The effects of lens distortion calibration patterns on the accuracy of monocular 3D measurements

Jason de Villiers
Council for Scientific and Industrial
Research, Pretoria, South Africa
Email: jdvilliers@csir.co.za

Fintan Wilson
Council for Scientific and Industrial
Research, Pretoria, South Africa

Fred Nicolls
Department of Electrical Engineering
University of Cape Town
Cape Town, South Africa

Abstract—A variety of lens distortion modelling techniques exist. Since they make use of different calibration metrics it is difficult to select one over the others. This work aims to compare lens distortion modelling techniques and calibration patterns in a unified and objective manner. A common image dataset is captured along with physical measurements and is used to quantify the photogrammetric accuracy of the different calibration techniques. Multiple calibration patterns and sizes are tested and compared to results obtained with industry standard calibration methods. Several sub-pixel accurate methods of finding calibration points in images are evaluated. Improvements of 20% over the method used in OpenCV are consistently obtained. This work opens up the possibility for improved distortion characterisation in the scientific community.

I. Introduction

There exists a number of lens distortion characterisation methods, each of which can make use of a number of calibration patterns. Each distortion characterisation technique makes use of a different metric and this hinders objective comparisons and selection between the different techniques. The aim of this work is to determine the calibration pattern which yields the best distortion characterisation in terms of 3 dimensional (3D) measurements when using a single camera. To put the results of this work into context, precision techniques are compared to popular methods for lens distortion characterisation.

A. Lens distortion characterisation techniques

The purpose of lens distortion characterisation is to ensure that straight lines in the real-world project into straight lines in image space, shown by figures 2(b) and 3(b) versus figures 2(a) and 3(a) respectively. The majority of the techniques are based on the plumb-line approach, first described by Brown in 1971 [1]. This involves the numerical refinement of a chosen subset of two infinite series (for radial and tangential distortion respectively) as described in the Brown/Conrady model [2], [3] given in eq. 1.

$$h_u = h_d + (h_d - h_c)(K_1 r^2 + K_2 r^4 + \ldots) + \tag{1}$$
$$(1 + P_3 r^2 + \ldots) \times$$
$$\Big(P_1(r^2 + 2(h_d - h_c)^2) + 2P_2(h_d - h_c)(v_d - v_c) \Big)$$

$$v_u = v_d + (v_d - v_c)(K_1 r^2 + K_2 r^4 + \ldots) +$$
$$(1 + P_3 r^2 + \ldots)) \times$$
$$\Big((2P_1(h_d - h_c)(v_d - v_c) + P_2(r^2 + 2(v_d - v_c)^2)) \Big)$$

where:

(h_u, v_u) = undistorted image point,

(h_d, v_d) = distorted image point,

(h_c, v_c) = centre of distortion,

K_n = N$^\text{th}$ radial distortion coefficient,

P_n = N$^\text{th}$ tangential distortion coefficient,

$r = \sqrt{(h_d - h_c)^2 + (v_d - v_c)^2}$, and

$\ldots$ = an infinite series.

B. Calibration patterns

The purpose of calibration patterns is to provide measurable points which are known to be collinear in the object space. The accuracy with which these points can be found in the image directly affects the accuracy of the lens distortion characterisation. To this end, it is necessary to obtain these points with a level of accuracy higher than the discrete pixel sampling of the image plane. This work only considers calibration patterns whose reference points can be determined with sub-pixel accuracy. Checker boards are an extremely popular choice (e.g. the open computer vision (OpenCV) library [4], Caltech Camera Calibration Toolbox [5]) as the intersections can be found extremely accurately by finding the saddle point of the intensity profile about the intersection as described by Lucchese and Mira [6] and expanded upon by Chen and Zhang in [7]. Circles are also a popular choice since the centres can be found, with high accuracy, by determining the centroid or fitting an ellipse. A less conventional method, used by Brown [1], makes use of straight line grids. This method can provide a significant increase in the number of points used in the lens distortion characterisation as many more points can be found on each line.

C. Axis and notation definition

The mathematical notation used in this paper is as follows: A 3D vector, V_{abc}, is a vector from point a directed towards point b expressed in terms of its projections on orthogonal coordinate system c's axes. V_{abc} is used when the magnitude
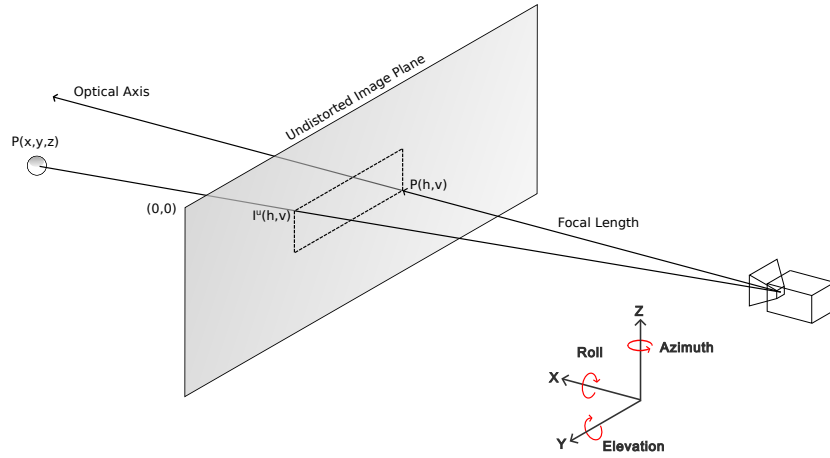
Fig. 1.   Axis definition.

of the vector is unknown or unimportant. T_{abc} represents the translation or displacement of point b relative to point a. R_{ab} is a 3-by-3 Euler rotation matrix expressing the rotation of an orthogonal axis system a relative to (and in terms of) its projections on) an orthogonal axis system b. Individual elements of 3 dimensional vectors are referred to as x, y or z whereas 2 dimensional (2D) vector's elements are referred to as horizontal (h) and vertical (v) to avoid confusion. Figure 1 defines the axis system used and the directions of positive rotation.

D. Paper organisation

The rest of this paper is organised as follows: section II describes the data capture methods for distortion characterisation and comparison. Section III details the methods used to compare the distortion characterisations in an unbiased manner. Thereafter, section IV provides a summary and discussion on the results obtained for each of the distortion characterisation methods. Finally, section V places the results of this work in context.

II. DISTORTION CHARACTERISATION METHODS

This section describes the camera equipment and methods used to capture and process data in order to calibrate it.

A. Equipment specification

A 1600-by-1200 Prosilica GE1660 Gigabit Ethernet machine vision camera was mated with a Schneider Cinegon 4.8mm/f1.4 lens for use in this work. This lens has an $82°$ horizontal field of view (FOV) with significant lens distortion and high modulation transfer function (MTF) making it particularly suitable for this work. The framework for live image transformation (flitr) [8] was used for the image capture and processing.

B. Distortion Calibration Patterns

A 46" Liquid Crystal Display (LCD) screen was used to display calibration patterns, and assumed to be sufficiently planar. This allowed for multiple variants of the calibration patterns to be tested. By shifting the calibration patterns a few pixels between captures, approximately 805 calibration points were captured. The camera was statically placed approximately perpendicular to the LCD such that the entire vertical FOV of the camera was occupied by the full extent of the LCD. This meant that there were two blind spots at the horizontal extremes of the LCD as the camera had 4:3 aspect ratio compared to the LCD's 16:9 ratio. In order to aid in the removal of background noise and ambient lighting effects, the calibration pattern capture was interleaved with capturing images of a blank LCD. Three different calibration patterns were considered, namely horizontal and vertical straight lines, checker intersections and circle arrays. Multiple sizes of the latter two patterns were captured in order to evaluate the effect of the calibration pattern size on calibration accuracy.

C. Pre-processing

This section details the distillation of raw captured image data into a set of accurate pixel positions for each of the reference marks for each calibration pattern.

Since a single line was captured at a time, simple thresholding of the background subtracted image yielded a dense list of all camera pixels on each line. Due to non-alignment of camera pixels and the non-perfect focus and non-zero line width of the line on the charge coupled device (CCD), the line in the captured image was several pixels wide. This meant that several hundred thousand calibration points were captured for the line patterns.

Two sub-pixel accurate methods for determining the intersection of checkers were evaluated, these being the window surface fitting method of Lucchese and Mira [6] and the Hessian matrix-based method proposed by Chen and Zhang [7]. Since the LCD coordinates of each checker are known in addition to the determined image coordinates of each checker it is possible to obtain many dozens of checkers for each row and each column as captured checkers in subsequent LCD frames can overlap physically.

The centre of the circle is the desired calibration image point for circle arrays. Two methods to find this point were

(a) Example distorted image. (b) Image undistorted as per [9].

Fig. 2.   Planar Reference Pattern, with Reference points marked.

evaluated, with the first being the centroid of the background subtracted circle image. The second is the numerical fitting of an ellipse to the determined set of pixels constituting the circle by minimising equation 2 (Leapfrog [10] was used for the minimisation):

$$metric = c_0 \pi ab + c_1(CS - ES) + c_2(WS - 2ES) \quad (2)$$

where

c_n = the nth weighting term,

CS = sum of intensities from centroid calculation,

$WS = \displaystyle\sum_{h,v \in W} I(h,v),$

$I(h,v)$ = image intensity at 2D coordinate (h,v),

$ES = \displaystyle\sum_{h,v \in W} \begin{cases} I(h,v) & \text{if } CR \leqslant ER, \\ \alpha I(h,v) & \text{if } ER < CR \leqslant ER+1, \\ 0 & \text{if } CR > ER+1, \end{cases}$

$CR = \| <h,v> - <E_h, E_v> \|$

$ER = \sqrt{\dfrac{a^2 b^2}{(b\cos(\theta))^2 + (a\sin(\theta))^2}}$

$\alpha = (1 - (CR - ER))$

$W = h \in (E_h - (a+3), E_h + (a+3)),$
$\quad v \in (E_v - (a+3), E_v + (a+3))$

(E_h, E_v) = centre of ellipse,

a = major axis of ellipse,

b = minor axis of ellipse,

θ = angle of major axis from horizontal.
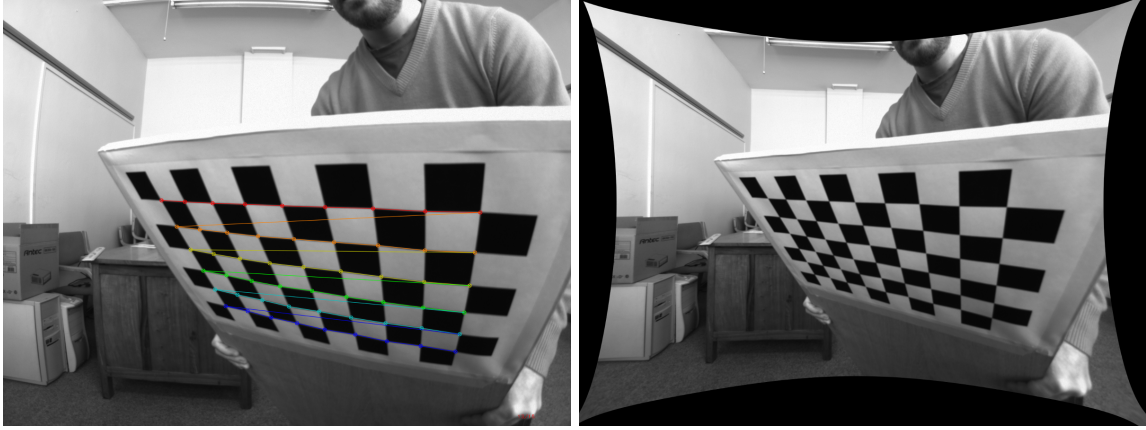
D. Precision distortion characterisation

The distortion characterisation method used in this work is that described by de Villiers *et. al* [11]. As suggested the Leapfrog algorithm [10] was used, as it is robust to errors and finds "low-local" minima and not merely the closest local minimum. With reference to eq. 1, five radial and three tangential parameters were determined as well as the optimal distortion centre. A genetic algorithm, implementing elitism and some "hill climbing" was run for 300 generations of 300 individuals to provide a robust starting point for Leapfrog to numerically refine. Thereafter all the points that had an error of more than three standard deviations from the mean were removed, and the Leapfrog algorithm was run again to further refine the distortion parameters.

The distortion measure that was minimised is that given in [11], namely the Root Mean Square (RMS) perpendicular distance of all the points on each row/column (of calibration reference points) from the best-fit straight line through the points that exist as part of that row/column. The calibration patterns and the determination of their reference points is described is sections II-B and II-C respectively.

E. OpenCV calibration

To better place this work in context the camera was also calibrated using the popular OpenCV library [4]. This calibration was performed by capturing 15 images of a checker board which had a total of 54 checker intersections, for a total of 810 intersections. The 15 images were chosen such that a subset of them covered the entire FOV of the camera with the checkerboard approximately orthogonal to the image axis and upright. The remainder of the dataset contained images of the checkerboard at non-orthogonal positions. Figure 3(a) contains one of the non-orthogonal images. As can be seen from figure 3(b), which contains the image as undistorted by OpenCV, the characterisation was successful. An average re-projection error of 0.770 pixels as well as the two radial and two tangential coefficients together with the distortion centre (see eq. 1) and focal length were returned to completely

(a) Example OpenCV calibration image.

(b) OpenCV undistorted image.

Fig. 3.   OpenCV [4] calibration.

characterise the distortion.

III. COMPARISON OF DISTORTION CHARACTERISTICS

A suitable metric is required in order to objectively compare the use of different calibration patterns and the resulting lens distortion characterisations. It was decided to compare a physical measurement, in the real-world, to an estimated photogrammetric measurement based on the distortion calibrations. To facilitate a monocular camera measurement, discernible points that exist in a 2D plane were required.

Fig. 4.   Calculated camera positions relative to planar reference.

A common dataset was gathered, consisting of nine images of a planar reference target as observed from different directions. Figure 4 shows the calculated position for the camera relative to the planar reference for each calibration/image pair, 5 DOF are shown as roll is not indicated. The end of the line indicates the translation of the camera and the line represents the optical axis extended to intersect with the planar reference. The reference points in each image were manually located. Distinct points were chosen such that the locations could be determined with a high level of accuracy. Figure 2(a) provides a sample of this dataset and figure 2(b) highlights the 6 points

used and also shows the result of the custom lens distortion correction. These 6 points allow for a total of 15 unique pairs with corresponding physically measured distances. Four additional reference checker intersections, whose relative positions are known, were placed on the planar reference to facilitate the pose estimation of the camera, as required to perform the monocular measurements. The subsequent sections describe the mathematical mechanics involved in the metric.

A. Pose estimation

Pose estimation was performed by minimising the offset between two bundles of vectors. The first vector bundle is created from the image data by transforming the pixel positions of the selected points in the planar reference. This is done using the distortion parameters and the focal length of the lens. The second vector bundle is created by hypothesising the 6 degree of freedom (DOF) position of the planar reference relative to the camera and calculating the vectors to the selected points. This method is sensitive to both translation and rotation as detailed by de Villiers [11]. A genetic algorithm, with 1000 generations of 1000 individuals, was used to generate an initial 6 DOF position for the Leapfrog algorithm to further refine.

B. Planar projection

The pose estimation information was used to project the location of the pixel position of the selected point in the image plane into the plane of the planar reference. This process is described mathematically as:

$$T_{rpr} = V_{cpr} \frac{T_{crr_x}}{V_{cpr_x}} + T_{crr} \qquad (3)$$

where

$$T_{crr} = R_{rc}^T T_{crc},$$
$$V_{cpr} = R_{rc}^T V_{cpc},$$

$$V_{cpc} = \begin{bmatrix} Focal\_Len \\ (P_h - I_{i_h}^u)pix\_w \\ (P_v - I_{i_v}^u)pix\_h \end{bmatrix},$$

$(P_h, P_v) =$ the principal point,

$(I_{i_h}^u, I_{i_v}^u) =$ the undistorted pixel position of planar

reference point i,

$pix\_w =$ the width of the pixels on the camera's CCD,

$pix\_h =$ the height of the pixels on the camera's CCD,

$R_{rc} =$ rotation of the planar reference to the camera,

$T_{crc} =$ the spatial offset of the planar reference origin

relative to the camera, and

$T_{rpr} =$ the distance from the planar reference origin to

the point.

C. Comparison metric

Equation 4 provides the comparison metric. It is the RMS error over all possible pairs of planar reference points of the difference between the physically measured distance and the corresponding distance as calculated from the projected image points (Eq. 3):

$$metric = \sqrt{\frac{2}{n^2 - n} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(\|T_{rp_i r} - T_{rp_j r}\|^2 - D_{i,j} \right)}$$

(4)

where

$n =$ the number of planar reference points,

$D_{i,j} =$ the measured distance between points i and j,

$T_{rp_i r} =$ projection of point i as per Eq. 3, and

$T_{rp_j r} =$ projection of point j as per Eq. 3.

IV. RESULTS

This section provides the results of the experiments and discusses the results.

Table I provides the results of the calibrations performed for the various lens distortion calibration patterns discussed in section II. Both the initial error and the resultant error after characterisation are given for comparison. The error values (in units of pixels RMS) are described in section II-D and [9]. The circle array patterns have two sets of values. The first is the characterisation when using the centroid of the captured circle, and the second when using the centre of the fitted ellipse (Eq. 2). The square patterns also have two sets of values: "L&M" refers to results obtained using [6] whereas "C&Z" indicates the results for checker intersections found by using [7].

Table II provides the 3D projection accuracy of each calibration method. This metric is described in detail in section III, and compares the photogrammetric measurements between the manually selected image points of the planar reference (see figure 2(b)) to their physically measured displacements. The RMS error for each of the 9 images taken from the planar

[1]OpenCV average reprojection error, not the metric described in section II-D

TABLE I
DISTORTION METRICS

Pattern type	Measurement method	Initial distortion (pixels RMS)	Optimal distortion (pixels RMS)
Open CV Calibration[1]		-	0.770
Circle, size 10	Centroid	347.645	0.081
	Ellipse	347.785	0.088
Circle, size 25	Centroid	335.622	0.078
	Ellipse	335.510	0.142
Square, size 15	L&M	386.059	0.256
	C&Z	340.954	0.082
Square, size 25	L&M	386.342	0.103
	C&Z	314.710	0.081
Square, size 50	L&M	386.695	0.099
	C&Z	251.682	0.060
Lines		290.917	3.026

reference is provided. The RMS error calculated over the first 5 images as well as all 9 images is provided for comparison. The interpretation of the two leftmost columns is the same as for table I.

The initial distortion values in table I vary slightly. This is due both to the accuracy with which the patterns are captured and inherent characteristics of the capturing process. The circle capture methods discard patterns that touch the image edges, ergo more of the smaller circles at the image periphery (which has higher distortion) were captured. Similarly, the implementation of Chen's checker intersection finding method [7] did not find intersections as close to the periphery as did the Lucchese [6] implementation.

In table II there is a clear increase in error moving from left to right. This can be attributed to increased obliqueness with which the planar logo was viewed. Furthermore, the limited depth of focus decreased the accuracy with which the manually tagged image points were located. This resulted in less accurate pose estimations and photogrammetric measurements.

The results indicate that a 25 pixel checker pattern located with the method described in [7] produced the best results. This method performs consistently well over the range of checker sizes. The same cannot be said for the method described in [6]. This method relies on fitting a curve to the intensity profile about the checker intersection. Using a larger checker ensures that the window of pixels used for the curve fitting will never extend beyond the bounds of the checker. This is evident in table I where a better distortion characterisation is achieved when a larger checker size is used. Furthermore, this trend propagates through to the results presented in table II. The 50 pixel checker outperforms the smaller checkers when considering the method of [6].

The circular calibration pattern results indicate that using the ellipse fitting method with a larger circle produces the best results. The reduction in the data obtained with smaller circles hampers the ability of the ellipse fitting process to accurately determine the centre of the ellipse in the image. As such, there exists no discernible difference in accuracy between the ellipse and centroid methods when smaller circles are used. With larger circles, however, the true centre of the circle will

TABLE II
3D ERROR OF CALIBRATION PATTERS

Pattern type	Measurement method	Error (mm) for image									Images 1-5 error (mm)	Global error (mm)
		1	2	3	4	5	6	7	8	9		
Open CV Calibration		3.48	3.91	3.19	2.67	3.36	4.92	5.19	3.82	8.17	3.66	4.58
Circle, size 10	Centroid	2.40	3.17	3.58	3.59	3.87	3.35	7.00	3.63	5.99	3.36	4.29
	Ellipse	2.44	3.26	3.84	3.38	3.20	3.52	7.03	4.30	5.63	3.26	4.28
Circle, size 25	Centroid	2.66	3.58	4.15	2.81	3.12	3.42	8.95	4.72	6.45	3.31	4.84
	Ellipse	2.38	3.40	3.10	2.76	3.33	2.94	5.09	3.94	6.71	3.02	3.95
Square, size 15	L&M	2.59	3.32	2.64	2.41	3.13	3.22	5.97	4.86	6.49	2.84	4.11
	C&Z	2.37	3.26	3.09	3.23	3.57	3.19	4.95	3.70	6.72	3.13	3.98
Square, size 25	L&M	2.51	3.47	3.41	2.71	3.01	3.45	4.97	3.85	6.31	3.05	3.91
	C&Z	2.46	3.34	3.02	3.00	3.27	3.23	4.73	3.73	6.28	3.03	3.83
Square, size 50	L&M	2.62	3.42	3.49	2.71	3.07	3.22	5.61	3.80	6.42	3.08	4.01
	C&Z	2.40	3.34	3.04	3.19	3.47	3.18	4.97	3.62	6.27	3.11	3.88
Lines		3.79	5.06	5.55	2.54	4.31	3.31	5.47	9.31	7.02	4.38	5.5

not coincide with the centre of the ellipse due to the non-linear radial compression introduced by the lens distortion as a function of the distance from the principal point. Additionally, due to the non-perfect orthogonality of the camera to the LCD, the circle pattern will be projected onto the CCD resulting in an ellipse. This results in a perspective induced difference between the projected centre of the circle, and the projection of the centre of the ellipse. These two effects therefore place an upper limit on the circle size.

The line calibration patterns produced the worst results overall. Coupled with the significant increase in computational requirements (due to the increased number of captured pixels) during the distortion characterisation phase, this makes the method the least desirable. An initial inspection of table I may lead one to believe that this was apparent. The poor distortion result of the line calibration in table I is due to the line width being greater than a single pixel on the CCD, as mentioned in section II-C.

A direct comparison between the results of table I is not meaningful as it is not a direct indication of the photogrammetric accuracy of the camera. This is evident in the comparatively poor distortion characterisation of the 25 pixel circles using ellipse fitting. This characterisation performed the best overall for the first five images (as shown in table II) and performed well when all the images were considered, contrary to the results presented in table I. Similarly, the calibration results would lead one to believe that characterisation of checkers size 15 pixels using Lucchese and Mira's method [6] would be 3 times worse than a majority of the other calibrations. This is not the case, as seen in table II. The same is true for the calibration using lines.

The OpenCV results are also presented in table II. They compare less favourably to the results obtained with the larger calibration patterns. When only the first 5 images are considered, OpenCV performs 29% worse than the best method. This drops to 20% when all nine images are considered. This comparison is pertinent as the data used for the OpenCV calibration (an example of which is shown in figure 3(a)), contains checkers of similar image size to the 50 pixel checker. Both OpenCV and the other distortion characterisations determined an optimal focal length for the lens, this value differed slightly from the manufacturer's claimed value.

V. Conclusions

The distortion metric is not indicative of photogrammetric accuracy.

Calibration with checker patterns yields better accuracy than using circular or line patterns. The method proposed by Chen and Zhang [7] is more robust to different checker sizes which leads to more consistent and accurate distortion characterisations when compared to the Lucchese and Mira [6] approach. Ellipse fitting provides better distortion characterisations than using the centroid when larger circles are used.

Accuracies of 20% improvement over the popular OpenCV method were obtained, with all but two calibration patterns leading to results superior to those of OpenCV.

References

[1] D. Brown, "Close range camera calibration," *Photogrammetric Engineering*, vol. 8, pp. 855–855, 1971.

[2] ——, "Decentering distortion of lenses," *Photogrammetric Engineering*, vol. 7, pp. 444–462, 1966.

[3] A. Conrady, "Decentered lens systems," *Monthly Notices of the Royal Astronomical Society*, vol. 79, pp. 384–390, 1919.

[4] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[5] J.-Y. Bouget, "Camera calibration toolbok for matlab," 2010. [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/

[6] L. Lucchese and S. Mira, "Using saddle points for subpixel feature detection in camera calibration targets," in *Proceedings of the Asia-Pacific Conference on Circuits and Systems*, vol. 2, 2002, pp. 191–195.

[7] D. Chen and G. Zhang, "A New Sub-Pixel Detector for X-Corners in Camera Calibration Targets," in *WSCG (Short Papers)*, 2005, pp. 97–100.

[8] J. Delport, "Framework for live image transformation," 2010. [Online]. Available: http://code.google.com/p/flitr/

[9] J. de Villiers, F. Leuschner, and R. Geldenhuys, "Centi-pixel accurate real-time inverse distortion correction," in *Proceedings of the 2008 International Symposium on Optomechatronic Technologies*, ser. ISOT2008, vol. 7266, 2008, pp. 1–8.

[10] J. Snyman, "An improved version of the original leap-frog dynamic method for unconstrained minimization: Lfop1(b)," *Applied Mathematics and Modelling*, vol. 7, pp. 216–218, 1983.

[11] J. de Villiers, "Real-time stitching of high resolution video on COTS hardware," in *Proceedings of the 2009 International Symposium on Optomechatronic Technologies*, ser. ISOT2009, vol. 9, 2009, pp. 46–51.