

The Mechanistic Deconvolutive Image Sensor Model for an Arbitrary Pan–Tilt Plane of View

S. H. Lim, and T. Furukawa

Abstract—This paper presents a generalized form of the mechanistic deconvolution technique (GMD) to modeling image sensors applicable in various pan–tilt planes of view. The mechanistic deconvolution technique (UMD) is modified with the given angles of a pan–tilt plane of view to formulate constraint parameters and characterize distortion effects, and thereby, determine the corrected image data. This, as a result, does not require experimental setup or calibration. Due to the mechanistic nature of the sensor model, the necessity for the sensor image plane to be orthogonal to its z-axis is eliminated, and it reduces the dependency on image data. An experiment was constructed to evaluate the accuracy of a model created by GMD and its insensitivity to changes in sensor properties and in pan and tilt angles. This was compared with a pre-calibrated model and a model created by UMD using two sensors with different specifications. It achieved similar accuracy with one-seventh the number of iterations and attained lower mean error by a factor of 2.4 when compared to the pre-calibrated and UMD model respectively. The model has also shown itself to be robust and, in comparison to pre-calibrated and UMD model, improved the accuracy significantly.

Keywords—Image sensor modeling, mechanistic deconvolution, calibration, lens distortion.

I. INTRODUCTION

LENS distortion inherent in off-the-shelf image sensors creates inaccuracy in their obtained information. Consequently, corrective sensor modeling—also known as *calibration*—is an important process for applications that require accurate geometric measurements [1]. The process has long been an important factor in photogrammetry and computer vision, and more recently, in the area of robotics and automation [2].

Conventionally, the correcting sensor model was derived by techniques that invoke an analytical closed-form solution with a set of linear equations formulated based on a distortion-free camera model. Since these computationally inexpensive techniques do not take distortion into account and yield inaccurate results [3], [5], they require combining the linear equations with a nonlinear set. Note that effective existing

optimization technique, such as Gauss–Newton and Newton–Raphson are applied in these techniques [3], [4]. Convergence, however, requires that the initial estimate be sufficiently near for accurate calibration equations to be attained and is computationally expensive. The disadvantages in these classical techniques have motivated researchers to develop new techniques, which may be categorized into two groups: the two-step method [3], [5]–[9], and self-calibration [10]–[12].

The two-step method was proposed by Tsai [3], where a closed-form solution is first derived using radial alignment constraints to estimate the extrinsic parameters and effective focal length. An optimization technique, with the estimated parameters as initial estimate, is then applied to compute nonlinear solution, and thereby retrieves the radial distortion parameters and the corrected effective focal length. This technique reduces the number of required iterations considerably. The accuracy of this method was improved by Weng et al. [5] by including tangential distortion. In an attempt to reduce the number of iterations, a number of other groups have also proposed techniques that create analytical closed-form solution to obtain radial distortion parameters [6]–[8]. Park and Hong [9] simplified Tsai’s technique by applying look-up-table (LUT) techniques such that it is applicable for real-time applications.

A different type of calibration technique, namely *self-calibration*, was developed by Maybank et al. [10] where the intrinsic parameters are assumed to be constant in order to reduce the computational expense. This technique, in contrast to the two-step method, does not also require a calibration setup which includes a calibration object with known 3D geometrical features. Despite having higher flexibility, the self-calibration requires at least three different orientations and achieves lower accuracy. This has been modified and extended to include different constraints, mainly camera motion and scene constraints, to increase its accuracy [11]. Self-calibration has been fused with the two-step method by Zhang [12], and was able to reduce the number of orientations whilst having better robustness, as compared to the self-calibration.

These techniques have a common approach in that the image sensor is evaluated with a known image. The models derived by these techniques have shown to reproduce parameters of image sensors successfully, negating the need to manually obtain the sensor’s mechanical and electrical properties. Nevertheless, these properties may also derive the

S. H. Lim is with the University of New South Wales, NSW, 2052 Australia (corresponding author to provide phone: 61-2-93854125; e-mail: shen.lim@student.unsw.edu.au).

T. Furukawa, was with University of New South Wales, NSW, 2052 Australia. He is now with Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0238 USA (e-mail: tomonari@vt.edu).

sensor model successfully and reduce the dependency on the image quality, and, acts as an alternative solution to calibration. A mechanistic deconvolution technique (UMD) has been proposed by the authors [13], and it utilizes the sensor's mechanical and electrical properties to model the image sensor. This technique does not require calibration, and hence, is not evaluated with a known image. The technique, however, requires the image plane to be orthogonal to its z-axis which limits its applicability.

This paper presents a generalized mechanistic deconvolutive image sensor model (GMD) for application in an arbitrary pan-tilt plane of view. In this approach, the UMD technique is modified with the given pan and tilt angles of a sensor plane of view to formulate the constraint parameters and characterize the distortion effects. The constraint parameters and distortion effects are then used to obtain the corrected image data. This approach eliminates the requirement for the image plane to be orthogonal to its z-axis while maintaining the advantage of UMD; that is removing the requirement of experimental evaluation of image sensors while remaining robust to changes in their properties.

The paper is organized as follows. Section II reviews the general formulation of existing problem and techniques, including UMD technique. The generalized form of the mechanistic deconvolutive model is presented in Section III. This is followed by numerical examples in Section IV, including experimental setup and procedure. Finally, Section V presents the conclusion of this paper.

II. CONVENTIONAL SENSOR MODELS AND MECHANISTIC DECONVOLUTION

A. Conventional Sensor Models

The objective of sensor modeling is to identify as many possible parameters of the sensor, to create the model in an attempt to correct image data. This sub-section describes the general approach of the conventional sensor models, which includes formulation of the pinhole camera model and the distortion model.

1) Pinhole Camera Model

The pinhole camera model, which is the ideal form of image sensor model, is based on a set of linear transformation equations. The position of the image plane coordinate system, $\mathbf{x}_u = [u, v, 1]^T$ is obtained using its z-component, and the position with respect to the camera 3D coordinate system $\mathbf{x}_c = [x_c, y_c, z_c]^T$ by

$$\mathbf{x}_u = \mathbf{A}\mathbf{x}_c, \quad (1)$$

where \mathbf{A} is a matrix that consists of intrinsic parameters. The matrix \mathbf{A} is formulated as

$$\mathbf{A} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{with} \quad \begin{cases} f_x = f/p_u \\ f_y = f/p_v \end{cases} \quad (2)$$

based on the assumption that the image plane axes are parallel to the world coordinate system, and as shown in \mathbf{A} , the

intrinsic parameters are the focal length f , principal point coordinates $[u_0, v_0]$ and pixel size $\mathbf{p}_s = [p_u, p_v]^T$. The position \mathbf{x}_c is generated based on the position of an object with respect to the world coordinate system $\mathbf{x}_w = [x_w, y_w, z_w]^T$ using

$$\mathbf{x}_c = \mathbf{R}\mathbf{x}_w + \mathbf{t} \quad (3)$$

where \mathbf{R} is the rotational matrix and \mathbf{t} is the translational vector. The parameters \mathbf{R} and \mathbf{t} are known as the extrinsic parameters.

2) Distortion Model

The radial and tangential distortion effects inherent in image sensors are characterized by the distortion model. The position \mathbf{x}_u , also known as undistorted coordinates, is then determined by the given distorted coordinates, $\mathbf{x}_d = [u_d, v_d]^T$, such as follows:

$$\mathbf{x}_u = \left(1 + \sum_i K_i (r_d^{2i} - 1) \right) \mathbf{x}_d + \mathbf{g}, \quad (4)$$

where K_{1i} are the radial distortion factors, and, \mathbf{g} and r_d are given by

$$\mathbf{g} = \begin{bmatrix} P_1 (r_d^2 + 2u_d^2) + 2P_2 u_d v_d \\ P_2 (r_d^2 + 2u_d^2) + 2P_1 u_d v_d \end{bmatrix} \quad \text{and} \quad (5)$$

$$r_d = \sqrt{u_d^2 + v_d^2} \quad (6)$$

where P_1 and P_2 are the tangential distortion factors. It is noted that the image data is highly affected by radial distortion.

3) General Conventional Models Approach

The techniques that derive the conventional sensor models initially invoke an analytical closed-form solution with the given target geometrical features, alignment constraints and formulations based on the pinhole camera model. The closed-form solution generates the extrinsic parameters and the focal length. These then combine the formulations based on the pinhole camera model with the nonlinear equations of the distortion model to generate nonlinear solution that estimates the final intrinsic and extrinsic parameters. In computing the nonlinear solution, the techniques apply optimization technique with the parameters obtained from the closed-form solution as the initial parameters.

B. Mechanistic Deconvolution (UMD)

In mechanistic deconvolution technique, the image sensor is modeled using its mechanical and electrical properties, instead of characteristics of a known image [13]. The undistorted coordinates \mathbf{x}_u are given by

$$\mathbf{x}_u = f(f_d, \mathbf{c}_p, \mathbf{x}_d, \boldsymbol{\delta}) \quad (7)$$

where f_d is the effective focal length, \mathbf{c}_p contains the constraint parameters of the sensor and $\boldsymbol{\delta}$ is the lens distortion factor offered by the sensor mechanical and electrical properties. The f_d is determined by the lens system approximated by a *thick lens*, as

$$f_d = \frac{nr_1r_2}{(n-1)(nr_2 - n_1r_1 + t(n-1))} + w^f \quad (8)$$

with the refractive index n , the radii curvature closest and farthest from the light source r_1 and r_2 , the lens thickness t and w^f as the combination of uncertainties from r_1 , r_2 and t . The constraint parameters \mathbf{c}_p , as required by (7), include the feature resolution constraint parameter $(\mathbf{b}_1 - \mathbf{b}_2)_{\min}$, the field of view constraint parameter $(\mathbf{b}_1 - \mathbf{b}_2)_{\max}$ and the depth of field constraint parameter h_d .

The $(\mathbf{b}_1 - \mathbf{b}_2)_{\min}$ shows the permissible smallest size of a scene feature by

$$(\mathbf{b}_1 - \mathbf{b}_2)_{\min} - \frac{l}{f_d} \mathbf{p}_s \geq 0, \quad (9)$$

using the given distance between the scene feature and sensor l . The points $\mathbf{b}_1 = [u_1, v_1]^T$ and $\mathbf{b}_2 = [u_2, v_2]^T$ are the reference points that define the size vector of the scene feature. On the other hand, the $(\mathbf{b}_1 - \mathbf{b}_2)_{\max}$ constrains the allowable largest size of a scene feature and is controlled by the sensor chip size $\mathbf{s}_a = [s_u, s_v]^T$, such as follows:

$$\frac{l}{f_d} \mathbf{s}_a - (\mathbf{b}_1 - \mathbf{b}_2)_{\max} \geq 0 \quad (10)$$

The final constraint parameter h_d controls the depth of field range D_f , which is the region where a known image is considered in focus. The D_f is defined by

$$D_f = D_f^+ - D_f^- \quad (11)$$

where D_f^+ and D_f^- are the far and near limits of depth of field respectively, and are formulated as

$$D_f^- = \frac{h_d l}{h_d + l}, \text{ and} \quad (12)$$

$$D_f^+ = \begin{cases} \frac{h_d l}{h_d - l}; & l < h_d \\ \infty & ; l > h_d \end{cases}$$

Here, h_d is also known as the hyperfocal distance and is given by

$$h_d = 1000d_a \sqrt{\frac{q}{2} + v^h}, \quad (13)$$

where d_a is the diameter aperture size, q is the linear constant introduced from the relationship between image brightness and d_a and v^h is the uncertainties due to d_a .

Instead of using the radial and tangential distortion factor in the distortion model, the lens distortion factor $\delta = [\delta_u, \delta_v]^T$ is determined by

$$\delta_i = \left(m_a \left(1 + \frac{f_d}{l} \right) - f_d \right) \tan \beta_i; i = u : v \quad (14)$$

using the sensor lens properties. The m_a is the aperture stop position offset from the principal plane of the approximated lens system and β_u and β_v are the refracted angles of the ray based on m_a in u -axis and v -axis respectively. Equation (14) is obtained from the differences of the image position

coordinates based on the sensor approximated lens system and the principal plane, as derived by

$$\delta_i = m_a \tan \gamma_i - f_d \tan \beta_i; i = u : v, \text{ and} \quad (15)$$

$$\tan \gamma_i = \frac{l - m_a}{l} \tan \beta_i; i = u : v. \quad (16)$$

The angles γ_u and γ_v are the refracted angles of the ray based on the position of the principal plane in u -axis and v -axis.

III. GENERALIZED MECHANISTIC DECONVOLUTION MODEL

Using mechanistic deconvolution technique as a basis, the proposed approach derives a model that is applicable in various pan and tilt angles of the plane of view, θ_u and θ_v . The model is presented in this section, which includes the formulation of the undistorted coordinates based on the reformulation of the constraint parameters and re-characterization of the distortion effects.

A. Formulation of Undistorted Coordinates for Given Pan and Tilt Angles

By introducing the given angles of the plane of view θ_u and θ_v , the undistorted coordinates \mathbf{x}_u are generated by

$$\mathbf{x}_u = \mathbf{x}_d + \mathbf{I}({}^{\theta} \delta)^2, \quad (17)$$

where ${}^{\theta} \delta$ is the modified lens distortion factor, and are constrained as follows:

$${}^{\theta} (\mathbf{b}_1 - \mathbf{b}_2)_{\min} \leq \mathbf{x}_u \leq {}^{\theta} (\mathbf{b}_1 - \mathbf{b}_2)_{\max} \quad (18)$$

with ${}^{\theta} (\mathbf{b}_1 - \mathbf{b}_2)_{\min}$ representing the modified feature resolution constraint parameter and ${}^{\theta} (\mathbf{b}_1 - \mathbf{b}_2)_{\max}$ representing the modified field of view constraint parameter. In this approach, the angles θ_u and θ_v are bound by

$$\phi_i - \pi/2 \leq \theta_i \leq \pi/2 - \phi_i; i = u : v, \quad (19)$$

where ϕ_u and ϕ_v are the halves of the sensor view angles in the u and v directions respectively, and can be written as

$$\tan \phi_i = \frac{p_i \tan(\theta_d/2)}{|\mathbf{p}_s|}; i = u : v \quad (20)$$

using the pixel size $\mathbf{p}_s = [p_u, p_v]^T$ and the sensor view angle θ_d . The effective focal length f_d and the depth of field constraint parameter h_d previously defined by (8) and (13) are not affected by θ_u and θ_v , and hence are not modified in this approach. The formulations for the lens distortion factor ${}^{\theta} \delta$ and the modified constraint parameters ${}^{\theta} (\mathbf{b}_1 - \mathbf{b}_2)_{\min}$ and ${}^{\theta} (\mathbf{b}_1 - \mathbf{b}_2)_{\max}$ are presented in the following sub-sections.

B. Derivation of Modified Constraint Parameters

The angles θ_u and θ_v affect the interpretation of the size of a scene feature in the plane of view ${}^{\theta} (\mathbf{b}_1 - \mathbf{b}_2)$, derived from the sensor image plane, in correspondence to the area covered by the regions, namely region I and II. Fig. 1 illustrates the area of regions due to the effects of θ_u and θ_v independently. As shown in Fig. 1, region I and II are determined using the v -axis for the θ_u case and the u -axis for θ_v . It is noted that both

axes intercept at principal coordinates of the plane of view $[x_0, y_0]$. The size of a scene feature in region I ${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)$ and in region II ${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)$ are given by

$$\left| {}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2) \cdot \mathbf{e}_i \right| = \frac{|(\mathbf{b}_1 - \mathbf{b}_2) \cdot \mathbf{e}_i| \cos \phi_i}{\cos(\theta_i + \phi_i)}; i = u : v \text{ and} \quad (21)$$

$$\left| {}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2) \cdot \mathbf{e}_i \right| = \frac{|(\mathbf{b}_1 - \mathbf{b}_2) \cdot \mathbf{e}_i| \cos \phi_i}{\cos(\phi_i - \theta_i)}; i = u : v, \quad (22)$$

where $(\mathbf{b}_1 - \mathbf{b}_2)$ represents the size in the plane of view orthogonal to the sensor z-axis.

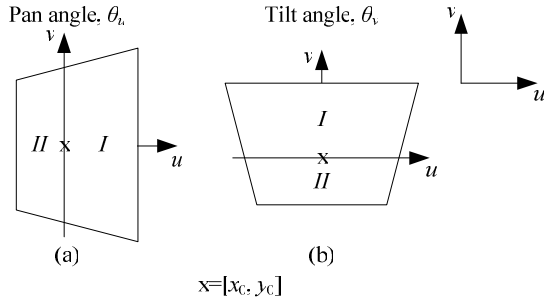


Fig. 1 Definition of region I and II in the plane of view based on the pan and tilt angles independently

Due to the effects of θ_u and θ_v , the feature resolution constraint parameter $(\mathbf{b}_1 - \mathbf{b}_2)_{\min}$ previously defined in (9), which allows the smallest permissible size of a scene feature, is required to be reformulated. The modified feature resolution constraint parameter ${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\min}$ is offered by

$${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\min} = {}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\min} > {}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\min} \quad (23)$$

with ${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\min}$ and ${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\min}$ representing the minimum size of the scene feature corresponding to region I and II respectively. The parameters ${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\min}$ and ${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\min}$ are formulated as

$$\left| {}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\min} \cdot \mathbf{e}_i \right| = \frac{|(\mathbf{b}_1 - \mathbf{b}_2)_{\min} \cdot \mathbf{e}_i| \cos \phi_i}{\cos(\theta_i + \phi_i)}; i = u : v \text{ and} \quad (24)$$

$$\left| {}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\min} \cdot \mathbf{e}_i \right| = \frac{|(\mathbf{b}_1 - \mathbf{b}_2)_{\min} \cdot \mathbf{e}_i| \cos \phi_i}{\cos(\phi_i - \theta_i)}; i = u : v. \quad (25)$$

The general formulation of ${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\min}$ from the sensor's mechanical and electrical properties, using (23) and (9), then becomes

$$\left| {}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\min} \cdot \mathbf{e}_i \right| - \frac{l_c \cos \phi_i}{f_d \cos(\phi_i + \theta_i)} p_i \geq 0; i = u : v, \quad (26)$$

where l_c is the given distance between $[x_0, y_0]$ and sensor, instead of the given distance l introduced by (9).

Similarly, the field of view constraint parameter $(\mathbf{b}_1 - \mathbf{b}_2)_{\max}$ that allows the largest permissible size of a scene feature, as previously defined in (10), is also required to be reevaluated. The modified field of view constraint parameter ${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\max}$ is formulated as

$${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\max} = {}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\max} + {}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\max} \quad (27)$$

where ${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\max}$ and ${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\max}$ are the maximum possible size of the scene feature for region I and II respectively, and these are given by

$$\left| {}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\max} \cdot \mathbf{e}_i \right| = \frac{|(\mathbf{b}_1 - \mathbf{b}_2)_{\max} \cdot \mathbf{e}_i| \cos \phi_i}{\cos(\theta_i + \phi_i)}; i = u : v \text{ and} \quad (28)$$

$$\left| {}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\max} \cdot \mathbf{e}_i \right| = \frac{|(\mathbf{b}_1 - \mathbf{b}_2)_{\max} \cdot \mathbf{e}_i| \cos \phi_i}{\cos(\phi_i - \theta_i)}; i = u : v. \quad (29)$$

Using (27) and (10), the ${}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\max}$ can be expressed in the general form

$$\frac{l_c \cos^2 \phi_i \cos \theta_i}{f_d (\cos^2 \phi_i - \sin^2 \theta_i)} s_i - \left| {}^{\theta}(\mathbf{b}_1 - \mathbf{b}_2)_{\max} \cdot \mathbf{e}_i \right| \geq 0; i = u : v \quad (30)$$

where $\mathbf{s}_u = [s_u, s_v]^T$ is the sensor chip size.

The final constraint parameter, which is the depth of field h_d , is not affected by θ_u and θ_v . The depth of field range D_f , however, is affected by θ_u and θ_v and defined by

$$D_f = {}^{\theta}D_f^+ - {}^{\theta}D_f^- \quad (31)$$

where ${}^{\theta}D_f^+$ and ${}^{\theta}D_f^-$ are the modified far and near limits of the depth of field respectively. The limits ${}^{\theta}D_f^+$ and ${}^{\theta}D_f^-$ are

$${}^{\theta}D_f^- = \frac{h_d l_c}{h_d + l_c} + |(\mathbf{b}_1 - \mathbf{b}_2)_{\max} \cdot \mathbf{e}_u| \cos \theta_u \text{ and} \quad (32)$$

$${}^{\theta}D_f^+ = \begin{cases} \frac{h_d l_c}{h_d - l_c} - |(\mathbf{b}_1 - \mathbf{b}_2)_{\max} \cdot \mathbf{e}_u| \cos \theta_u; l_c < h_d \\ \infty & ; l_c > h_d \end{cases}$$

for the pan angle θ_u , and

$${}^{\theta}D_f^- = \frac{h_d l_c}{h_d + l_c} + |(\mathbf{b}_1 - \mathbf{b}_2)_{\max} \cdot \mathbf{e}_v| \cos \theta_v \text{ and} \quad (33)$$

$${}^{\theta}D_f^+ = \begin{cases} \frac{h_d l_c}{h_d - l_c} - |(\mathbf{b}_1 - \mathbf{b}_2)_{\max} \cdot \mathbf{e}_v| \cos \theta_v; l_c < h_d \\ \infty & ; l_c > h_d \end{cases}$$

for the tilt angle θ_v , where l_c is previously defined as the given distance between the plane of view principal point coordinates and sensor.

C. Characterization of Distortion Effects

The distortion effects of the sensor, similar to the constraint parameters, are affected by θ_u and θ_v . Due to the complexity of the pan-tilt plane of view, γ_u and γ_v previously defined in (16) are not able to be determined, and this results in inapplicability of the derivation of the lens distortion factor $\delta = [\delta_u, \delta_v]^T$ using (15). The proposed approach modifies the distortion model in the mechanistic deconvolution technique, and defines the modified lens distortion factor, ${}^{\theta}\delta = [{}^{\theta}\delta_u, {}^{\theta}\delta_v]^T$ as

$${}^{\theta}\delta_i = f_d \tan\left(\frac{{}^{\theta}\beta_i}{2}\right); i = u : v \quad (34)$$

where ${}^{\theta}\beta_u$ and ${}^{\theta}\beta_v$ are the modified refracted angles of the ray based on the position of aperture stop offset in the u -axis and

v-axis respectively. The angles ${}^{\theta}\beta_u$ and ${}^{\theta}\beta_v$ are generated by

$$\tan {}^{\theta}\beta_i = \frac{|y|l_c}{f_d(l_c - m_a) - y_i m_a \tan \phi_i}; i = u : v \quad (35)$$

with the position $\mathbf{y} = [y_u, y_v]^T$ representing the position in the image plane, in comparison to the image plane principal point coordinates $[u_0, v_0]$, ϕ_u and ϕ_v , as given by (20).

The angles ${}^{\theta}\beta_u$ and ${}^{\theta}\beta_v$ produced in (35) are then utilized to derive the general formulation of ${}^{\theta}\delta$

$${}^{\theta}\delta_i = \sum_{j=1}^3 \frac{n_j l_c^j |y|^j}{f_d^{j-1} \left[l_c - m_a \left(1 + \frac{y_i (\tan \phi_u + \tan \phi_v)}{f_d} \right) \right]^j} + n_4 f_d; i = u : v \quad (36)$$

with the sensor lens properties, and using the approximated formulation of ${}^{\theta}\beta_i/2$, denoted by

$$\tan \left(\frac{{}^{\theta}\beta_i}{2} \right) = \frac{1}{2} \tan {}^{\theta}\beta_i + \sum_{j=1}^3 n_j \tan^j ({}^{\theta}\beta_i) + n_4 \quad (37)$$

where $[n_1, n_2, n_3, n_4]$ are the constant values defined from the third-order approximation between ${}^{\theta}\beta_i/2$ and ${}^{\theta}\beta_i$. Equation (36) can be simplified to obtain an alternate solution of the lens distortion factor $\delta = [\delta_u, \delta_v]^T$, such as follows:

$$\delta_i = \sum_{j=1}^3 \frac{n_j l_c^j |y|^j}{f_d^{j-1} (l - m_a)^j} + n_4 f_d; i = u : v \quad (38)$$

which eliminates the necessity of β_u and β_v introduced in (14)

IV. NUMERICAL EXAMPLES

An experimental setup was constructed to evaluate the behavior of the generalized mechanistic deconvolutive model. The evaluations of the model include accuracy, robustness to changes in image sensor properties and in pan and tilt angles without recalibration. The grid pattern of 20 box X 18 box was used as the target in this experimental setup. Two image sensors with different specifications, namely Sensor 1 and Sensor 2, were used in the evaluation process, as shown in Table I. It is noted that Sensor 1 has better specifications than Sensor 2. The only similarity between image sensors is; the focus of the sensors is manually adjusted. In this experiment, these sensors were tilted by 20 degrees, in comparison to the plane of view.

TABLE I
SPECIFICATIONS OF THE IMAGE SENSORS

Specifications	Sensor 1	Sensor 2
Pixel resolution (MPixel)	3.2	0.4
Sensor type	CCD	CMOS
Sensor size, s_a (mm ²)	11.8X7.9	2.6X2.13
Aperture stop position offset, m_a (mm)	0.1	0.3
Lens thickness, t (mm)	6	4

The generalized mechanistic deconvolutive model (GMD) was compared with a model derived by mechanistic

deconvolution technique (UMD) described in Section II, and a conventional sensor model developed by a conventional technique illustrated in Section II. The optimization technique used by the conventional technique was Gauss–Newton method. In this example, the radii mean error was used to measure the accuracy of both models, when compared with the ideal form of the grid pattern. The robustness of the models was then tested based on two types of changes, which are two factors of sensor properties and the pan and tilt angles θ_u and θ_v . The two factors of sensor properties are the aperture stop position offset from the principal plane, m_a and lens thickness, t .

Fig. 2 shows the number of iterations required by the conventional, UMD and GMD models to achieve minimal percentage mean error for both sensors. The UMD model yields mean errors of 2.2% and 8.9% for Sensor 1 and Sensor 2. The mean errors produced by the conventional and GMD models, in comparison to the UMD model, are lower by a factor of 2.4 for Sensor 1 and 3.9 for Sensor 2, which are approximately 0.9% and 2.3% for Sensor 1 and Sensor 2. The GMD model, however, manages to obtain the mean errors using one-seventh and one-tenth of the iterations required by the conventional model for Sensor 1 and Sensor 2 respectively.

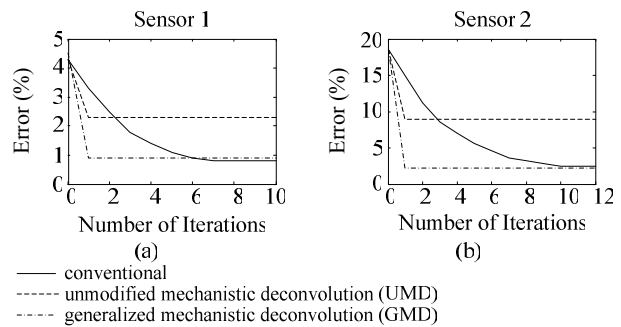


Fig. 2 Accuracy evaluation of the models in corresponding to number of iterations for Sensor 1 and Sensor 2

Fig. 3 illustrates the accuracy of the models due to incremental changes in sensor properties, m_a and t . The UMD model attains consistent mean errors due to changes in m_a and t for Sensor 1 with the average mean errors of 3.1% and 2.9% respectively, as illustrated in Fig. 3(a) and 3(c). Similarly, the GMD model generates consistent mean errors for Sensor 1 with lower average mean errors of 1.1% and 1% for changes in m_a and t , in comparison to the UMD model. These results show that UMD and GMD models are not affected by the changes in m_a and t . The conventional model, despite having lower mean errors than the UMD model at the given specifications of Sensor 1 in Table I, is unable to achieve consistent mean errors due to changes in m_a and t for Sensor 1 and obtains maximum mean errors of 10.1% and 12.5% respectively.

The conventional model also obtains inconsistent mean

errors due to changes in m_a and t for Sensor 2 and the mean errors escalate to maximum of 20.1% for changes in m_a and 19.8% for changes in t , as shown in Fig. 3(b) and 3(d). The UMD model, similar to the results produced for Sensor 1, generates consistent mean errors due to changes in m_a and t for Sensor 2 with the average mean errors of 9.3% and 9.8% respectively. The GMD model yields consistent and low mean errors due to changes in m_a and t for Sensor 1 with the average mean errors of 2.4% for changes in m_a and 2.7% for changes in t . These results show that the GMD model provides higher and more consistent accuracy, in comparison to the conventional and UMD models.

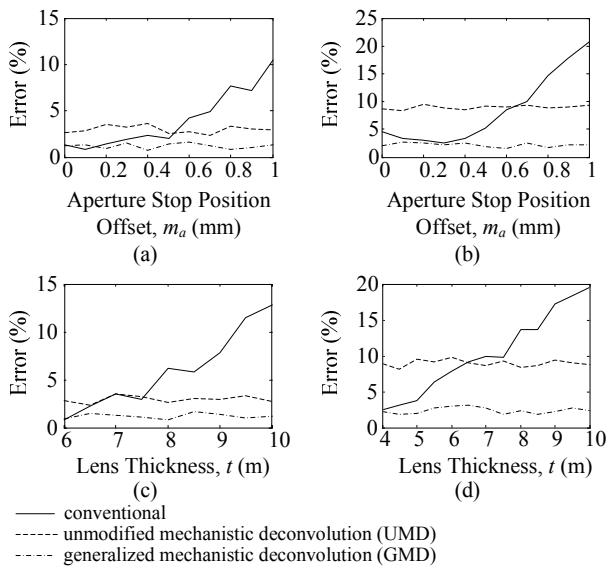


Fig. 3 Accuracy evaluations of the models due to incremental changes in aperture stop position offset, m_a and lens thickness, t for Sensor 1 (left column) and Sensor 2 (right column)

Fig. 4 shows the accuracy of the models due to incremental changes in θ_u and θ_v , independently. It is noted that θ_v is set to zero during the incremental changes in θ_u and likewise. Fig. 4(a) and 4(c) show that the conventional model achieves linearly increasing mean errors due to increasing changes in θ_u and θ_v for Sensor 1, with maximum mean errors of 2.1% and 2.3%. The mean errors produced by the UMD model also increases with the increasing changes in θ_u and θ_v for Sensor 1, with higher maximum mean errors of 4.3% and 4.5% respectively. The GMD model, in comparison to the conventional and UMD models, attains consistent and lower mean errors with average mean errors of 1.3% for both changes in θ_u and θ_v . These results show that GMD model is not affected by the changes in θ_u and θ_v .

Fig. 4(b) and 4(d) illustrate that the mean errors generated by the GMD model for Sensor 2, similar to results obtained for Sensor 1, are consistent with low average mean errors of 2.4% and 2.5% for changes in θ_u and θ_v . The conventional model

also produces consistent mean errors due to increasing changes in θ_u and θ_v for Sensor 2, and obtains higher average mean errors of 2.8% and 2.9% respectively when compared with the GMD model. The UMD model, on the contrary, is unable to attain consistent mean errors due to increasing changes in θ_u and θ_v , and the mean errors rise to maximum of 14.9% for θ_u and 14.8% for θ_v . Fig. 4 shows that all the models have similar patterns for the comparison between θ_u and θ_v for both sensors.

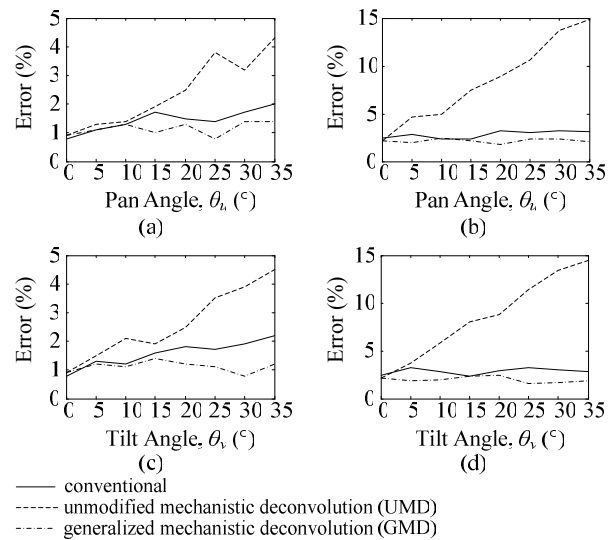


Fig. 4 Accuracy evaluation of the models due to incremental changes in pan and tilt angles, θ_u and θ_v for Sensor 1 (left column) and Sensor 2 (right column)

V. CONCLUSION

A generalized mechanistic deconvolution technique to modeling an image sensor for an arbitrary pan-tilt plane of view has been presented. The model constructed by the proposed approach (GMD) achieved accuracy similar to the conventional model while considerably reducing the number of iterations. In contrast to the conventional model, it also demonstrated high robustness to changes in sensor properties. In addition, the GMD model proved insensitive to changes in pan and tilt angles, and showed an improvement in accuracy when compared with the unmodified mechanistic deconvolution model (UMD). Thus it does not require recalibration if sensor parts are modified, and is applicable for various pan and tilt angles of the sensor plane of view. In conclusion, the model has capably demonstrated its capacity to correct for distortion without a dependence on image data. Continuing investigation of this approach will aim towards an extension to include tangential distortion, further improving the accuracy of corrected data.

REFERENCES

- [1] F. Remondino and C. Fraser, "Digital camera calibrations: Considerations and comparisons," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 36, no. 5, pp. 266–272, 2006.
- [2] G.N. DeSouza and A.C. Kak "Vision for mobile robot navigation: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 2, pp. 237–267, 2002.
- [3] R.Y. Tsai, "A versatile camera calibration technique for high-Accuracy 3D machine vision metrology using off-the-Shelf TV cameras and lenses," *IEEE Journal of Robotics and Automation*, vol. Ra-3, no. 4, pp. 323–344, 1987.
- [4] D.C. Brown, "Close-range camera calibration," *Photogrammetric Engineering*, vol. 37, no. 8, pp. 855–866, 1971.
- [5] J. Weng, P. Cohen and M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 10, pp. 965–980, 1992.
- [6] S. Graf and T. Hanning, "Analytical solving radial distortion parameters," *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 1104–1109, 2005.
- [7] J. Mallon and P.F. Whelan, "Precise radial un-distortion of images," *Proceedings of the 17th International Conference on Pattern Recognition*, vol. 1, 2004.
- [8] D.G. Bailey, "A new approach to lens distortion correction," *Proceedings Image and Vision Computing New Zealand*, pp. 59–64, 2002.
- [9] S. Park and K. Hong, "Practical ways to calculate camera lens distortion for real-time camera calibration," *Pattern Recognition*, vol. 34, no. 6, pp. 1199–1206, 2004.
- [10] S.J. Maybank and O.D. Faugeras, "A theory of self-calibration of a moving camera," *International Journal of Computer Vision*, vol. 8, no. 2, pp. 123–151, 1992.
- [11] E.E. Hemayed, "A survey of camera self-calibration," *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance*, pp. 352–357, 2003.
- [12] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [13] S. H. Lim, T. Furukawa, "Calibration-free image sensor modelling using mechanistic deconvolution," *Sensors and Transducers Journal*, vol. 90, Special Issue, pp. 195–208, April 2008.