M. Sangi and M. Jahed, "A fast 3D hand model reconstruction by stereo vision system," in *Proc. 2nd Int. Conf. Comput. Autom. Eng. (ICCAE)*, Feb. 2010, pp. 545–549.

[4] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*. Cambridge, MA, USA: MIT Press, Jan. 1993.
[5] R. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE J. Robot. Autom.*, vol. 3, no. 4, pp. 323–344, Aug. 1987.
[6] S. J. Maybank and O. D. Faugeras, "A theory of self-calibration of a moving camera," *Int. J. Comput. Vis.*, vol. 8, no. 2, pp. 123–151, Aug. 1992.
[7] S. Bougnoux, "From projective to Euclidean space under any practical situation, a criticism of self-calibration," in *Proc. 6th Int. Conf. Comput. Vis.*, Jan. 1998, pp. 790–796.
[8] R. Mohedano and N. Garcia, "Capabilities and limitations of mono-camera pedestrian-based autocalibration," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 4705–4708.
[9] Z. Tang, Y.-S. Lin, K.-H. Lee, J.-N. Hwang, J.-H. Chuang, and Z. Fang, "Camera self-calibration from tracking of moving persons," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 265–270.
[10] B. Caprile and V. Torre, "Using vanishing points for camera calibration," *Int. J. Comput. Vis.*, vol. 4, no. 2, pp. 127–139, Mar. 1990.
[11] L. Ling Wang and W.-H. Tsai, "Camera calibration by vanishing lines for 3-D computer vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 4, pp. 370–376, Apr. 1991.
[12] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.
[13] L. Tan, Y. Wang, H. Yu, and J. Zhu, "Automatic camera calibration using active displays of a virtual pattern," *Sensors*, vol. 17, no. 4, p. 658, 2017.
[14] P. Rojtberg and A. Kuijper, "[POSTER] efficient pose selection for interactive camera calibration," in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR-Adjunct)*, Oct. 2017, pp. 182–183.
[15] C. Ricolfe-Viala and A.-J. Sanchez-Salmeron, "Optimal conditions for camera calibration using a planar template," in *Proc. 18th IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 853–856.
[16] A. S. Alturki and J. S. Loomis, "Camera principal point estimation from vanishing points," in *Proc. IEEE Nat. Aerosp. Electron. Conf. (NAECON) Ohio Innov. Summit (OIS)*, Jul. 2016, pp. 307–313.
[17] M.-T. Lu and J.-H. Chuang, "Fully automatic camera calibration for principal point using flat monitors," in *Proc. 25th IEEE Int. Conf. Image Process.*, Oct. 2018, pp. 3154–3158.
[18] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.
[19] O. Chum and J. Matas, "Optimal randomized RANSAC," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 8, pp. 1472–1482, Aug. 2008.
[20] A. F. Siegel, "Robust regression using repeated medians," *Biometrika*, vol. 69, no. 1, pp. 242–244, 1982.
[21] P. J. Rousseeuw, "Robust regression using repeated medians," *J. Amer. Stat. Assoc.*, vol. 79, pp. 871–880, Jan. 1984.
[22] R. Wilcox, "A note on the theil-sen regression estimator when the regressor is random and the error term is heteroscedastic," *Biometrical J.*, vol. 40, no. 3, pp. 261–268, Jul. 1998.
[23] R. Fernandes and S. G. Leblanc, "Parametric (modified least squares) and non-parametric (Theil–Sen) linear regressions for predicting biophysical parameters in the presence of measurement errors," *Remote Sens. Environ.*, vol. 95, no. 3, pp. 303–316, Apr. 2005.
[24] P. Woit, *Quantum Theory, Groups and Representations: An Introduction*, vol. 13. New York, NY, USA: Columbia Univ., Mar. 2016.
[25] A. Datta, J.-S. Kim, and T. Kanade, "Accurate camera calibration using iterative refinement of control points," in *Proc. IEEE 12th Int. Conf. Comput. Vis. Workshops, ICCV Workshops*, Sep. 2009, pp. 1201–1208.
[26] L. Yang, C. Tian, V. A. Vaishampayan, and A. R. Reibman, "An automatic grid corner extraction technique for camera calibration," in *Proc. 19th IEEE Int. Conf. Image Process.*, Sep. 2012, pp. 349–352.
[27] J. H. Park and S. H. Park, "Improvement on Zhang's camera calibration," *Appl. Mech. Mater.*, vols. 479–480, pp. 170–173, Dec. 2013.
[28] L. Deng, G. Lu, Y. Shao, M. Fei, and H. Hu, "A novel camera calibration technique based on differential evolution particle swarm optimization algorithm," *Neurocomputing*, vol. 174, pp. 456–465, Jan. 2016.
[29] W. Burger, "Zhang's camera calibration algorithm: In-depth tutorial and implementation," Dept. Digit. Media, Univ. Appl. Sci. Upper Austria, School Inform., Commun. Media, Softwarepark, Hagenberg, Austria, Tech. Rep. HGB16-05, May 2016.
[30] L. N. Trefethen and D. Bau III, *Numerical Linear Algebra*, 1st ed. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, Jun. 1997, ch. 31, pp. 234–240.
[31] K. Ueda and N. Yamashita, "On a global complexity bound of the Levenberg-Marquardt method," *J. Optim. Theory Appl.*, vol. 147, no. 3, pp. 443–453, 2010.

Jen-Hui Chuang (Senior Member, IEEE) received the B.S. degree in electrical engineering from National Taiwan University in 1980, the M.S. degree in electrical and computer engineering from the University of California at Santa Barbara in 1883, and the Ph.D. degree in electrical and computer engineering from the Urbana-Champaign University of Illinois at Urbana-Champaign in 1991. Since then, he has been on the faculty of the Department of Computer and Information Science, National Chiao Tung University (NCTU). From 2004 to 2005, he was the Chairman of the Department of Computer and Information Science, NCTU, where he was the Associate Dean of the College of Computer Science from 2006 to 2007. From 2017 to 2020, he served as the Dean for the College of Computer Science, NCTU. His research interests include signal and image processing, computer vision, and pattern recognition, robotics, and potential-based 3-D modeling. He was the President of the Chinese Image Processing and Pattern Recognition Society from 2015 to 2016, and served as a Governing Board Member of the International Association of Pattern Recognition from 2013 to 2016. He currently serves as a Governing Board Member of IEEE Taipei Chapter. He served as an Associate Editor for *Signal Processing* from 2005 to 2010, and has been an Associate Editor of *Journal of Information Science and Engineering* since 2009.

Chih-Hui Ho received the B.S. degree in EECS from National Chiao Tung University, Taiwan, in 2016, and the M.S. degree in computer science and engineering from the University of California at San Diego, USA, in 2019, where he is currently pursuing the Ph.D. degree in electrical and computer engineering. His research interests include machine learning and computer vision. This work was done, when he was a Research Assistant with National Chiao Tung University.

Ardian Umam received the B.S. degree in electrical engineering from Gadjah Mada University, Indonesia, in 2014, and the M.S. degree in electrical engineering and computer science from National Chiao Tung University, Taiwan, in 2018. He is currently a Faculty Member of the School of Electrical Engineering and Informatics, Bandung Institute of Technology, Indonesia. His research interests include computer vision, machine learning, and data science.

Hsin-Yi Chen received the B.S. degree in applied mathematics from National Chung Hsing University, Taiwan, in 2005, and the M.S. degree in computer science and information engineering from National Central University, Taiwan, in 2007. Then, she served as a Deputy Engineer with the Industrial Technology Research Institute, mainly focusing on optimization researches, from 2007 to 2012. From 2012 to 2015, she served as an Engineer and mainly focused on algorithm and system development in automatic optical inspection (AOI) field. From 2015 to 2019, she served as a Senior Engineer and focused on artificial intelligence (AI) applications, which include smart factory and interactive system of transparent display. In 2018, she joined the full-time doctoral program at the Department of Computer Science, National Chiao Tung University.

Jenq-Neng Hwang (Fellow, IEEE) received the B.S. and M.S. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1981 and 1983, respectively, and the Ph.D. degree from the University of Southern California. In Summer 1989, he joined the Department of Electrical and Computer Engineering (ECE), University of Washington, Seattle, where he was promoted to a Full Professor in 1999. He served as the Associate Chair for Research from 2003 to 2005 and from 2011 to 2015. He is currently the Associate Chair for Global Affairs and International Development in the ECE Department. He is the Founder and Co-Director of the Information Processing Laboratory, which has won CVPR AI City Challenges awards consecutively in the past years. He has written more than 350 journal, conference papers, and book chapters in the areas of machine learning, multimedia signal processing, and multimedia system integration and networking, including an authored textbook on "*Multimedia Networking: From Theory to Practice*" (Cambridge University Press). He has close working relationship with the industry on multimedia signal processing and multimedia networking. He has been a fellow of IEEE since 2001. He is currently on the Editorial Board of *ZTE Communications*, *ETRI Journal*, IJDMB, and JSPS journals. He is a Founding Member of the Multimedia Signal Processing Technical Committee of IEEE Signal Processing Society and was the Society's representative to IEEE Neural Network Council from 1996 to 2000. He is currently a member of the Multimedia Technical Committee (MMTC) of IEEE Communication Society and of the Multimedia Signal Processing Technical Committee (MMSP TC) of IEEE Signal Processing Society. He received the 1995 IEEE Signal Processing Society's Best Journal Paper Award. He served as the Program Co-Chair for IEEE ICME 2016 and was the Program Co-Chair for ICASSP 1998 and ISCAS 2009. He served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON NEURAL NETWORKS, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON IMAGE PROCESSING, and the *IEEE Signal Processing Magazine* (SPM).

Tai-An Chen received the B.S. degree in computer science information engineering from the National Taipei University of Technology, Taiwan, in 2008, and the M.S. degree in computer science from National Chiao Tung University, Taiwan, in 2018. Since 2013, he has been serving as a Research Engineer with the Industrial Research Institute of Technology, mainly focusing on augmented reality and virtual reality researches. He is currently a Junior Manager with E.Sun Bank, Taiwan.