Generalized Radial Alignment Constraint for Camera Calibration

Avinash Kumar, Narendra Ahuja
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
Email: {avinash,n-ahuja}@illinois.edu

***Abstract*—In camera calibration, the radial alignment constraint (RAC) has been proposed as a technique to obtain closed form solution to calibration parameters when the image distortion is purely radial about an axis normal to the sensor plane. But, in real images this normality assumption might be violated due to manufacturing limitations or intentional sensor tilt. A misaligned optic axis results in traditional formulation of RAC not holding for real images leading to calibration errors. In this paper, we propose a generalized radial alignment constraint (gRAC), which relaxes the optic axis-sensor normality constraint by explicitly modeling their configuration via rotation parameters which form a part of camera calibration parameter set. We propose a new analytical solution to solve the gRAC for a subset of calibration parameters. We discuss the resulting ambiguities in the analytical approach and propose methods to overcome them. The analytical solution is then used to compute the intersection of optic axis and the sensor about which overall distortion is indeed radial. Finally, the analytical estimates from gRAC are used to initialize the non-linear refinement of calibration parameters. Using simulated and real data, we show the correctness of the proposed gRAC and the analytical solution in achieving accurate camera calibration.**

I. Introduction

Camera calibration estimates the physical (intrinsic) properties of the camera and its pose (extrinsic) with respect to a known world coordinate system using known locations of 3D scene points and their measured image coordinates. Typically camera calibration is a two step procedure. In the first step, either all or a subset of unknown calibration parameters are linearly estimated by using a linear constraint, e.g. DLT [1], collinearity of a scene point and its image [2] under the assumption of no image distortion or image noise. In the second step, image distortion and noise are taken into account and calibration parameters are non-linearly optimized [3]. This step is typically initialized by the calibration estimates obtained in the first step.

Assuming radial distortion as the major source of image distortion, Tsai [4] observed that the location vectors of a scene point and its distorted image point should be radially aligned about the optic axis of the lens and thus their cross product must be zero. This was termed as the ***R**adial **A**lignment **C**onstraint (RAC)* and could be analytically solved for a subset of calibration parameters. The major assumption of RAC was that the optic axis is normal to the sensor at the ***C**enter **o**f radial **D**istortion (CoD)* and was known a priori. Although, later it was shown that the RAC could itself be used to compute the CoD [5].

But in a generic imaging setting, the optic axis may not be normal to the image sensor due to manufacturing limitations in aligning lens elements or assembling lens-sensor planes exactly parallel to each other. Although, sometimes an intentional tilting of sensor can prove useful in obtaining slanted depth-of-field effects like tilt-shift imaging [6], omnifocus imaging [7] and depth from focus estimation [8]. Under such a setting where sensor is non-frontal to the lens, the RAC can be interpreted in the following two ways, both of which we show to be inaccurate: (1) RAC can be modeled about an *"effective" optic axis* which is normal to the sensor at the location denoted as the *principal point*. But the total distortion about this point is a combination of radial and decentering [2] distortion and thus the world and distorted image point are not radially aligned. (2) If RAC is formulated about the physical optic axis, then even though the world and image point lie on the same 3D plane, they are not parallel to each other and thus are not radially aligned.

Thus, in this paper we propose the ***g**eneralized **R**adial **A**lignment **C**onstraint (gRAC)* to handle the more generic case of sensor non-frontalness. We first model the lens-sensor configuration by an explicit rotation matrix about the optic axis [9] and include it as a part of intrinsic calibration parameter set. Second, the rotation parameters are used to project the observed image points on the non-frontal sensor on to a hypothesized frontal sensor assuming that the pixel size (in metric) are known a-priori. The gRAC constraint is then derived for these frontal image points (Sec. IV) about the optic axis and the CoD. As this constraint is different than RAC [4], it requires a new analytical method to solve it for a subset of calibration parameters (Sec. V). Third, the analytical technique is used to computationally estimate the CoD (Sec. V-C). Sec. III describes the RAC from [4]. Sec. II describes the coordinate system and the generic lens-sensor configuration for which gRAC will be derived. Sec. VI describes the results obtained on synthetic and real data.

II. Calibration Coordinate Systems

In this section, we describe the **C**oordinate **S**ystems (CS) used in this paper for the task of camera calibration (Fig. 1).
(1) **World Coordinate System,** where the location of world points in metric units is known, e.g. corners of a checkerboard (CB) of known dimensions.
(2) **Image Coordinate System,** where the observed image points are measured in pixels.
(3) **Lens Coordinate System,** whose origin lies at the lens center (center of projection) and whose z axis coincides with the optic axis. It has metric units.
(4) **Sensor Coordinate System,** whose origin is at the CoD,

the z axis coincides with the optic axis and the xy plane lies on the sensor surface. It has metric units.

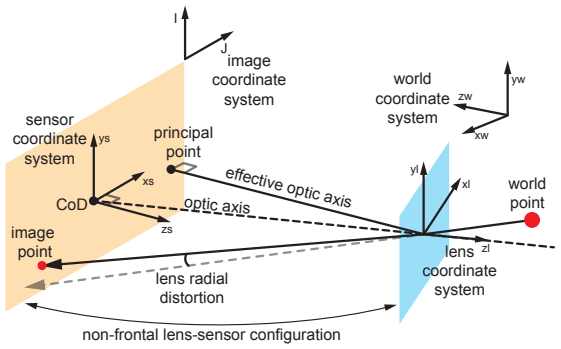

Fig. 1. Coordinate systems for camera calibration.

III. Tsai's Radial Alignment Constraint

In this section, we describe the radial alignment constraint as proposed in [4]. Consider Fig. 2(a) which describes the coordinate system used in [4]. The lens and the sensor coordinate system are assumed to be parallel to each other with a common z-axis (z_l or z_s) as the "effective" optic axis and O_s as the principal point. The image of world point $P_w = (x_w, y_w, z_w)$ on the sensor is formed at $P_d = (x_d, y_d)$. Assuming only radial lens distortion about O_s, this point would ideally be imaged at $P_u = (x_u, y_u)$ such that the triplet O_s, P_d, P_u are collinear. Let P_w be denoted as $P_l = (x_l, y_l, z_l)$ in the lens coordinate system, Then the normal from P_l onto the "effective" optic axis will be incident at $P_{oz} = (0, 0, z_l)$.

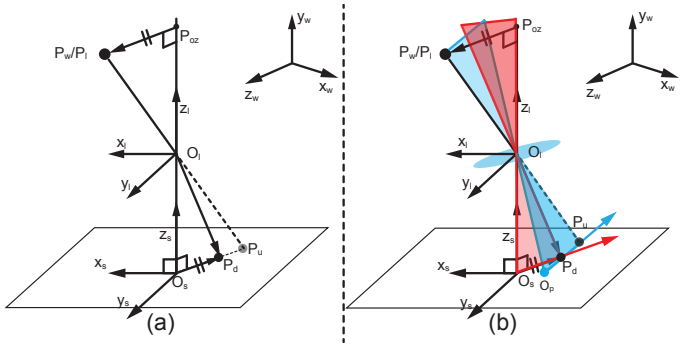

Fig. 2. (a) Imaging model for RAC [4]. (b) An Illustration of RAC not holding true in real images, when the sensor maybe non-frontal with respect to the lens plane.

Then, the RAC says that the vector $\overrightarrow{P_{oz}P_l}$ is radially aligned to the vector $\overrightarrow{O_sP_d}$ or $\overrightarrow{P_{oz}P_l} \| \overrightarrow{O_sP_d}$, as the two vectors are normal to the same line, namely "effective" optic axis and also lie on the same 3D plane formed by the points O_s, P_u, O_l. Thus, we get the RAC constraint $\overrightarrow{P_{oz}P_l} \times \overrightarrow{O_sP_d} = 0$, which is solved to obtain a subset of calibration parameters. Furthermore, assuming that radial distortion was symmetric about O_s, the RAC constraint was also used to estimate the principal point O_s [5].

But, while the imaging model in [4] assumed that radial distortion was symmetric about "effective" optic axis, in reality this is inaccurate for real images. Here, radial distortion is symmetric about the physical optic axis which may not coincide with the former due to unintentional lens misalignment or intentional sensor tilt (See Sec. I). Thus, a more generic image formation model is required (See Fig. 2(b)) [9], [10] where the non-alignment of lens and sensor is explicitly modeled via a rotation matrix and the distorted (P_d) and undistorted (P_u) image points are radially aligned about the CoD (O_p). It can be seen that the world point P_l lies on the 3D plane formed by triplets $\{O_p, P_u, O_l\}$ (shown in blue in Fig. 2(b)). In comparison, the 3D plane formed by $\{O_s, P_d, O_l\}$ (shown in red in Fig. 2(b)) is different from the blue plane as O_p is not a part of this plane. For RAC to hold, the two vectors: $\overrightarrow{O_sP_d}$ and $\overrightarrow{P_lP_{oz}}$, should be radially aligned which constrains them to lie on the same plane. Since $\overrightarrow{O_sP_d}$ belongs to red plane and P_l belongs to the blue plane which does not coincide with red plane, P_l is out of plane with respect to red plane. Thus the normal $\overrightarrow{P_lP_{oz}}$ from P_l normally incident onto the "effective" optic axis (edge of red plane) can never be coplanar with $\overrightarrow{O_sP_d}$, or traditional RAC [4] cannot hold. Thus, next we propose the gRAC for a generic non-frontal sensor model.

IV. Generalized Radial Alignment Constraint (gRAC)

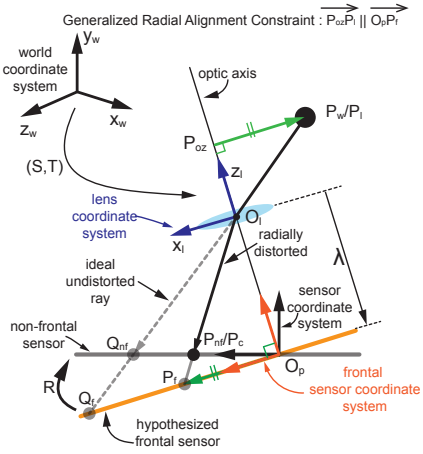

Fig. 3. Illustration of generalized radial alignment constraint (gRAC).

In this section, we derive the gRAC (See Fig. 3). Here the lens and sensor planes are assumed to be not parallel to each other but related via a rotation transformation R [9], [10], about the optic axis. Under these settings, if a world point P_w (P_l in lens coordinate system) is imaged at P_{nf} on the sensor (in sensor coordinate system) , then as per RAC, $\overrightarrow{P_{oz}P_l}$ is not parallel to $\overrightarrow{O_pP_{nf}}$. But, if the relative rotation R between lens and sensor coordinate system is known, then the projected frontal image point P_f of P_{nf} on a hypothesized frontal sensor gets radially aligned with the world point P_l. In other words, we will have that $\overrightarrow{O_pP_f} \| \overrightarrow{P_{oz}P_l}$. In the following we will derive this constraint as a function of R and then solve it to get a closed from solution to a subset of calibration parameters including R.

We define R as a rotation matrix which aligns the lens coordinate system with the sensor coordinate system and is parameterized by two Euler angles (ρ, σ) corresponding to clockwise rotations about its x and y axis respectively. The rotation of lens coordinate system about the z axis is considered redundant as the lens is symmetric about its z axis. Thus, the final Euler angle representation of rotation is $R(\rho, \sigma, 0)$ where:

$$R(\rho,\sigma,0) = \begin{bmatrix} \cos(\sigma) & \sin(\rho)\sin(\sigma) & \cos(\rho)\sin(\sigma) \\ 0 & \cos(\rho) & -\sin(\rho) \\ -\sin(\sigma) & \sin(\rho)\cos(\sigma) & \cos(\rho)\cos(\sigma) \end{bmatrix} \tag{1}$$

Let r_{ij} denoted the i^{th} row and j^{th} column entry of R. Next, we derive the gRAC by analyzing the geometric relationship between a given known 3D scene point P_w and its corresponding observed distorted image point P_{nf} as a function of various calibration parameters.

Consider the imaging configuration in Fig. 3, where a known world point $P_w = (x_w, y_w, z_w)$ in world coordinate system gets imaged at the pixel location $P_c = (I, J)$ in image coordinate system. Let the world and the lens coordinate system be related by a rotation $S = (s_{ij} : 1 \leq (i, j) \leq 3)$ parameterized by Euler angles (θ, ϕ, ψ) and a 3×1 translation $T = (t_x, t_y, t_z)$. Then, P_w can be expressed as $P_l = (x_l, y_l, z_l)$ in lens coordinate system, where $P_l = SP_w + T$. Thus,

$$\begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix} = \begin{bmatrix} s_{11}x_w + s_{12}y_w + s_{13}z_w + t_x \\ s_{21}x_w + s_{22}y_w + s_{23}z_w + t_y \\ s_{31}x_w + s_{32}y_w + s_{33}z_w + t_z \end{bmatrix}. \tag{2}$$

Let the imaged point P_c be expressed in sensor coordinate system as $P_{nf} = (xd_{nf}, yd_{nf})$, where

$$xd_{nf} = (I + I_0)s_x, \quad yd_{nf} = (J + J_0)s_y \tag{3}$$

and (I_0, J_0) is the location of the CoD in pixels and (s_x, s_y) are the pixel sizes (in metric units e.g. mm) along the x and y axis of sensor coordinate system.

Now, we compute the projection of P_{nf} on a hypothesized frontal sensor, so that a radial alignment constraint can be deduced between the frontal projected sensor point and the world point P_l expressed in lens coordinate system. Let the projected point on the frontal sensor be $P_f = (xd_f, yd_f)$. Given the sensor tilt parameterized by rotation R, the distance λ between the lens and frontal sensor coordinate system along the optic axis and the collinearity of center of projection O_l,P_{nf} and P_f, we get the coordinates of P_f as:

$$\begin{bmatrix} xd_f \\ yd_f \end{bmatrix} = \begin{bmatrix} \frac{-(r_{11}xd_{nf} + r_{21}yd_{nf})\lambda}{r_{13}xd_{nf} + r_{23}yd_{nf} - \lambda} \\ \frac{-(r_{12}xd_{nf} + r_{22}yd_{nf})\lambda}{r_{13}xd_{nf} + r_{23}yd_{nf} - \lambda} \end{bmatrix}. \tag{4}$$

Next, we project world point P_l on the optic axis to obtain $P_{oz} = (0, 0, z_l)$. Then, we have that location vectors $\overrightarrow{O_pP_f}$ and $\overrightarrow{P_{oz}P_l}$ are coplanar lying on a plane formed by points (O_p, P_f, P_l) and are also parallel and radially aligned to each other. From the radially aligned constraint, we have $\overrightarrow{O_pP_f} \times \overrightarrow{P_{oz}P_l} = 0$, which given $\overrightarrow{O_pP_f} = xd_f\hat{i} + yd_f\hat{j}$ and $\overrightarrow{P_{oz}P_l} = x_l\hat{i} + y_l\hat{j}$ (both in lens sensor coordinate system) simplifies to the generalized radial alignment constraint (gRAC):

$$xd_f \cdot y_l = yd_f \cdot x_l. \tag{5}$$

If it is assumed that the subset

$$U_1 = (I_0, J_0, s_x, s_y) \tag{6}$$

of calibration parameters is known, then $P_{nf} = (xd_{nf}, yd_{nf})$ can be computed using Eq. 3. Given known P_{nf}, Eq. 4 can be used to obtain hypothesized frontal points $P_f = (xd_f, yd_f)$ as a function of unknown calibration parameters (R, λ). Also, using Eq. 2, $P_l = (x_l, y_l)$ can be obtained in terms of unknown extrinsic calibration parameters $(\theta, \phi, \psi, t_x, t_y, t_z)$ and known world points $P_w = (x_w, y_w, z_w)$. Thus, Eq. 5 can be simplified to obtain the linear equation $\mathbf{Aq} = \mathbf{b}$, relating i^{th} world-image point observation as

$$\underbrace{\begin{bmatrix} xd_{nf}x_w & xd_{nf}y_w & xd_{nf}z_w & xd_{nf} & yd_{nf}x_w & yd_{nf}y_w & yd_{nf}z_w \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} q_1 \\ \vdots \\ q_7 \end{bmatrix}}_{\mathbf{q}} = \underbrace{yd_{nf}}_{\mathbf{b}}. \tag{7}$$

Here, $(\mathbf{A}, \mathbf{b})$ are known, while $\mathbf{q} = \{q_1, \cdots, q_7\}$ encodes seven calibration parameters denoted here as U_2:

$$U_2 = (\underbrace{\rho, \sigma}_{R}, \underbrace{\theta, \phi, \psi}_{S}, t_x, t_y). \tag{8}$$

via the following non-linear relationships:

$$q_1 = \frac{r_{11}s_{21} - r_{12}s_{11}}{r_{22}t_x} \quad (9) \qquad q_2 = \frac{r_{11}s_{22} - r_{12}s_{12}}{r_{22}t_x} \quad (10)$$

$$q_3 = \frac{r_{11}s_{23} - r_{12}s_{13}}{r_{22}t_x} \quad (11) \qquad q_4 = \frac{r_{11}t_y - r_{12}t_x}{r_{22}t_x} \quad (12)$$

$$q_5 = \frac{-s_{11}}{t_x} \quad (13) \quad q_6 = \frac{-s_{12}}{t_x} \quad (14) \quad q_7 = \frac{-s_{13}}{t_x} \quad (15)$$

As $(\mathbf{A}, \mathbf{b})$ are known, Eq. 7 can be solved in least squares sense given four or more observations of scene points to obtain an estimate $\overline{\mathbf{q}}$. This estimate can be used to analytically solve the set of non-linear relationships in Eq. (9-15) to obtain U_2 as we shown in Sec. V. It can be noted that in Tsai's RAC [4], R was an identity matrix and their solution was derived based on this assumption. In the gRAC case, the derivations are comparatively more involved due to the inclusion of R parameter. For calibrating the remaining calibration parameters, namely

$$U_3 = (\lambda, t_z). \tag{16}$$

we adopt the technique of [4] as shown in Sec. V. Thus, from Eq. 6,8,16, the final set of camera calibration parameters to be calibrated is $U = \{U_1, U_2, U_3\}$.

V. Analytical Solution to gRAC

In this section, we analytically solve Eq. (9-15) for the seven calibration parameters U_2 (Eq. 8) assuming that U_1 (Eq. 6) is known. Later, we will show a technique similar to [5] and estimate U_1 given optimal estimates of U_2 applied to the gRAC based linear Eq. 7. We use $|x|$ to denote that magnitude of x without knowing the sign.

A. Stage 1: Determining sign ambiguous estimates

1) ***Solving for*** t_x*:* Squaring and adding Eq. (13-15) and from orthonormality of first row of extrinsic rotation matrix S

(Eq. 2), t_x can be computed with a sign ambiguity as

$$|t_x| = \frac{1}{\sqrt{(q_5^2 + q_6^2 + q_7^2)}} \tag{17}$$

*2) **Solving for** s_{11}, s_{12}, s_{13}:* Given t_x^*, using Eq. (13,14,15), we get we get

$$s_{11} = -q_5 t_x \quad s_{12} = -q_6 t_x \quad s_{13} = -q_7 t_x \tag{18}$$

*3) **Solving for** s_{21}, s_{22}, s_{23}:* Adding the product of Eq. (9,13), Eq. (10,14), Eq. (11,15) we get,

$$\frac{r_{12}}{r_{22}} = t_x^2 \underbrace{(q_1 q_5 + q_2 q_6 + q_3 q_7)}_{M} \tag{19}$$

Also, adding the squares of Eq. (9,10,11) and using the orthonormality of first and second row of S, we obtain

$$\frac{r_{11}^2 + r_{12}^2}{r_{22}^2} = t_x^2 \underbrace{(q_1^2 + q_2^2 + q_3^2)}_{N} \tag{20}$$

$$\implies \frac{r_{11}^2}{r_{22}^2} = N t_x^2 - M^2 t_x^4 \quad \text{(Using Eq. 19)} \tag{21}$$

As $r_{11} = \cos(\sigma) > 0$ and $r_{22} = \cos\rho > 0$ from Eq. 1, the ratio $\frac{r_{11}}{r_{22}}$ from Eq. 21 can be determined uniquely as

$$\frac{r_{11}}{r_{22}} = \underbrace{\sqrt{N t_x^2 - M^2 t_x^4}}_{P}. \tag{22}$$

Applying Eq. (19,22) and Eq. (13-15) to Eq. (9-11) respectively we can solve for s_{21}, s_{22}, s_{23} with sign ambiguity as

$$s_{21} = \frac{(q_1 - t_x^2 M q_5) t_x}{P} \tag{23}$$

$$s_{22} = \frac{(q_2 - t_x^2 M q_6) t_x}{P} \tag{24}$$

$$s_{23} = \frac{(q_3 - t_x^2 M q_7) t_x}{P} \tag{25}$$

*4) **Solving for** s_{21}, s_{22}, s_{23} **uniquely**:* Assuming right hand coordinate system, the cross product of the first (Eq. 18) and second (Eq. (23-25)) row of S can be used to determine the third row of S: (s_{21}, s_{22}, s_{23}). These estimates are unique as the it involves terms of t_x^2 which is greater than 0 and all other terms involving q_i are uniquely known.

*5) **Solving for** t_y:* Applying Eq. (19,22) to Eq. 12, we get

$$t_y = \frac{(q_4 + t_x^2 M) t_x}{P} \tag{26}$$

*6) **Solving for** $\{r_{11}, \cdots, r_{33}\}$:* The left hand side of Eq. 19 and Eq. 22 can be expressed in terms of Euler angle (ρ, σ) via Eq. 1, which expresses sensor rotation matrix R in terms of its component Euler angles as follows

$$\frac{r_{12}}{r_{22}} = \frac{\sin\rho \sin\sigma}{\cos\rho} = \underbrace{t_x^2 M}_{L}, \quad \text{and,} \tag{27}$$

$$\frac{r_{11}}{r_{22}} = \frac{\cos\sigma}{\cos\rho} = P. \tag{28}$$

These two equations can be solved for (ρ, σ) with a sign ambiguity to obtain

$$\rho = \pm \cos^{-1}\left(\frac{L^2 + P^2 + 1 - \sqrt{(L^2 + P^2 + 1)^2 - 4P^2}}{2P^2}\right) \tag{29}$$

$$\sigma = \pm \sin^{-1}\left(\sqrt{1 - P^2 \cos^2(\rho)}\right) \tag{30}$$

Although the individual signs of (ρ, σ) are not known uniquely, the relative sign of (ρ, σ) with respect to each other can be determined from the sign of L in Eq. 27 as the denominator in Eq. 27 is always positive ($\cos\rho > 0$). The ambiguity here arises from the fact that gRAC is designed for a frontal coordinate system which is obtained by projecting the non-frontal sensor coordinates P_{nf} onto a frontal sensor to give P_f. Since, this projection involves taking the cosine of tilt angles encoded in R, it is many-to-one leading to sign ambiguity in analytical estimate of (ρ, σ).

B. Stage 2: Determining the sign of estimates

In Sec. V-A, we determined partial set of extrinsic and intrinsic parameters denoted here as $U_e = \{S, t_x, t_y\}$ and intrinsic parameters denoted here as $U_i = \{R\}$ respectively with sign ambiguity. While the sign ambiguity in determining U_e resulted from not knowing t_x uniquely in Eq. 17, the ambiguity in U_i was inherent to the gRAC constraint due to many-to-one projection map from a non-frontal sensor configuration to a frontal sensor configuration. Next, we present a technique to retrieve the sign of t_x uniquely (similar but not same as in Tsai [4]), thus determining U_e uniquely. This is followed by a method to uniquely determine U_i.

We also note that given all sign ambiguities in $\{U_e, U_i\}$, there are four possible solution sets for $\{U_e, U_i\}$, corresponding to the combinations: $\texttt{sign}(t_x) = \pm$ and either $\texttt{sign}(\rho, \sigma) = (+,+)/(-,-)$ or $\texttt{sign}(\rho, \sigma) = (+,-)/(-,+)$. This is so as the relative sign of (ρ, σ) is uniquely determined from $\texttt{sign}(L)$ (Eq. 27). Lets assume the two rotation matrices obtained from sign ambiguity of (ρ, σ) are R_1 and R_2.

*1) **Determining** λ, t_z **and the sign of** t_x **by ignoring lens distortion:*** Let us redefine

$$u = r_{13} x d_{nf} + r_{23} y d_{nf} \tag{31}$$

$$v = -(r_{11} x d_{nf} + r_{21} y d_{nf}) \tag{32}$$

Then from Eq. 4, we have $x d_f = \frac{v\lambda}{u - \lambda}$. Also, if we ignore lens distortion, then world point P_l and frontal image point P_f can be related as

$$x d_f = -\lambda \frac{x_l}{z_l} \tag{33}$$

Replacing for $x d_f$ we get

$$\frac{v\lambda}{u - \lambda} = -\lambda \frac{x_l}{\underbrace{w + t_z}_{z_l}} \tag{34}$$

where $w = s_{31} x_w + s_{32} y_w + s_{33} z_w$ from Eq. 2. This equation can be simplified to set up the following linear equation

$$\begin{bmatrix} -x_l & v \end{bmatrix} \begin{bmatrix} \lambda \\ t_z \end{bmatrix} = -u x_l - v w \tag{35}$$

where, (u, v) are functions of R from Eq. 31-32 and (x_l, w) are functions of t_x from Eq. 2, Eq. 18 and Eq. 23-25.

Now, given multiple world-image point observations, Eq. 35 can be solved for (λ, t_z) using each of the four possible values of $\{U_e, U_i\}$. Graphically, the four possible solutions to $\{U_e, U_i, \lambda, t_z\}$ can be visualized in Fig. 4, where on the left we have the ground truth imaging and on the right are the four imaging hypothesis labeled as A, B, C and D. As can be seen all four solutions satisfy the perspective (we had assumed no distortion earlier) imaging of P_w to P_i but each correspond to different calibration parameters. Based on this analysis, solution C and D can be rejected by checking the sign of λ obtained from Eq. 35 as λ cannot be negative. The correct solution among A and B can be obtained by analyzing model fitting error for radial distortion coefficients as described next.

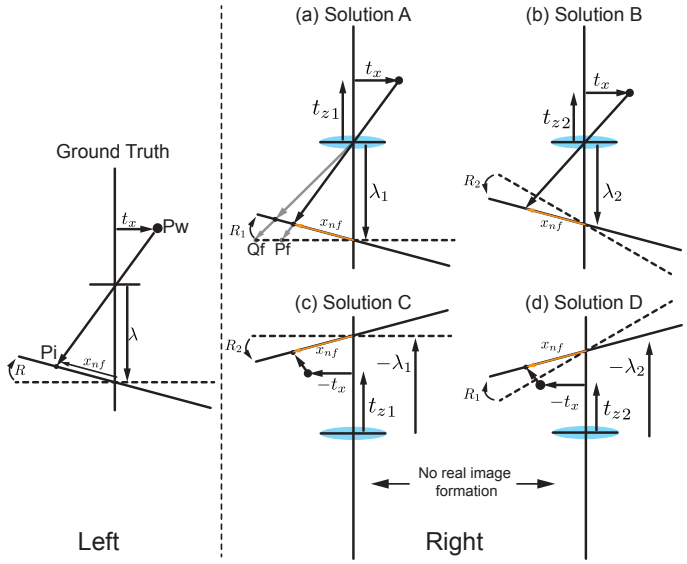

Fig. 4. (Left) Ground truth image formation. (Right) (a) Solution A: $(t_x, R_1, \lambda_1, t_{z1})$. (b) Solution B: $(t_x, R_2, \lambda_2, t_{z2})$. (c) Solution C: $(-t_x, R_2, -\lambda_1, t_{z1})$. (d) Solution D: $(-t_x, R_1, -\lambda_2, t_{z2})$. Solution C and D can be rejected based on λ being negative. The better solution between A and B is selected by analyzing radial distortion model fitting error.

*2) **Determining R**:* From Fig. 4(Right,a-b), we observe that among the two solutions A and B, only solution A coincides with a rotation which will result in a frontal sensor parallel to the lens plane. This implies that the projected frontal points in A will fit the symmetric radial distortion model better than in B. For each set of calibration parameters in A and B, we first compute the radial distortion parameters of (k_1, k_2) by solving the linear equation $P_f - Q_f(1 + k_1 r^2 + k_2 r^4) = 0$ for a set of world-image point observations. Here $Q_f = (x_f, y_f)$ is ideally projected frontal image sensor points, $r^2 = x_f^2 + y_f^2$ and $P_f = (xd_f, yd_f)$. The radial distortion model fitting error E_{rad} can then be obtained as:

$$E_{rad} = P_f - Q_f(1 + k_1 r^2 + k_2 r^4) \tag{36}$$

The solution with least E_{rad} is selected, e.g. in Fig. 4, solution A will get selected. Thus, $R(\rho, \sigma), t_x, \lambda, t_z$ are estimated uniquely. Furthermore t_x can then be used to estimate S uniquely from Eq. 18 and Eq. (23-25). Also, applying t_x to Eq. 26, t_y can be estimated uniquely. Thus, we uniquely determine the calibration parameters $\{U_e = \{S, t_x, t_y\}, U_i = R, \lambda, t_z\} = \{U_2, U_3\}$ (from Eq. 8,16). Next, we estimating the remaining calibration parameters of CoD (I_0, J_0) in U_1.

C. Iterative determination of CoD

The RAC [4] as well as the proposed gRAC are formulated in sensor coordinate system (metric), while the image measurements are in the image coordinate system (pixels). This requires conversion from pixels to metric domain as per Eq. 3, which is a function of $U_1 = \{I_0, J_0, s_x, s_y\}$. Since, the gRAC has rank seven which is same as the size of U_2, there are no additional analytical constraints to determine U_1 completely. Thus, we first assume (s_x, s_y) are known. s_y is typically known [2], [4] as it defines the reference scale over which λ, s_x are defined and s_x can be obtained reliably from the sensor data-sheet.

If the principal point were same as the CoD (I_0, J_0), [5] showed that the residual error in RAC (Sec. III) when applied to measured image points on frontal sensor is quadratic with respect to error in the assumed location (I_0, J_0). Thus, [5] proposed to compute (I_0, J_0) by nonlinear minimization of residual RAC [4] error. But, in our imaging model (Sec. IV), the measured image points are on a non-frontal sensor plane. They need to be converted to frontal sensor coordinates requiring knowledge of R (Eq. 4). While R can be computed from the analytical technique of Sec. V, this technique in turn requires correct estimates of (I_0, J_0). To solve this "chicken and egg problem", we propose an iterative solution similar to [11] as follows:

1) Uniformly sample an image region for (I_0, J_0).
2) For each hypothesized (I_0, J_0) obtain gRAC (Eq. 7).
3) Solve gRAC for R (Sec. V) and obtain frontal coordinates (Eq. 4).
4) Nonlinearly minimize residual RAC [4] obtained from frontal coordinates to get optimal (I_0^*, J_0^*).
5) Compute the difference error $E = abs(I_0 - I_0^*) + abs(J_0 - J_0^*)$, giving an estimate of how good the initial assumed (I_0, J_0) was. Select the point with minimum error E. Stop if E is less then a threshold, otherwise goto Step 6.
6) Refine the sampling around the selected point in Step 5 and repeat Steps 2,3,4,5.

Finally, the initial estimates obtained from gRAC are used as initialization for non-linear refinement of parameters to obtain U^* . This process incorporates radial lens distortion and minimizes the re-projection error over all observed scene and image points [2], [10].

VI. Experiments

Wee present and compare the results of proposed analytical solution to gRAC on synthetic distorted and real data with traditional RAC [4].

A. Synthetic Data

A camera was simulated with intrinsic parameters $\lambda = 8.4$ mm, $\rho = 0, \sigma = 4$ degrees, $s_x = 0.01, sy = 0.01$ mm, $I_0 = 240, J_0 = 320$ pixels, $k_1 = 0.0021966, k_2 = -1.3001e-05$

and extrinsic parameters $\theta = 0.10, \phi = 43.31, \psi = 0.02$ degrees, $t_x = -65.09, t_y = -41.04, t_z = 102.2$ mm. Synthetic world points P_w are generated and projected (Sec. IV) using simulated camera parameters to obtain image points. Then, Gaussian noise with standard deviation $\{0.05, 0.1, \cdots, 1.0\}$ pixels is added to the synthesized image points to simulate measurement error. The gRAC constraint (Eq. 7) is applied and analytical calibration estimates are computed. This procedure is repeated 100 times and the mean of all the trials is taken and compared with the ground truth data. Fig. 5(a-d) shows the relative error(%) in estimation of $R(\rho, \sigma)$, $S(\theta, \phi, \psi)$, t_x, t_y, t_z and λ respectively. The error bars in Fig. 5 indicates the std. dev. in the estimation of respective calibration parameters. The relative error in parameter estimates increases with increasing noise. For lower noise levels, this error as well as the std. dev. is low for all calibration parameters. As, the measurement error in our real data is close to 0.11 pixels, the simulation gives confidence that for real data, gRAC based analytical solution should be robust to image noise.

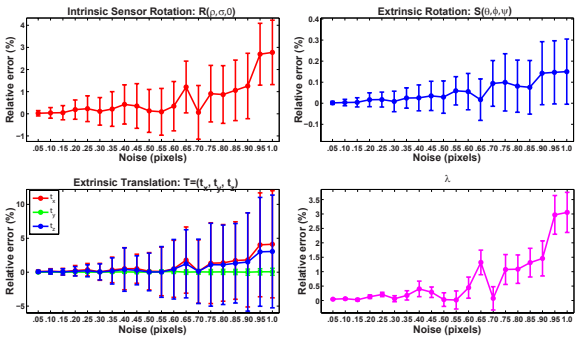

Fig. 5. Relative error vs noise(in pixels) using gRAC on synthetic data.

B. Real Data

The camera used for calibration is a custom made AVT Marlin F-033C camera with sensor tilted $\approx$ 3-4 degrees and acquiring 640×480 resolution images. The corners of a checkerboard (CB) calibration pattern with 20×20 squares of length 5 mm and positional accuracy of .001 mm are used as known 3D scene points. A 2.5D image data is captured by moving the CB along its surface normal and imaging each discrete CB position. A set of 11 such 2.5D datasets are captured by placing the camera at different locations in-front of the CB. The corners in the acquired calibration images are computed using [12]. We compute calibration parameters by using RAC and gRAC and then refine them via non-linear minimization. The results obtained are shown in Tab. I. Comparing the re-projection errors in the last row of Tab. I, we observe that calibration based on the analytical estimates obtained from gRAC leads to smaller re-projection error as compared to traditional RAC. The image center (I_0, J_0) from analytical gRAC has been obtained using the technique proposed in Sec. V-C. It can be seen that it is quite different from the one obtained by RAC [5] indicating that the optic axis is indeed not orthogonal to the lens and thus the sensor is tilted. The analytical tilt estimate from gRAC is 3.81^o and after refinement it is 4.23^o. The small difference arises since analytical solution ignores noise and is thus sensitive to measurement errors.

TABLE I. CALIBRATION ESTIMATES USING THE TWO TECHNIQUES.

Method		RAC [4]		gRAC(proposed)	
		analytical	non-linear	analytical	non-linear
	$\lambda_x = \frac{\lambda}{s_x}$	829.57	823.64	855.25	854.56
	$\lambda_y = \frac{\lambda}{s_y}$	833.63	827.67	855.25	855.19
Principal Point	I_0	225.845	226.15	239.30	238.91
	J_0	331.632	330.53	330.59	330.83
Radial	k_1	–	-0.0021	–	-0.0022
	k_2	–	$2.33e-05$	–	$3.37e-05$
	ρ	–	–	-0.49	0.13
	σ	–	–	**3.81**	**4.23**
Re-projection Error		–	0.082064	–	**0.057119**

VII. CONCLUSIONS

In this paper, we have proposed a generalized radial alignment constraint (gRAC) which takes possible misalignment between lens and sensor planes into account. We have developed an analytical solution to solve the gRAC constraint for a subset of calibration parameters. Then, we have shown that the center of radial distortion can also be computed based on the analytical solution using an iterative approach. Finally, we have shown that non-linear calibration with gRAC initialization leads to lower re-projection error than RAC [4] based initialization.

ACKNOWLEDGMENTS

This work was supported by US Office of Naval Research grant N00014-12-1-0259.

REFERENCES

[1] Y. I. Abdel-Aziz and H. M. Karara, "Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry," in *Proceedings of the Symposium on Close-Range photogrammetry*, vol. 1, 1971.

[2] J. Weng, P. Cohen, and M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *PAMI*, 1992.

[3] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *CVPR*, 1997.

[4] R. Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *IJRA*, 1987.

[5] R. Lenz and R. Tsai, "Techniques for calibration of the scale factor and image center for high accuracy 3d machine vision metrology," in *ICRA*, vol. 4, mar 1987, pp. 68 – 75.

[6] R. T. Held, E. A. Cooper, J. F. O'Brien, and M. S. Banks, "Using blur to affect perceived distance and size," *ACM TOG*, vol. 29, April 2010.

[7] A. Kumar and N. Ahuja, "A generative focus measure with application to omnifocus imaging," in *ICCP*, 2013.

[8] A. Krishnan and N. Ahuja, "Range estimation from focus using a non-frontal imaging camera," in *Proceedings of the eleventh national conference on Artificial intelligence*, ser. AAAI'93, 1993, pp. 830–835.

[9] D. Gennery, "Generalized camera calibration including fish-eye lenses," *IJCV*, 2006.

[10] A. Kumar and N. Ahuja, "Generalized pupil-centric imaging and analytical calibration for a non-frontal camera," in *CVPR*, 2014.

[11] D. Scaramuzza, A. Martinelli, and R. Siegwart, "A toolbox for easily calibrating omnidirectional cameras," in *IROS*, 2006.

[12] J.-Y. Bouguet, "Camera calibration toolbox for matlab," Website, 2000, http://www.vision.caltech.edu/bouguetj/calib_doc/.