Photogrammetric Record, xx(xxx): 000–000 (Month 20##)

MODELLING EXTREME WIDE-ANGLE LENS CAMERAS

DEREK D. LICHTI* (ddlichti@ucalgary.ca)
WYNAND TREDOUX (wynand.tredoux@ucalgary.ca)
University of Calgary, Calgary, Canada
REZA MAALEK (reza.maalek@kit.edu)
Karlsruhe Institute of Technology, Karlsruhe, Germany
PETRA HELMHOLZ (petra.helmholz@curtin.edu.au)
Curtin University, Australia
ROBERT RADOVANOVIC (rradovanovic@mcelhanney.net)
McElhanney Ltd., Calgary, Canada

* *Corresponding author*

Abstract

The use of consumer cameras fitted with extreme wide angle (EWA) lenses for photogrammetric measurement is increasing. Conventional modelling of EWA systems relies on the pinhole camera model and up to five radial lens distortion terms. Aiming to reduce model complexity, this paper reports an investigation into an alternate approach using fisheye lens models for EWA systems, despite them not falling strictly into to the fisheye category. Four fisheye models were tested on four different cameras under laboratory conditions. The self-calibration results show superior model fit for all fisheye models over the pinhole plus radial model in terms of residual RMS. The number radial distortion of terms required for the fisheye models was lower in all cases, so model complexity was reduced. Independent assessment revealed very similar 3D reconstruction accuracy for all models. The results suggest that fisheye modelling is an advantageous alternative for EWA lens systems.

KEYWORDS: camera calibration, extreme wide-angle lenses, fisheye lenses, radial lens distortion, barrel distortion.

INTRODUCTION

WIDE ANGLE LENSES are finding increased photogrammetric use thanks to their rising popularity in consumer cameras and the large field-of-view (FoV) they offer in

comparison to standard lenses. Consumer uses include smart phone cameras, web cams, action cameras and multi-sensor panoramic imaging systems. Photogrammetric interest lies in a number of application areas including, but not limited to, mobile mapping (Rau et al., 2016), unmanned aerial vehicle mapping (Wierzbicki, 2018a), archaeological recording (Ballarin et al., 2015), biomechanics (Bernardina et al., 2017), modelling indoor environments (Teo, 2015) and underwater mapping (Bosch et al., 2015).

Cameras fitted with extreme wide-angle (EWA) lenses are sometimes casually referred to as fisheye systems. This is perhaps natural due to the high degree of barrel distortion found in EWA imagery that is associated with the fisheye effect. Large radial lens distortion, in excess of 100% of the distance between principal point and image format corner, has been reported for EWA cameras (Helmholz and Lichti, 2020; Lichti et al., 2020). However, as pointed out by Miyamoto (1964), a fisheye lens is not "an extension of a wide-angle lens". The severe barrel effect is not an aberration, as it is for a standard lens, but is the result of its specialized design that maps a hemisphere to a plane. So, whilst the appearance of distorted imagery from fisheye and EWA cameras may be similar, the underlying cause is different.

According to Ray (1994), a wide-angle lens is one for which the "effective focal length is significantly less than the format diagonal". Ray (1994) categorizes lenses in terms of the angular FoV computed with the sensor format diagonal (Table I).These range from the standard lens up to the 140° fisheye limit beyond which distortion-free imaging is not possible. Other delineations can be found in the literature. For example, Pernechele (2016) considers a lens having a FoV greater than 100° to be very wide-angle. The terminology of Ray (1994) will be used herein.

TABLE I. Photographic lens classification based on FoV (Ray, 1994).

Category	*FoV Range (°)*
Standard lens	42 - 54
Semi-wide-angle (SWA)	55 - 70
Wide-angle (WA)	71 - 90
Extreme wide-angle (EWA)	91 - 140
Fisheye	> 140

The focus of this paper is modelling EWA cameras that exhibit large barrel distortion. A selection of action cameras, panoramic systems and mobile phones is examined. In recent years, a number of authors have reported on the calibration of action cameras—GoPro cameras in particular—due to their increased ubiquity, compact and lightweight design as well as robustness (Ballarin et al., 2015; Balletti et al., 2014; Bernardina et al., 2017; Hastedt et al., 2016; Helmholz and Lichti, 2020; Teo, 2015; Wierzbicki, 2018b). Similarly, panoramic camera systems featuring multiple, integrated sensors have also received attention in terms of the geometric calibration, notably Ladybug systems (Bosch et al., 2015; Lichti et al., 2020; Rau et al., 2016) and the Panono system (Khoramshahi and Honkavaara, 2018). The predominant means to account for the distortion in such EWA cameras uses the standard central perspective model, the collinearity condition, augmented with the odd-power polynomial model for radial lens distortion. The number of terms employed for the latter varies from two to five.

Fisheye lens cameras require the use of a different projection. The conventional pinhole model is not applicable because the incident (α) and exit chief ray angles (β) preceding the aperture are not preserved (Pernechele, 2016); see Figs. 1a and 1b. Schneider et al. (2009) report a detailed investigation into modelling fisheye systems using four different projection models (equidistant, equisolid angle, orthographic and stereographic), as well as the pinhole model, augmented with radial (and decentring) lens distortion terms. They quantify the performance of each with a number of tests including comparison of photogrammetrically-determined coordinates with independently-determined reference values. Hughes et al. (2010) perform a comparative test of the four fisheye models plus three others: the polynomial fisheye transform, the fisheye transform and the field-of-view model. Their assessment is made in terms of model fit, which varies considerably among the models. Other reports of fisheye lens modelling include calibration of a multi-camera system (Campos et al., 2018) and a study of the impacts of target field geometry on fisheye system calibration quality (Choi et al., 2019).

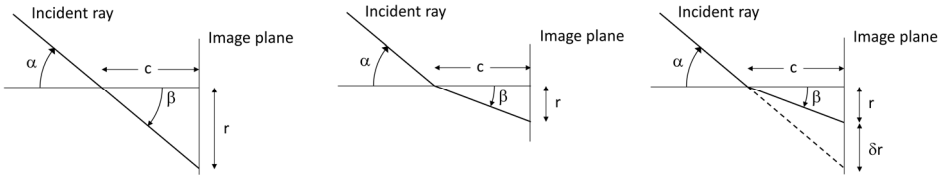

FIG. 1. Chief ray entry (α) and exit (β) angle relationships: a) pinhole camera model; b) fisheye lens model; c) pinhole model with barrel distortion.

The successful use of the central perspective model with a five-term radial lens distortion curve for EWA systems has been demonstrated for GoPro cameras (Helmholz and Lichti, 2020) and the Ladybug system (Lichti et al., 2020). As shown in the latter work, inclusion of the extra terms was needed to completely describe the barrel distortion and to maximize 3D reconstruction accuracy. The additional terms also proved beneficial for generating orthorectified panoramic mosaic imagery of indoor scenes (Jarron et al., 2019a). In spite of these successes, high correlation coefficients (> 0.99) existed among the distortion coefficients.

Despite the different origins of distortion in fisheye and EWA systems, their common attribute of a high degree of barrel distortion has motived this investigation into whether EWA systems can be better modelled by one of the four fisheye projections. The search for a viable alternate model for EWA systems is also driven by the principle that the aforementioned high correlations between radial lens distortion coefficients, which indicate strong functional dependencies, should be reduced wherever possible. The hypothesis of the fisheye modelling approach is that a significant amount of the barrel distortion found in EWA systems, indicated by δr in Fig. 1c, can be compensated by the fisheye projection model. Any residual radial lens distortion is expected to be of much lower magnitude and can be modelled with fewer terms than are required for the pinhole model. This reduction of model complexity should alleviate the parameter correlation issue.

An initial investigation into the use of the equidistant fisheye model for Ladybug and GoPro self-calibration showed promising results in terms of model fit and accuracy

(Jarron et al., 2019c). Here, a much more detailed and comprehensive investigation using 22 datasets captured with 14 separate EWA cameras from four different manufacturers is reported. All four fisheye projection models as well as the pinhole and radial lens distortion model (hereafter referred to as the pinhole+radial model) were tested. The effectiveness of the modelling is quantified in terms of model fit, parameter precision and correlations and the accuracy of 3D reconstruction. To this end, the next section features a review of relevant models as background for the proposed EWA system modelling approach. The experiments are then described in terms of camera hardware, calibration facilities and software for image point measurements and self-calibrating bundle adjustments. This is followed by presentation and detailed analyses of the experimental results and, ultimately, conclusions about the investigation.

EWA MODELLING

Pinhole Camera Model

The geometry of the pinhole camera model corresponding to a standard, distortion-free lens is depicted in Fig. 2. In this situation the entry angle, α, and the exit angle, β, are equal (α=β; Fig. 1a) so the collinearity equations are applicable.

$$x_{ij} + \varepsilon_{x_{ij}} = x_{p_j} - c_j \frac{U_{ij}}{W_{ij}} + \Delta x_{ij} \tag{1}$$

$$y_{ij} + \varepsilon_{y_{ij}} = y_{p_j} - c_j \frac{V_{ij}}{W_{ij}} + \Delta y_{ij} \tag{2}$$

where

$$\begin{pmatrix} U_{ij} \\ V_{ij} \\ W_{ij} \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \begin{pmatrix} X_i - X_j^c \\ Y_i - Y_j^c \\ Z_i - Z_j^c \end{pmatrix} \tag{3}$$

$\left(x_{ij}, y_{ij}\right)$ are the observations of object point i in image j; $\left(X_i, Y, Z_i\right)$ are the object space coordinates of point i; $\left(X_j^c, Y_j^c, Z_j^c\right)$ are the object space coordinates of the perspective centre (PC) of image j; m_{pq} are the elements of the rotation matrix from object space to image space; $\left(x_{p_j}, y_{p_j}\right)$ and c_j are the principal point offset and principal distance, respectively, of image j; $\left(\varepsilon_{x_{ij}}, \varepsilon_{y_{ij}}\right)$ are additive random error terms; and $\left(\Delta x_{ij}, \Delta y_{ij}\right)$ are additive systematic error terms.

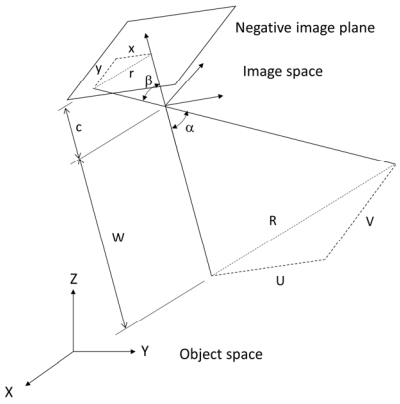

FIG. 2. Central perspective imaging geometry.

Fisheye Lens Models

The FoV of a lens can be increased by decreasing the focal length, but this comes at the cost of a significant drop off in illumination at the image periphery. Thus, a different projection is needed to extend the FoV beyond the fisheye limit (Ray, 1994). Four projections can be used for the design of fisheye lenses: equidistant; equisolid angle; orthographic and stereographic. In all cases, $\alpha \neq \beta$ (Fig. 1b). Whereas in the central perspective model the relationship between radial image distance, r, and incidence angle, α, is given by $r = c\tan(\alpha)$, the models for the equidistant, equisolid angle, orthographic and stereographic projections are $r = c\alpha$, $r = 2c\sin(\alpha/2)$, $r = c\sin(\alpha)$ and $r = 2c\tan(\alpha/2)$, respectively. The observation equations corresponding to these four designs are given as follows:

Equidistant

$$x_{ij} + \varepsilon_{x_{ij}} = x_{p_j} - c_j \frac{U_{ij}}{\sqrt{U_{ij}^2 + V_{ij}^2}} \arctan\left(\frac{\sqrt{U_{ij}^2 + V_{ij}^2}}{W_{ij}}\right) + \Delta x_{ij} \tag{4}$$

$$y_{ij} + \varepsilon_{y_{ij}} = y_{p_j} - c_j \frac{V_{ij}}{\sqrt{U_{ij}^2 + V_{ij}^2}} \arctan\left(\frac{\sqrt{U_{ij}^2 + V_{ij}^2}}{W_{ij}}\right) + \Delta y_{ij} \tag{5}$$

Equisolid angle

$$x_{ij} + \varepsilon_{x_{ij}} = x_{p_j} - c_j \frac{2U_{ij}}{\sqrt{U_{ij}^2 + V_{ij}^2}} \sin\left\{\frac{1}{2}\arctan\left(\frac{\sqrt{U_{ij}^2 + V_{ij}^2}}{W_{ij}}\right)\right\} + \Delta x_{ij} \tag{6}$$

$$y_{ij} + \varepsilon_{y_{ij}} = y_{p_j} - c_j \frac{2V_{ij}}{\sqrt{U_{ij}^2 + V_{ij}^2}} \sin\left\{\frac{1}{2}\arctan\left(\frac{\sqrt{U_{ij}^2 + V_{ij}^2}}{W_{ij}}\right)\right\} + \Delta y_{ij} \tag{7}$$

Orthographic

$$x_{ij} + \varepsilon_{x_{ij}} = x_{p_j} - c_j \frac{U_{ij}}{\sqrt{U_{ij}^2 + V_{ij}^2}} \sin\left\{\arctan\left(\frac{\sqrt{U_{ij}^2 + V_{ij}^2}}{W_{ij}}\right)\right\} + \Delta x_{ij} \tag{8}$$

$$y_{ij} + \varepsilon_{y_{ij}} = y_{p_j} - c_j \frac{V_{ij}}{\sqrt{U_{ij}^2 + V_{ij}^2}} \sin\left\{\arctan\left(\frac{\sqrt{U_{ij}^2 + V_{ij}^2}}{W_{ij}}\right)\right\} + \Delta y_{ij} \tag{9}$$

Stereographic

$$x_{ij} + \varepsilon_{x_{ij}} = x_{p_j} - c_j \frac{2U_{ij}}{\sqrt{U_{ij}^2 + V_{ij}^2}} \tan\left\{\frac{1}{2}\arctan\left(\frac{\sqrt{U_{ij}^2 + V_{ij}^2}}{W_{ij}}\right)\right\} + \Delta x_{ij} \tag{10}$$

$$y_{ij} + \varepsilon_{y_{ij}} = y_{p_j} - c_j \frac{2V_{ij}}{\sqrt{U_{ij}^2 + V_{ij}^2}} \tan\left\{\frac{1}{2}\arctan\left(\frac{\sqrt{U_{ij}^2 + V_{ij}^2}}{W_{ij}}\right)\right\} + \Delta y_{ij} \tag{11}$$

Note that whereas all models are formulated for observations made in the positive image plane, the negative image plane is shown in Figs. 1 and 2 for the sake of clarity.

Lens Distortion Models

The radial lens distortion model used herein comprises up to six terms of the odd-power polynomial.

$$\Delta x = \bar{x}\left(k_1 r^2 + k_2 r^4 + k_3 r^6 + k_4 r^8 + k_5 r^{10} + k_6 r^{12}\right) \tag{12}$$

$$\Delta y = \bar{y}\left(k_1 r^2 + k_2 r^4 + k_3 r^6 + k_4 r^8 + k_5 r^{10} + k_6 r^{12}\right) \tag{13}$$

Whereas standard, SWA and WA lenses may require up to three terms, more are required for modelling EWA lens cameras in combination with the collinearity condition. High correlation exists among the individual parameters, but generally does not exist between these coefficients and the other interior orientation parameters or the exterior orientation parameters (Fraser, 1997).

Decentring distortion is modelled with the Brown-Conrady model

$$\Delta x = p_1\left(r^2 + 2\bar{x}^2\right) + 2p_2\bar{x}\bar{y} \tag{14}$$

$$\Delta y = p_2\left(r^2 + 2\overline{y}^2\right) + 2p_1\overline{xy} \tag{15}$$

that is often displayed graphically in terms of the profile function

$$P(r) = \sqrt{p_1^2 + p_2^2}\, r^2 \tag{16}$$

EXPERIMENTS

Cameras

Four camera systems, namely the Sony ICX655, GoPro Hero5, Samsung Galaxy Note 10+ and Cobra Adventure HD, were used in this investigation. Some of their salient details are summarized in Table II. The Ladybug5 is a panoramic imaging system comprising six integrated Sony ICX655 cameras. Five of the cameras (LB-0 to LB-4) are horizontal looking whilst the sixth (LB-5) is oriented toward zenith. Each camera was calibrated independently without any relative orientation constraints that enforce the rigid construction of the system. Two sets of GoPro Hero5 Black action cameras were tested: four from Curtin University and two from the University of Calgary. The Curtin University cameras (CU-1 to CU-4) were calibrated in air but each was inside a waterproof housing, which effectively added a planar port to the optical system. The two from the University of Calgary (UC-1 and UC-2) were calibrated without any housing. All six GoPro cameras were operated in photo mode with wide FoV. Images from the Samsung Galaxy Note10+ (SG) were captured in ultrawide mode with "ultra wide shape correction" (lens distortion correction) disabled. The Cobra action camera (CA) was operated in single shot, photo mode at 12 M image resolution.

The fields-of-view listed in Table II were computed using the principal distance obtained from the pinhole+radial self-calibration adjustment. Mean values from multiple datasets were used where applicable. Based on these computations, the Sony, GoPro and Samsung cameras fall into the EWA category. The manufacturer's FoV specification for the Cobra is 120°, suggesting it is an EWA system, but the computed value places it in the wide-angle category. It was selected for the experiments due to its high degree of barrel distortion and is henceforth denoted as an EWA camera based on the manufacturer's characterization.

TABLE II. Tested camera details.

	Abbreviated name	*Number tested*	*Sensor size (pixels)*	*Nominal c (mm)*	*Pixel spacing (μm)*	*Diagonal FoV (°)*
Sony ICX655 (Ladybug5 cameras)	LB	6	2448 × 2048	4.4	3.45	104
GoPro Hero5 Black with plane port	CU	4	4000 × 3000	3	1.54	110
GoPro Hero5 Black	UC	2	4000 × 3000	3	1.54	110
Samsung Galaxy Note10+	SG	1	4608 × 3456	1.8	1.0	115

Cobra Adventure HD (COBRA5200)	CA	1	4032 × 3024	5	1.45	72

Target Fields

All self-calibrations were performed from imagery captured over signalized target fields. With the exception of the four Curtin GoPro cameras, all datasets used a dedicated indoor calibration facility at the University of Calgary that is described by Lichti et al. (2020). This temperature-controlled, 11.6 m × 11.3 m × 4.1 m room has some 500 targets mounted on the walls, floor and ceiling (Fig. 3). Each target comprises a white circle (several diameters are used) with black background printed directly onto a 4 mm thick extruded polypropylene board. The full Calgary facility was used for the calibration of the Ladybug system's cameras, which was performed twice over a ten-month period. Two orthogonal planes comprising a 4.0 m × 1.9 m × 3.3 m A-shape configuration (Choi et al., 2019) were used for the two Calgary GoPro, the Samsung and the Cobra cameras.

The Curtin calibration data were collected over a 3.8 m × 2.5 m × 0.7 m field of some 150 targets (Fig. 4). In this case, the targets comprise a 5 mm retro-reflective circle with a black background. Though nominally planar, depth variation was present thanks to targets placed on a structural column and a removable pole at the network periphery and the scale bar in the foreground. Three datasets were collected consecutively with one of the Curtin GoPros, while single datasets were captured with the other three.

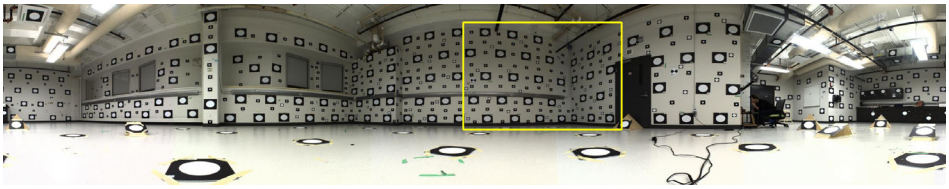

FIG. 3. University of Calgary target field. The yellow box indicates the A-shape section used for some of the camera calibrations.

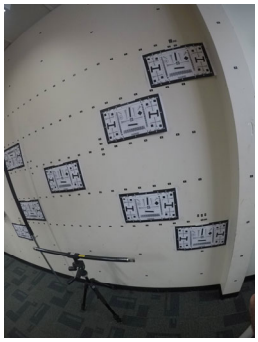

FIG. 4. Curtin University target field.

Strong first order design principles were followed for the image capture with each camera. These measures included convergent imagery, rolled imagery and filling the format with targets. The network geometry for the zenith-looking camera of the first Ladybug dataset was weaker than that of the horizontal cameras since it lacked scale

variation in the imagery. However, it was stronger in terms of superior roll diversity in comparison to the horizontal cameras. The second Ladybug calibration dataset was collected with additional targets as well as tilted imagery to improve the network geometry, particularly in terms of roll diversity of the horizontal cameras.

Target Measurement and Self-Calibration

Target centre measurement was dataset dependent. The ellipse fitting method reported by (Jarron et al., 2019b) was used for the Ladybug imagery. Target centres were measured using the centroid method available in the *Australis* software (v. 8.33) for the Curtin GoPro camera imagery. All other target measurements were made with the ellipse fitting method with target eccentricity correction described by Maalek and Lichti (2021).

Imagery output by the Cobra camera exhibited strong overshoot in brightness values at edges. Though the exact cause of this artefact is not known, a possible source is interpolation of the brightness values in an on-board up-sampling process from the lower native resolution of the acquired imagery. To remove this artefact, the Cobra camera imagery was lowpass filtered with Gaussian filter (σ=4 pix) prior to target measurement.

Custom software was written to perform the self-calibrating bundle adjustments. The user has the ability to select one of the five models to use. All five were applied to every single dataset. The program simultaneously solves for the interior orientation parameters (IOPs; the principal point offset, the principal distance and lens distortion coefficients); the exterior orientation parameters (EOPs) and object point coordinates. The x- and y-observations were weighted equally in each adjustment with the same standard deviation used for all five models for a given camera. All adjustments were minimally-constrained using inner constraints. A sequential model identification process was performed for each dataset in order to determine the exact number of radial distortion terms needed for a given camera and projection model combination (Lichti et al., 2020). Two decentring distortion terms used for all datasets and all projection models.

Accuracy Assessment

Independent accuracy assessment was performed for all cameras used at the University of Calgary. Additional images collected for this purpose were excluded from the self-calibrations. For each of the two Ladybug datasets, ten additional images (two per horizontal camera) were acquired from two stations (2.8 m separation) within the calibration room. For the GoPro, Samsung and Cobra cameras, an extra a stereo pair of images was collected. The convergence angle ranged between 12° and 26°.

Each set of images was oriented by bundle adjustment in which four well-distributed targets were treated as fixed control points and the remaining targets were estimated as unknown tie points. The IOPs determined from the self-calibration were treated as constants. Note that all ten images from each Ladybug dataset were included in the same adjustment. The tie point coordinates were independently determined to millimetre-level accuracy using a Faro Focus3D terrestrial laser scanner (TLS; Lichti et al., 2019b). Differences between tie point coordinates and TLS coordinates were computed to quantify the 3D reconstruction accuracy of each of the five projection models.

RESULTS

Self-Calibration Summary Data

Pertinent metadata for the self-calibration datasets are provided in Table III. They are summarized with ranges where there are multiple cameras and/or repeat datasets. All datasets are highly redundant as can be seen from the degrees-of-freedom (dof) entries, which are only indicated for the pinhole+radial model for the sake of brevity. The low-end values shown for the Ladybug camera correspond to the zenith-looking camera. The upper-end values are representative of the quantities for the five horizontal cameras.

TABLE III. Self-calibration dataset metadata. Numbers in parentheses indicate repeat sets. dof indicates degrees-of-freedom.

Camera (dataset)	*Images per camera*	*Number of Object points*	*Image points*	*Pinhole+radial model dof*
LB (1)	40-44	53-110	976-1078	1376-1596
LB (2)	29-37	85-289	539-1829	646-2596
CU (all)	12	149-154	1045-1162	1562-1791
UC (all)	10	129-144	1088-1147	1704-1799
SG	10	152	1271	2022
CA	10	123	861	1291

Model Fit

The self-calibrating bundle adjustment residuals indicate how well the observations fit a given functional model. As such, they are important quantities to analyse when comparing the performance of different models. If all systematic errors have been modelled and outliers have been identified and removed, then the residuals represent random error estimates. The quality of the fit of the observations to each model is measured with the RMS_{xy} of the self-calibration residuals (Fig. 5). The values range from 0.14 pix to 0.65 pix. Comparison of the performance between cameras is not made here since it is a function of camera and target measurement algorithm. However, the relative performance of the five projection models can be compared.

The results show that, without exception, the pinhole+radial camera model yielded the poorest quality fit. The RMS_{xy} is at least 50% worse than the fisheye models in most cases. Comparison of the results for the six Ladybug cameras shows the worst pinhole results were realized for the zenith-looking camera (LB-5), for which network geometry was weaker. These findings suggest that the use of the pinhole+radial model for EWA cameras should be avoided.

Each fisheye model offers considerable improvement in model fit over the pinhole+radial model. The top performing model in terms of RMS_{xy} was the orthographic for all Ladybug cameras datasets, the Calgary GoPros, the Samsung and the Cobra. This was followed by the equisolid angle and then the equidistant. The stereographic model performance ranked fourth (second-worst) in all cases. For all Curtin GoPro datasets the lowest RMS_{xy} was the equisolid angle model. The second-ranked model was the equidistant in four cases and the orthographic in two. The different performance of the Curtin GoPro cameras relative to the Calgary GoPros may be related to the presence of

the plane port. Overall, the orthographic provides the best fit in 16 out of 22 datasets (73% of the cases), which would suggest it is the preferred model to use for these EWA cameras.

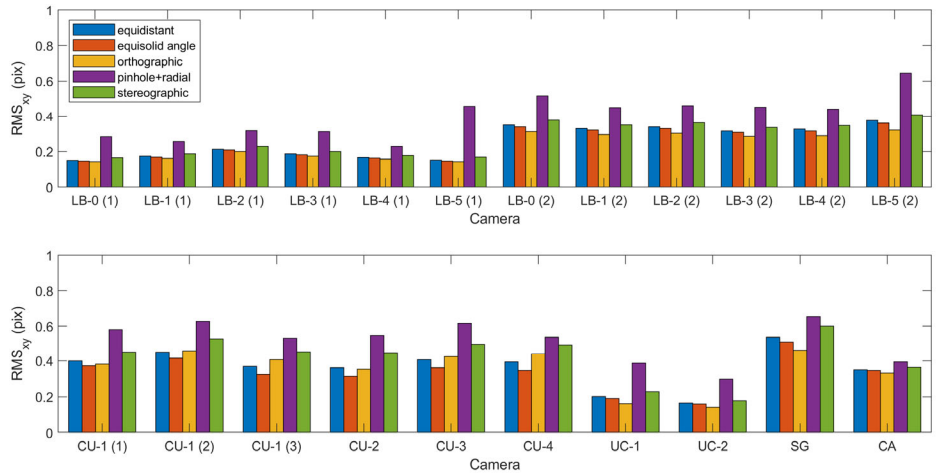

FIG. 5. Self-calibration residual RMS_{xy}. Numbers in parentheses indicate repeat sets.

Principal Distance

The principal distance is the most fundamental parameter defining the interior geometry of a camera, so a comparative analysis of the estimates obtained from the different projection models is warranted. The estimated principal distance and its precision (95% confidence error bars) for all cameras and all projections are presented in Fig. 6. Note that constants have been subtracted from the Samsung and Cobra estimates for display purposes. The amounts to be added are indicated in the figure. Also note the small size of the error bars for many of the datasets.

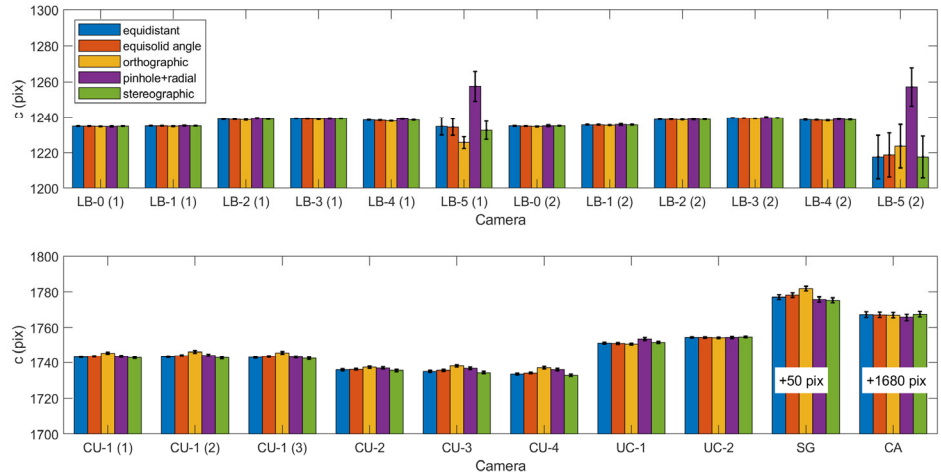

FIG. 6. Estimated principal distances and 95% confidence intervals. Numbers in parentheses indicate repeat sets.

In general, little variation is evident in the principal distances for a given camera. The exception to this observation is again the zenith-looking Ladybug camera, for which estimates vary considerably. The pinhole+radial model principal distances in particular appear to be larger than might be expected. This is likely due to the weaker geometry that was realized for the calibration of the vertical camera. The lower precision supports this hypothesis. It is interesting that the pinhole+radial model is most affected by the weaker network geometry. This suggests that the fisheye models may be more robust in this regard, but this would require further investigation. Some statistically significant differences exist between principal distance estimates from different models for a given camera, but there does not appear to be a pattern common to all cameras.

Principal Point Offset

Estimates of the principal point offset coordinates, also fundamental IOPs, are shown for all models with 95% confidence regions (ellipses) in Fig. 7. In many cases the estimates from each projection model are tightly clustered together and not significantly different. From all 22 datasets, the mean difference in principal point offset among the five projection models is 0.98 pix in x and 0.97 pix in y, with respective maximum values of 4.0 pix and 2.7 pix. Some significantly different principal point offset estimates do exist. There are several cases for the pinhole+radial model and a couple for the orthographic model, though a systematic pattern is difficult to discern. Unlike the situation found for the principal distance, these outlier cases are not simply confined to the zenith-looking Ladybug camera. Also of note is the greater principal point precision for the vertical camera in the first Ladybug dataset.

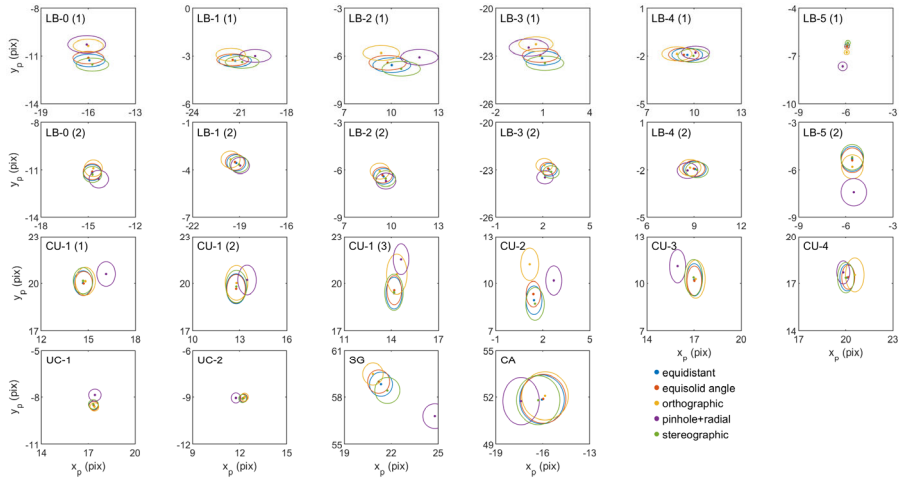

FIG. 7. Estimated principal point offsets and 95% confidence ellipses. Numbers in parentheses indicate repeat sets. The range of all axes is the constant: 6 pix in both x and y.

Radial Lens Distortion

The premise of the fisheye modelling approach for EWA systems is that the barrel distortion can be largely compensated by the projection model and any residual systematic error can be modelled with fewer radial lens distortion terms. The number of terms required for each camera and each model is summarized in Table IV. Representative distortion profiles for each camera model are shown in Fig. 8. The amount of distortion is plotted in terms of proportion of the radial distance from image centre to format corner in order to allow direct comparison of results. Maximum distortion magnitudes are provided in Table IV.

TABLE IV. Required number of radial lens distortion coefficients (k) and the maximum distortion magnitude (δr_{max}). Mean values are reported for multiple cameras and/or repeats. Distortion magnitudes have been rounded to the nearest integer for compactness. Numbers in parentheses indicate repeat sets.

Camera	*Equidistant*		*Equisolid angle*		*Orthographic*		*Pinhole+radial*		*Stereographic*	
(dataset)	*k*	δr_{max} *(pix)*	*k*	δr_{max} *(pix)*	*k*	δr_{max} *(pix)*	*k*	δr_{max} *(pix)*	*k*	δr_{max} *(pix)*
LB (1)	2	4	2	110	2	411	5	2636	2	264
LB (2)	2	6	2	112	2	412	5	2617	2	262
CU	2	229	2	384	2	806	5	3931	2	156
UC	3	218	3	376	3	810	5	3767	3	168
SG	4	948	4	1044	4	1288	6	451	4	755
CA	3	306	3	229	2	1	4	1173	3	475

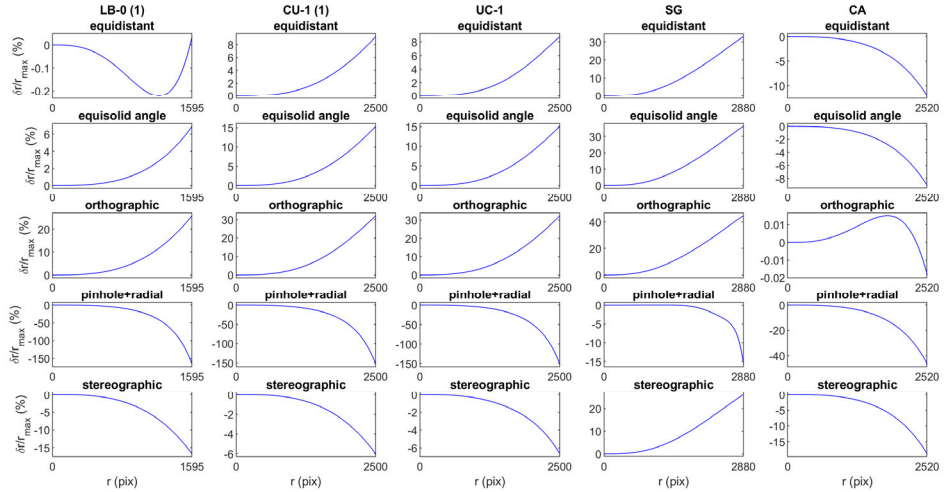

FIG. 8. Sample radial lens distortion curves for each camera type. Numbers in parentheses indicate repeat sets.

As expected, the pinhole+radial model required the most terms: five for all cameras except the Cobra, which required four, and six for the Samsung. Whereas the Ladybug and the GoPro cameras exhibited more than 150% distortion (equivalent to several thousand pixels) at the format corner, the amount of distortion in the Samsung only reached 15% (about 450 pixels). Although the magnitude of the Samsung distortion was

much lower than that of the other cameras, it required more terms to completely describe the systematic effect. The Cobra camera distortion reached nearly 1200 pixels at the format corner.

The fisheye models required fewer distortion terms: two or three in all cases except for the Samsung, which needed four. This outcome is in line with the objective of reducing model complexity. While all pinhole+radial model distortion profiles represent barrel (negative distortion), the type of curve profile varied with both the camera and the fisheye model used. The change in distortion from barrel to pincushion (or compound) suggests overcompensation of the barrel distortion by the fisheye model. This is not seen as a problem since model complexity is reduced. Furthermore, the magnitude of the distortion is a function of these two factors. In general, the use of a fisheye projection model led to a distortion curve with lower maximum magnitude. In some cases, the reduction was several orders of magnitude. The exception is the Samsung camera, for which all fisheye models increased the amount of distortion by up to three times. However, fewer terms were required to effectively model the Samsung distortion profiles.

A good example of the model and camera dependence is the equidistant projection. Both the GoPro and Samsung cameras' curves show pincushion distortion. Barrel distortion is found for the Cobra and the Ladybug example shows a compound pattern. The equidistant projection seems to be the most appropriate model for the Ladybug camera since the distortion magnitude is very low, only a few pixels at the format corners. It is possible that the lens coupled to the Sony camera is in fact of equidistant design, but this is not known with any certainty. All equisolid angle model distortion curves are pincushion type except for Cobra, which is barrel in nature. All orthographic curves are also pincushion except once again the Cobra for which the profile changes sign. Furthermore, the amount of distortion in this case is very low, less than one pixel at the format corner. This may be an indication that the Cobra camera's lens is of orthographic design, but again this is not known. Finally, all stereographic model distortion curves are of the pincushion variety except for the barrel profile of the Samsung.

Decentring Lens Distortion

The decentring distortion coefficients were found to differ between cameras and projection models. All decentring distortion effects were small. The profile function (Equation 16) was computed at the maximum radial distance for all datasets and all projection models. The maximum value of $P(r)$ was less than one pixel for 81 out of 110 cases (73.6%) and less than two pixels in 108 cases (98.2%). In only two cases did $P(r)$ exceed the latter threshold. Both were the from the pinhole+radial model: one of the Ladybugs (2.35 pix) and the Samsung (4.08 pix). No further analyses are conducted due to the low magnitude of the decentring distortion.

Parameter Correlations

A key indication of the quality of a functional model is the correlation among variables. High parameter correlations indicate functional dependencies that can degrade the condition of the normal equations matrix and, ultimately, can compromise object

space reconstruction accuracy. The ideal correlation matrix is an identity matrix, though this structure is seldom achieved in practice. In the context of self-calibration, correlations between EOPs and IOPs are of paramount interest. Correlation among the set of IOPs must also be given due consideration.

With such a large number of self-calibration adjustments (110 cases in total) and parameters, a tractable means to effectively analyse the correlations is required. For each adjustment, the EOP-IOP parameter correlation matrix for each image was computed. Next, the mean of the absolute values of all entries was computed from this set of matrices. The maximum off-diagonal value of the resulting matrix was then used for the comparison. For the IOP-IOP correlations, the largest (in magnitude) off-diagonal value was determined for each calibration case. It is recognized that these values are generalizations of the complex functional relationships between variables and there will be individual coefficients larger than the mean values. However, these measures should be effective for highlighting any large differences between models.

The EOP-IOP correlations (Fig. 9) are generally low; only three exceed 0.8. The largest values are from the first Ladybug calibration dataset. The improvements to the network design are obvious in the overall lower correlations of the second dataset. Strong network design is also evident for the other datasets. In almost all cases, the largest correlation is between the rotation angles and the principal point offset, but there are a few exceptions. The principal distance of the vertical Ladybug camera (LB-5) was correlated with the perspective centre coordinates due to the aforementioned geometric weakness. For the first Ladybug dataset, the highest mean correlation was between k_1 and the PC for the equisolid angle, equidistant and orthographic models but was between the rotation angles and principal distance for the pinhole+radial , and k_1 and the rotation angles for the stereographic. For all cameras of the second Ladybug dataset the highest correlation was between the PC and k_1. In some of the Calgary GoPro and Samsung cases the highest mean correlation was between the PC and the rotation angles.

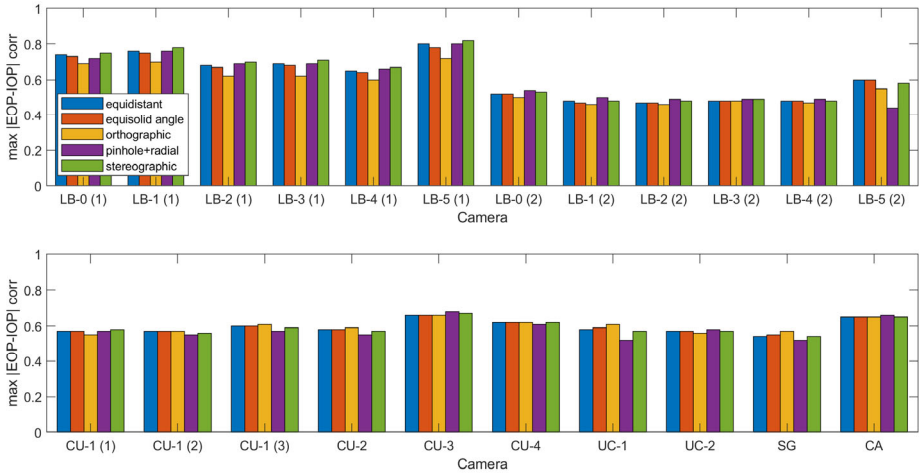

FIG. 9. Maximum EOP-IOP correlation coefficients. Numbers in parentheses indicate repeat sets.

The differences in the largest correlations between models are small overall, less

than 0.1. The largest variation (up to 0.16) is for the vertical Ladybug camera. This is an exceptional case due to inherently weaker network geometry. Among the 22 datasets, the orthographic model ranked first in terms of lowest EOP-IOP maximum correlation 12 times. The equidistant and pinhole+radial models ranked first on three and seven occasions, respectively. The equisolid angle and stereographic models never produced the lowest maximum correlation. Although there is no clear "winner", it can be stated that overall there is little effect of projection model choice on EOP-IOP functional dependence.

The maximum IOP-IOP correlations (Fig. 10) exist between the radial lens distortion terms and are generally much larger in magnitude. With the exception of the two vertical Ladybug camera cases, the largest correlation in all cases is attributed to the pinhole+radial model. This is expected due to the large number of terms required for the distortion profile. So, the aim of reducing functional dependencies is realized by using the fisheye models that require fewer distortion coefficients. The improvements are, however, modest because the basis functions comprising the polynomial (i.e. the individual terms of Equations 12 and 13) do not constitute an orthogonal basis. Therefore, high correlations will always exist even in a lower-order radial lens distortion model.

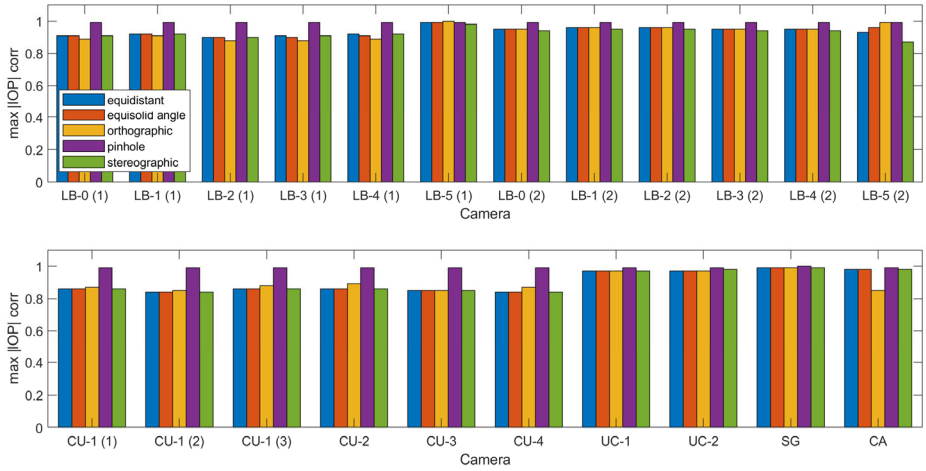

FIG. 10. Maximum IOP-IOP correlation coefficients. Numbers in parentheses indicate repeat sets.

Deeper insight into the dependence among IOPs is provided in Figure 11, which shows the maximum correlation between the principal distance and radial lens distortion terms for each dataset and each model. Beginning with the horizontal Ladybug cameras, the highest correlation is experienced by the pinhole+radial model, but the coefficients are not particularly high: 0.8 or less for the first dataset and 0.9 or less for the second. The next-highest correlations are for the orthographic model. Interestingly, these correlations are very low for the pinhole+radial model of both vertical Ladybug camera datasets. In this degenerate case, the strong functional dependence instead exists between the principal distance and the EOPs. Conversely, all fisheye models, for which principal distance-EOP correlations are lower, exhibit high correlations. For all other cameras' datasets, the orthographic model exhibits the highest correlation, but it is 0.8 or less. All

others models show much lower correlation with no discernible pattern between them apart from the stereographic being the lowest.

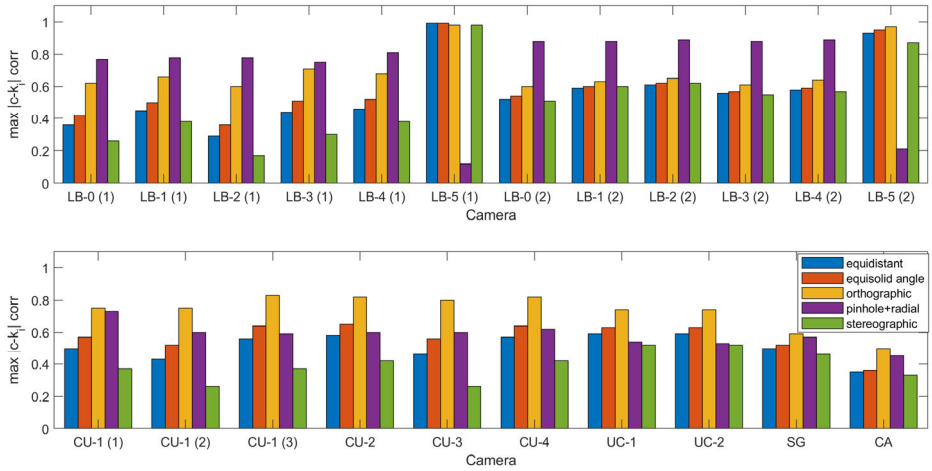

FIG. 11. Maximum principal distance-radial lens distortion parameter correlation coefficients. Numbers in parentheses indicate repeat sets.

3D Reconstruction Accuracy

The aim of photogrammetry can be simply stated as the reconstruction of a scene from imagery. Accordingly, the ultimate test of an alternate functional model is the accuracy with which a 3D scene can be reconstructed. This has been done under controlled conditions by comparing photogrammetrically-determined and independently-determined coordinates of signalized targets. Reconstruction accuracy for each model is quantified by the $RMSE_{XYZ}$ of the differences between the two sets of coordinates.

Prior to comparing model performance, a reference case is presented for the pinhole+radial model. Table V reports the reconstruction accuracy for a basic set of IOPs, determined from separate self-calibration adjustments, and for the full set of IOPs. The basic IOP set varies depending on the camera. The parameters shown were those required to achieve convergence of the solution. Up to two orders of magnitude improvement was achieved for the GoPro and Cobra cameras. About 25 times improvement was realized for the Ladybug. This factor is slightly lower since some of the radial lens distortion—the main source of bias—is modelled by the basic IOP set. The Samsung improvement (5x) is lower simply due to the lower amount of lens distortion in this camera.

The final column of Table V provides context. Object point precision for each network is measured by the mean semi-major axis of the 95% confidence ellipsoids from the pinhole model. As a result of the modelling, all $RMSE_{XYZ}$ values are lower than the expected precision. Overall, these results give a very clear indication of the improvement that is realized by performing the rigorous sensor modelling, even with a large number of radial lens distortion terms.

TABLE V. Model improvement: full IOP set vs. basic IOP set, pinhole+radial model only. Numbers in parentheses indicate repeat sets.

	Basic IOP set		*Full IOP set*		
Camera(s)	*Parameters*	*$RMSE_{XYZ}$ (mm)*	*$RMSE_{XYZ}$ (mm)*	*Improvement*	*$a_{95\%}$ (mm)*
LB-0 to LB-4 (1)	x_p, y_p, c, k_1 (k_2 for 2 cameras)	131.5	5.1	26x	19.7
LB-0 to LB-4 (2)	x_p, y_p, c, k_1 (k_2 for 1 camera)	137.2	5.8	24x	20.9
UC-1	x_p, y_p, c	209.6	2.2	95x	3.5
UC-2	x_p, y_p, c	344.3	2.1	163x	3.4
SG	x_p, y_p, c	23.5	4.7	5x	4.8
CA	x_p, y_p, c	196.8	1.5	129x	2.0

Fig. 12 shows the accuracy achieved for each of the five functional models. Although the pinhole+radial model performs worse in terms of RMS_{XYZ} by about a millimetre in a couple cases, overall the results are very similar among the different models. There is not one model that clearly stands out as being superior to the rest. Furthermore, the achieved accuracy is at a similar level to the accuracy of the TLS reference coordinates, so a definitive statement about which model is best cannot be made. What is clear is that all models perform at about the same level under the controlled conditions of this experiment. Importantly, there is no degradation in accuracy as a result of using the fisheye lens projection models for EWA cameras.

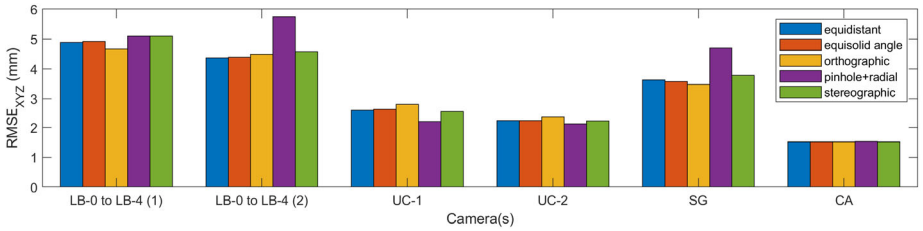

FIG. 12. Accuracy assessment statistics.

CONCLUSIONS

Camera systems having extreme wide angle lenses feature high degrees of barrel distortion that must be modelled if they are to be used for photogrammetry. Modelling EWA systems with the conventional pinhole camera model has the drawback of model complexity, specifically the large number of radial lens distortion terms required. This paper has reported an investigation into the use of four fisheye camera models (equidistant, equisolid angle, orthographic and stereographic) as possible alternatives to the pinhole model in order to reduce model complexity. Experiments were conducted on four different camera models in controlled environments.

All four fisheye models were found to provide superior data fits, as measured by the RMS of residuals, in comparison to the pinhole+radial model. The differences among fisheye models were small, but the orthographic model provided the best fit in 73% of the cases. In terms of the basic interior orientation parameters, all five models produced very similar principal distance and principal point offset estimates. A few principal point offset

estimates differed by more than a pixel, though no systematic pattern was observed.

It was found that some of the EWA lens barrel distortion was indeed compensated by the fisheye modelling. The profiles of the un-modelled errors varied markedly across models and cameras in terms of both general behaviour (barrel, pincushion or compound) and magnitude. In all cases except one, the magnitude of the distortion was reduced by up to three orders of magnitude. For the Samsung Galaxy Note10+, which exhibited an overall lower magnitude of pinhole model distortion, the fisheye modelling actually increased the distortion but without any detriment. For all datasets and fisheye models, the number of terms required to model the residual radial lens distortion was reduced from the number required for the pinhole model. This was even true for the Samsung.

Little change was observed in terms of functional dependencies between model variables. The fewer number of lens distortion terms required for the fisheye models did result in a slight reduction of mean correlation coefficient magnitude. The ultimate test was the independent accuracy assessment. Model choice had no bearing on 3D reconstruction accuracy, which was at the level of 2 mm to 6 mm, depending on the camera. Thus, the reduction of distortion model complexity without sacrificing 3D reconstruction accuracy suggests that the use of a fisheye model is advantageous for the photogrammetric exploitation of EWA lens camera systems. It is hoped that the findings of this work can be used to guide developers who wish to implement their own software for modelling EWA systems and/or users who may have a choice of which model to use among existing software packages.

ACKNOWLEDGEMENTS

Funding for this research was provided by the Natural Sciences and Engineering Research Council of Canada (CRDPJ 505367-16 and RGPIN/03775-2018), McElhanney Ltd, the Australian Research Council and the Curtin Institute for Computation. The authors thank Mozhdeh Shahbazi for use of the Calgary GoPro cameras.

REFERENCES

BALLARIN, M., BALLETTI, C., GUERRA, F., 2015. Action cameras and low-cost aerial vehicles in archaeology, in: Remondino, F., Shortis, M.R. (Eds.), *SPIE Optical Metrology*, Munich, Germany, p. 952813.

BALLETTI, C., GUERRA, F., TSIOUKAS, V., VERNIER, P., 2014. Calibration of action cameras for photogrammetric purposes. *Sensors* 14, 17471–17490.

BERNARDINA, G.R.D., CERVERI, P., BARROS, R.M.L., MARINS, J.C.B., SILVATTI, A.P., 2017. In-air versus underwater comparison of 3D reconstruction accuracy using action sport cameras. *Journal of Biomechanics* 51, 77–82.

BOSCH, J., GRACIAS, N., RIDAO, P. and RIBAS D., 2015 Omnidirectional Underwater Camera Design and Calibration. *Sensors*, 15: 6033-6065

CAMPOS, M.B, TOMMASELLI, A.M.G., MARCATO JUNIOR, J. and HONKAVAARA, E., 2018. Geometric model and assessment of a dual-fisheye imaging system. *Photogrammetric Record,* 33(162): 245-263.

CHOI, K., KIM, Y., KIM, C., 2019. Analysis of fish-eye lens camera self-calibration. *Sensors* 19, 1218.

FRASER, C.S., 1997. Digital camera self-calibration. *ISPRS Journal of Photogrammetry and Remote Sensing*. 52, 149-159.

HASTEDT, H., EKKEL, T., LUHMANN, T., 2016. Evaluation of the quality of action cameras with wide-angle lenses in UAV photogrammetry. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. XLI-B1: 851–859.

HELMHOLZ, P., LICHTI, D.D., 2020. Investigation of chromatic aberration and its influence on the processing of underwater imagery. *Remote Sensing* 12, 3002.

HUGHES, C., DENNY, P., JONES, E., GLAVIN, M., 2010. Accuracy of fish-eye lens models. *Applied Optics*. 49, 3338

JARRON, D., LICHTI, D., SHAHBAZI, M., and RADOVANOVIC, R., 2019a. Multi-Camera Panoramic Imaging System Calibration. In: Proceedings of the *11th International Conference on Mobile Mapping Technology*. Shenzhen, China, 6-8 May. 70-77.

JARRON, D., SHAHBAZI, M., LICHTI, D., and RADOVANOVIC, R., 2019b. Automatic detection and labelling of photogrammetric control points in a calibration test field. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W13: 1673-1680.

JARRON, D., SHAHBAZI, M., LICHTI, D., and RADOVANOVIC, R., 2019c. Modelling wide-angle lens cameras for metrology and mapping applications. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W7: 79–86.

KHORAMSHAHI, E., and HONKAVAARA, E., 2018. Modelling and automated calibration of a general multi-projective camera. *Photogrammetric Record,* 33(161): 86-112.

LICHTI, D.D., GLENNIE, C.L., JAHRAUS, A. and HARTZELL, P., 2019a. New approach for low-cost TLS target measurement. *ASCE Journal of Surveying Engineering,* 145(3): 04019008-1 – 04019008-12.

LICHTI, D.D., GLENNIE, C.L., AL-DURGHAM, K., JAHRAUS, A. and STEWARD, J., 2019b. Explanation for the seam line discontinuity in terrestrial laser scanner point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*. 154, 59-69.

LICHTI, D.D., JARRON, D., TREDOUX, W., SHAHBAZI, M., RADOVANOVIC, R., 2020. Geometric modelling and calibration of a spherical camera imaging system. *Photogrammetric Record* 35(170): 123–142.

MAALEK, R., LICHTI, D.D., 2021. New confocal hyperbola-based ellipse fitting with applications to estimating parameters of mechanical pipes from point clouds. *Pattern Recognition* 116, 107948

MIYAMOTO, K., 1964. Fish Eye Lens. *Journal of the Optical Society of America,* 54, 1060–1061.

PERNECHELE, C., 2016. Hyper hemispheric lens. *Optics Express* 24, 5014.

RAU, J.Y., SU, B.W., HSIAO, K.W. and JHAN, J.P., 2016. Systematic calibration for a backpacked spherical photogrammetry imaging system. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B1: 695–702.

RAY, S.F., 1994. *Applied Photographic Optics*, 2nd ed. Focal Press, Oxford.

SCHNEIDER, D., SCHWALBE, E., MAAS, H.-G., 2009. Validation of geometric models for fisheye lenses. *ISPRS Journal of Photogrammetry and Remote Sensing.* 64, 259–266.

TEO, T., 2015. Video-based point cloud generation using multiple action cameras. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XL-4/W5, 55–60.

WIERZBICKI, D., 2018a. Multi-camera imaging system for UAV photogrammetry. *Sensors* 18, 2433.

WIERZBICKI, D., 2018b. Calibration Low-Cost Cameras with Wide-Angle Lenses for Measurements. Journal *of Automation, Mobile Robotics and Intelligent Systems.* 12, 45–51

Résumé

L'histoire de l'appariement d'images remonte à plus de cinquante ans, lorsque les premières …

Zusammenfassung

Die digitale Bildzuordnung hat seit den ersten analogen Ansätzen für die automatisierte ...

Resumen

La correspondencia de imágenes tiene una historia de más de 50 años, desde los primeros …

摘要

影像匹配技术在模拟摄影测量中首次应用开始，已经有50年的发展 …