Simple fish-eye calibration method with accuracy evaluation

Ciarán Hughes*, Martin Glavin and Edward Jones

Connaught Automotive Research Group,
Electrical and Electronic Engineering, National University of Ireland, Galway, Ireland

Abstract

In this paper, a simple fish-eye radial distortion calibration procedure is described. This method avoids costly minimisation and optimisation algorithms, and is based on trivial concentricity of three extracted points. The results show that this simplicity is at the expense of increased deviation of results (and thus increased error). However, this deviation can be reduced significantly by the use of simple averaging, such that it is only marginally greater than the current state-of-the-art.

Key Words: Fisheye, Equidistant projection, Camera calibration, Distortion centre.

1   Introduction

Fish-eye cameras are so-named because they relate to the $180°$ view of the world that a fish has observing the water surface from below. Wood [1] originally coined the term, and constructed a fish-eye camera by taking a pin-hole camera and filling it with water. This water-based camera was replaced with a hemispherical lens by Bond [2], with further early work completed by Hill [3] and Beck [4]. Indeed, the first commercially available camera with fish-eye lens was called the "Robin Hill Camera" or "Hills Cloud Camera" [5]. Fish-eye cameras are designed such that a greater portion of a real world scene can be captured by an image sensor, resulting in cameras with a field-of-view of up to $180°$ (or even greater; the Fish-eye Nikkor lens has a FOV of $220°$ [6]). However, this greatly increased FOV comes at the expense of distortion being introduced into the resultant image.

Typically, the early use of fish-eye cameras was restricted to scientific measurement, such as in whole-sky photography for meteorological purposes [3, 4], forest canopy research [5, 7], medical imagery (endoscopy) [8, 9], or other scientific measurements [10, 11, 12], and these continue to be popular applications for fish-eye lenses. Wide-angle optics are also becoming more common in automotive applications [13], with many major automotive manufacturers including wide-angle cameras as features.

Fish-eye distortion causes points on the image plane to be displaced in a nonlinear fashion from their ideal position in the rectilinear pin-hole camera model, along a radial axis from the distortion centre in the fish-eye image plane. The visual effect of this displacement in fish-eye optics is that the image will have a higher

---

Correspondence to: <ciaran.hughes@valeo.com>
* Currently with Valeo Vision Systems, Tuam, Co. Galway, Ireland

resolution in the foveal areas, with the resolution decreasing nonlinearly towards the peripheral areas of the image. Figure 1 shows the difference between a standard FOV camera and a fish-eye camera. Both cameras are positioned at the same distance from the target, which consists of equally spaced bands of black and white. While it can be clearly seen that the fish-eye camera images a far greater portion of the scene than the standard FOV camera, severe radial distortion is also introduced that results in the straight lines becoming curved (very mild radial distortion can also be observed in the standard FOV image, though this is negligible compared to the level of distortion in the fish-eye image).
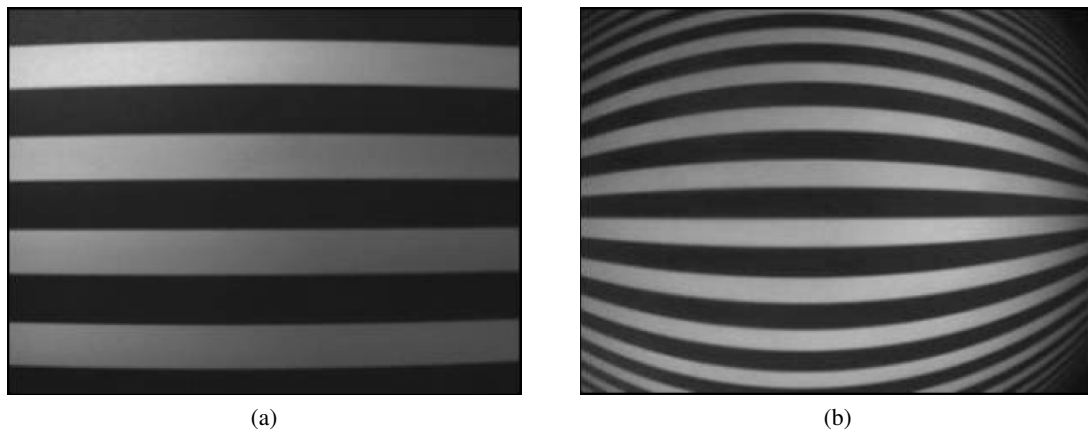

(a)   (b)

Figure 1: Difference between (a) a standard field-of-view camera (b) and a fish-eye camera which images a far greater portion of the scene.

The aim of this paper is to introduce and describe in detail a simple fish-eye calibration method. The proposed method does not include any minimisation or optimisation functions, which has several advantages, as it avoids issues such of potential non-convergence, development of nonlinear iterative optimisation methods and local minima. However, this is at the cost of somewhat decreased accuracy of parameter estimation. The calibration procedure itself is based on the trivial concentricity of three points on a distorted line extracted from a fish-eye image for the extraction of the distortion centre. The equidistant parameter that returns the minimum area of the triangle that is described by these three points is found using a simple search algorithm. We compare our results with those presented in the state-of-the-art literature [14, 15, 16].

1.1 Projection functions

The rectilinear pin-hole camera model is typically considered to be the ideal and intuitive model, whereby straight lines in the real-world are mapped to straight lines in the image. The rectilinear projection function is given as [17]:

$$r_u = f \tan (\theta) \tag{1}$$

where r_u is the height of the projection on the image plane (the subscript u being used to denote the undistorted projection), f is the equidistant model parameter of the camera and θ is the angle the projected ray makes with the optical axis (Figure 2a).

For normal and narrow FOV cameras, the effects of radial distortion can be considered negligible for most applications. However, in wide-angle and fish-eye cameras, radial distortion can cause severe problems, not only visually but for further processing in computer vision applications such as object detection, recognition and classification. Additionally, the radial distortion introduced by these lenses does not preserve the rectilinearity of an object in its transformation from real-world coordinates to the image plane. The visual effect of this is that straight lines in the real world are distorted to curves in the image plane.

Fish-eye cameras are typically designed to follow one of several spherical projection functions. The projection functions associated with fish-eye cameras are the equidistance, the equisolid, the orthogonal (ortho-

graphic) and the stereographic projection functions [17, 18]. The equidistant projection function is the most common projection function that fish-eye lenses are designed to follow, and for this reason the equidistance projection function is used in this paper. Equidistant fish-eye lenses are often used for scientific measurement where the measurement of angles is necessary. Thus, it is also sometimes referred to as an equiangular fish-eye lens. According to Miyamoto [18], the equidistant lens is "preferable for measurement of zenith angles and azimuth angles. The effect of error of lens position is small, and the linear relation of r_d and θ is convenient to analyse". That is, the angle of the projected ray in radians translates linearly to the distorted radial distance on the image plane. Additionally, the projected distorted distance r_d is equivalent to the length of the arc segment between the z-axis and $\mathbf{x_p}$, which is the point of intersection of the projection ray of point $\mathbf{X}$ with the projection sphere (Figure 2b). Thus, the projection function is:

$$r_d = f\theta \tag{2}$$

where the subscript d is used to denote a radial distance in distorted fish-eye image space. To determine the distortion function, i.e. the function that converts a fish-eye point to its rectilinear equivalent, (1) and (2) are solved in terms of θ and equated to get:

$$r_d = f \arctan\left(\frac{r_u}{f}\right) \tag{3}$$

and the inverse is:

$$r_u = f \tan\left(\frac{r_d}{f}\right) \tag{4}$$

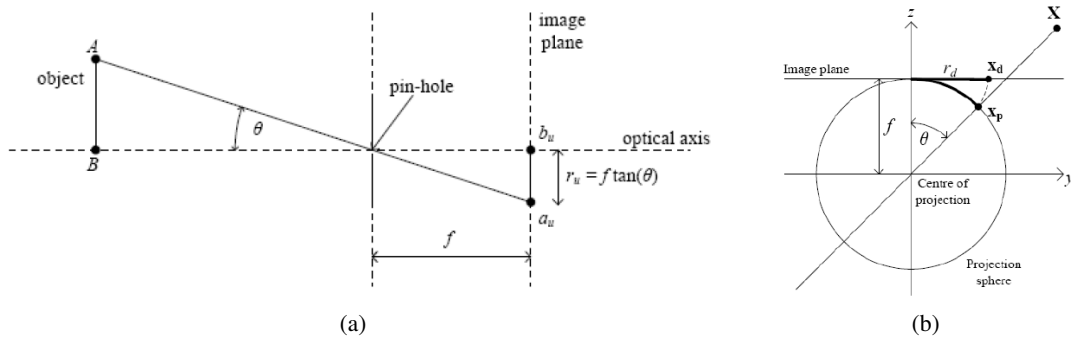These equations describe the conversion between rectilinear image space and the fish-eye image space, and vice-versa.

Figure 2: Projection functions; (a) Rectilinear projection; and (b) Spherical equidistant projection.

2   Description of the method

2.1   Extracting three points on each curve

The first step in this calibration procedure is to capture an image of a checkerboard calibration diagram using the fish-eye camera. For each imaged line in the calibration diagram, the aim is to extract three points, as shown in Figure 3a. As the distortion centre estimation is based on the trivial fitting of a circle to these three points, three points are selected that are as distant as possible from each other along each imaged line. This is due to the fact that the sensitivity of the circle fit decreases as the angle the points make with the circle centre increases [19]. The extraction of the points on the curves can be done manually. However, here an automatic procedure to extract the points is preferred. The Sobel edge detection operator is applied to the calibration image [20].

This need only be applied along lines towards the periphery and through the numerical image centre of the calibration image, as shown in Figure 3a. If $\mathbf{I}$ is the source image, the Sobel operator is:

$$\mathbf{G}_u = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{I} \tag{5}$$

$$\mathbf{G}_u = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * \mathbf{I} \tag{6}$$

where $*$ denotes convolution, and $\mathbf{G}_u$ and $\mathbf{G}_v$ are the resultant gradient images. The overall gradient magnitude is determined as $\mathbf{G} = \sqrt{\mathbf{G}_u^2 + \mathbf{G}_v^2}$. If the vertical gradient $\mathbf{G}_u$ is greater than the horizontal gradient $\mathbf{G}_v$, the point is determined to be on one of the horizontal lines. Conversely, if $\mathbf{G}_v > \mathbf{G}_u$, the point is determined to be on one of the vertical lines. The edge detection will be discontinuous at corners of the checkerboard. However, these discontinuities can be detected and corrected using a suitable corner detection algorithm (e.g. [21]).
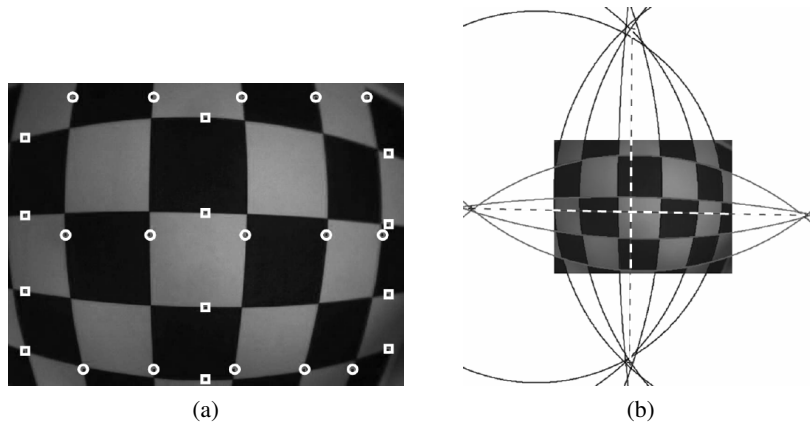

(a)   (b)

Figure 3: The distortion centre estimation algorithm; (a) Extraction of points; and (b) Circle fit to three points and distortion centre estimation.

Figure 3a shows the points extracted from each calibration edge. The circles show the points selected on the vertical edges, and the boxes show the ones selected on the horizontal edges. The points selected were either along a centre line of the image or near the edge of the image, though not at the edge of the image, to avoid issues with the Sobel filter at image edge pixels (as the kernel window passes outside the image boundary).

2.2   Estimating the Distortion Centre

The distortion centre (c_u, c_v) estimation is based on the same perspective principles that are introduced in [14] and described in greater detail [15]. There are three key points from [15]; (a) the projection of a set of lines in the scene to the equidistant image plane can be accurately modeled as a set of circles; (b) the circles from the projection of parallel scene lines will converge on two vanishing points; and (c) the line that intersects the two vanishing points also intersects the distortion centre.

However, in the interest of simplicity, circles are fit to the three points extracted from each image edge, instead of using nonlinear optimisation to determine the overall error of circles fit to each curved image line through two vanishing points, as is done in [14, 15]. For each circle fit, the intersection point with the other circles is determined, and the average of the intersection points is taken as the vanishing points. Figure 3b shows the circles fitted to the points extracted from each of the edges for a calibration image. Also shown are the vanishing points. The distortion centre is the point at which the lines between the vanishing points intersect [14].

2.3   Estimating the equidistance parameter

Each set of three points extracted from the calibration image edges form the vertices of a triangle (Figure 4a). When the three points are collinear, the area of this triangle is zero. The area A of a triangle is given by:

$$A = \frac{1}{2}\left| \det\left(\begin{bmatrix} u_{u,1} & u_{u,2} & u_{u,3} \\ v_{u,1} & v_{u,2} & v_{u,3} \\ 1 & 1 & 1 \end{bmatrix} \right) \right| \tag{7}$$

where $[u_{u,1}, v_{u,1}]$, $[u_{u,2}, v_{u,2}]$ and $[u_{u,3}, v_{u,3}]$ are the vertices of the triangles; i.e. they are the points that are extracted, as shown in Figure 4a, with the inverse equidistant distortion function 4 applied. However, given the relationship of similar triangles, $u_u/r_u = u_d/r_d$ and $v_u/r_u = v_d/r_d$, and using 4:

$$u_u = \frac{u_d}{r_d}r_u = \frac{u_d}{r_d}f\tan\left(\frac{r_d}{f}\right) \tag{8}$$

$$v_u = \frac{v_d}{r_d}r_u = \frac{v_d}{r_d}f\tan\left(\frac{r_d}{f}\right) \tag{9}$$

Therefore, the area A of the triangle becomes:

$$A = \frac{1}{2}\left| \det\left(\begin{bmatrix} \frac{u_{d,1}}{r_{d,1}}f\tan\left(\frac{r_{d,1}}{f}\right) & \frac{u_{d,2}}{r_{d,2}}f\tan\left(\frac{r_{d,2}}{f}\right) & \frac{u_{d,3}}{r_{d,3}}f\tan\left(\frac{r_{d,3}}{f}\right) \\ \frac{v_{d,1}}{r_{d,1}}f\tan\left(\frac{r_{d,1}}{f}\right) & \frac{v_{d,2}}{r_{d,2}}f\tan\left(\frac{r_{d,2}}{f}\right) & \frac{v_{d,3}}{r_{d,3}}f\tan\left(\frac{r_{d,3}}{f}\right) \\ 1 & 1 & 1 \end{bmatrix} \right) \right| \tag{10}$$

where $[[u_{d,1}, v_{d,1}]$, $[[u_{d,2}, v_{d,2}]$ and $[[u_{d,3}, v_{d,3}]$ are the vertices of the triangle that are extracted from the fish-eye image (Figure 4a), and $r_{d,1}$, $r_{d,2}$ and $r_{d,3}$ are the corresponding radial distances from the already determined distortion centre to each of the vertices. Given that the distorted vertices of each triangle and the radial distances of the vertices are known, and thus constant, as they have been extracted from the fish-eye image, it becomes clear that the area A of the undistorted triangle vertices is a function of the equidistant parameter f. Figure 4b shows the triangular areas for a range of values of f for a sample triangle from the calibration image shown in Figure 4b.
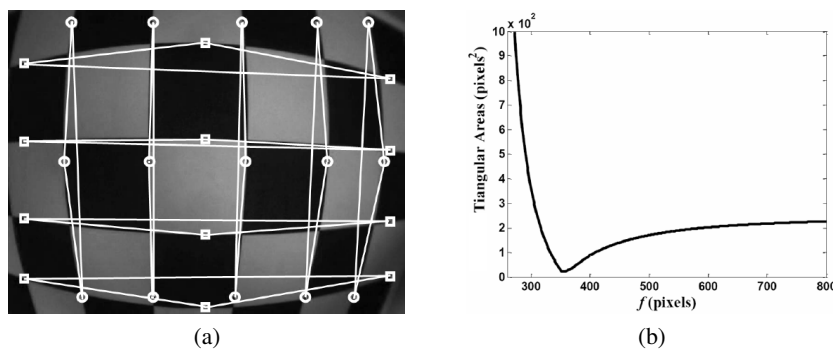

(a)   (b)

Figure 4: Estimation of f; (a) triangles; and (b) triangular areas for a range of value of f.

To find the value of f that returns the minimum triangle area, we utilise a simple golden section search [22]. However, this search requires an upper and a lower bound for the values of f. The lower bound is selected as the minimum possible value for f. The aim is to provide an undistorted image, and as such, ray incident angles of over $90°$ cannot be considered (due to the fact that $\tan(90°)$ is infinitely large). If a sensor pixel resolution of 640×480 is assumed, the maximum radial distance from the centre is 400 pixels. Therefore, from 4, the minimum value for f becomes:

$$f_{\min} = \frac{r_{d,\max}}{\theta_{d,\max}} = \frac{800}{\pi} = 254.65 \text{ pixels} \tag{11}$$

Thus, a value of 255 is used as a lower bound for f. Of course, this can be adjusted using the same methods for any sensor pixel resolution.

The upper bound cannot be derived as analytically as the lower bound. Theoretically, the f can be infinitely large. However, a value of f of 800 pixels is equivalent to a maximum incident ray $\theta_{\max}$ of:

$$\theta_{\max} = \frac{400}{800} \text{ radians} = 28.6° \tag{12}$$

The angular field-of-view of such a camera would be FOV $= 2\theta_{\max} = 57.2°$. This FOV is far lower than any fish-eye camera, and as such an upper bound of 800 pixels is selected (for a 640×480 pixel resolution). Figure 4b shows the area for the values of f between the bounds $(255, 800)$.

This search is repeated for all of the extracted triangles in the calibration image (Figure 4a), and the value of f that minimises the total area of all of the triangles is selected as the estimate of that parameter.

3   Results

In this section, the calibration results for three cameras (Table 1) are presented. Table 2 shows the mean and standard deviations of the fish-eye parameter estimations for 15 calibration images captured using each of the cameras, and Figure 5 shows the distribution of the estimations.

Table 1: List of lens types used in testing

	Camera type	Lens
Camera 1	Fish-eye	178° FOV $f\theta$
Camera 2	Fish-eye	178° FOV $f\theta$
Camera 3	Wide-angle	103° FOV $f\theta$

This method is considerably simpler than either the circle fitting or the conic fitting methods described in [14, 15], as it avoids the implementation of a nonlinear optimisation algorithm and requires the extraction and use of a very limited number of data points from the fish-eye calibration image, this is with the trade-off of increased sensitivity to noise (note that the results presented here are the exact cameras that the authors previously used in [15], and as such a direct comparison is possible). This is evident in the larger distributions of the results. For example, the standard deviation for the equidistant parameter f results from Camera 1 is 7.81 pixels, as opposed to 1.87 for the circle fit method and 1.13 for the conic fit method from [15].

As with the methods proposed in [14, 16] averaging can be used to overcome the sensitivity to noise. Figure 6 shows the distributions of the results after "5 of 15" and "10 of 15" averaging, with the expected reduction in standard deviations shown in Table 3. After the "10 of 15" averaging, it can be seen by comparing the results

Table 2: Mean and standard deviations of the intrinsic parameter estimation using the simple calibration method

	Camera 1		Camera 2		Camera 3	
	μ	ρ	μ	ρ	μ	ρ
f	273.86	7.81	277.73	10.81	365.53	10.94
c_u	253.35	5.26	206.62	8.21	252.93	6.05
c_v	277.45	10.28	300.59	6.01	283.93	8.14

The mean μ and standard deviation ρ are in pixels. The pixel resolution for all cameras is 640×480
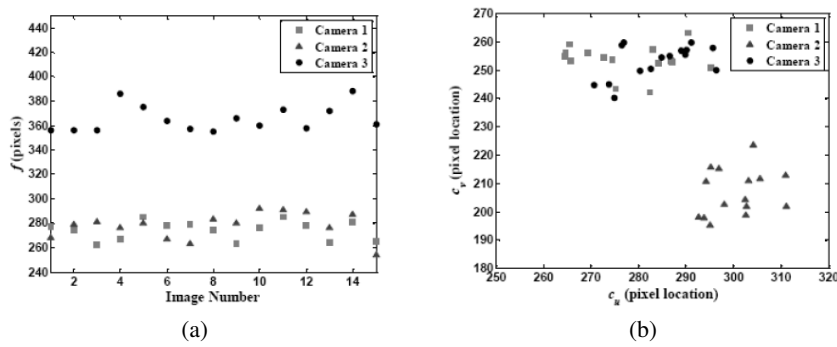
| (a) | (b) |

Figure 5: Distribution of the fish-eye parameter estimations; (a) equidistant parameter f; and (b) distortion centre (c_u, c_v).

Table 3: Standard deviations of the simple calibration estimations after averaging

	f	c_u	c_v
Averaging 5 of 15			
Camera 1	3.01	2.25	3.56
Camera 2	3.91	2.92	2.23
Camera 3	3.86	2.37	3.40
Averaging 10 of 15			
Camera 1	1.27	0.96	1.91
Camera 2	1.98	1.51	1.08
Camera 3	1.95	1.17	1.49

All measurements are in pixels

presented here with those presented in [15] that the standard deviations are comparable. Typically, the standard deviations of the results presented here after averaging are not more than 1 pixel greater than those presented by the single-pass conic fitting method in [15]. Figure 7 shows an example of a photographic image corrected using parameters estimated from the proposed calibration method.

We validated the assertion that the equidistant fish-eye model was an appropriate model for these cameras using the method described in [23].

4   Conclusions

In this paper we have presented a simple fish-eye calibration based on the capture of a calibration image from a standard checkerboard diagram. The proposed method is inexpensive, and avoids the use of iterative methods and nonlinear optimisations, and related issues, such as the potential for non-convergence and local minima, etc. The results presented show, however, that the increased simplicity of the algorithm is at the expense of a greater distribution (and thus error) of the estimated parameters using the proposed method, when compared the method described in [15]. However, this can be mitigated through the use of averaging over a number of calibrations for a given camera, which has also been utilised in recent calibration techniques [14, 16].
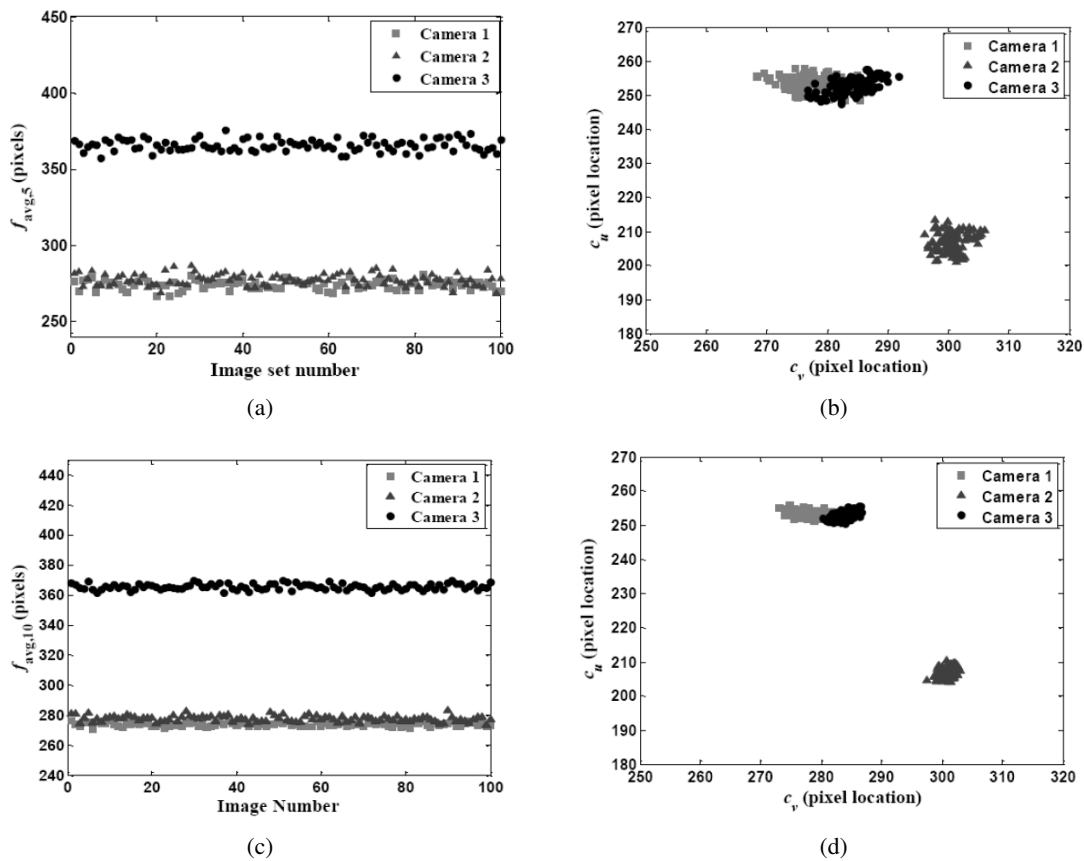
Acknowledgements

Figure 6: Results of "5 of 15" and "10 of 15" averaging; (a) "5 of 15" averaging of f; (b) "5 of 15" averaging of (c_u, c_v); (c) "10 of 15" averaging of f; and (d) "10 of 15" averaging of (c_u, c_v).

References

[1] R.W. Wood, "Fish-eye views, and vision under water", *Philosophical Magazine* 12(6):159–162, 1906.

[2] W.N. Bond, "A wide angle lens for cloud recording", *Philosophical Magazine* 44(263):999–1001, 1922.

[3] R. Hill, "A lens for whole sky photographs", *The Quarterly Journal of the Royal Meteorological Society* 50(211):227–235, 1924.

[4] C. Beck, "Apparatus to photograph the whole sky", *Journal of Scientific Instrumentation* 2:135–139, 1925.

[5] G.C. Evans, D.E. Coombe, "Hemisperical and woodland canopy photography and the light climate", *The Journal of Ecology* 47:(1)103–113, 1959.

[6] http://www.mir.com.my/rb/photography/companies/nikon/nikkoresources/fisheyes/6mmf28.htm, accessed November 2011.

[7] M.C. Anderson, "Studies of the woodland light climate: I. the photographic computation of light conditions", *The Journal of Ecology* 52(1):27–41, 1964.

[8] J. Brubaker, "Lens data tables and design calculations for open-tube endoscopic, cavity, and close-up photography", *Journal of the Biological Photographic Association* 20(2):81–85, 1952.

[9] A. Soulas, J.M. Dubois Dedenthon Dontreynaud, R.J. Edwards, A.J. Gladu, "Bronchoscopy and television", *Chest Journal* 31(5):580–584, 1957.

(a)   (b)

Figure 7: Example of corrected images; (a) original photographic example; and (b) corrected photographic example.

[10] R.M. Langer, "New form of matter found in cosmic ray bombardments", *The Science News-Letter* 31(839):291, 1937.

[11] J. Burger, "The use of a fish-eye lens to study nest placement in Franklins gulls", *Ecology* 53(2):362–364, 1972.

[12] G.E. Walsberg, "Quantifying radiative heat gain in animals", *American Zoologist* 32(2):217–224, 1992.

[13] C. Hughes, M. Glavin, E. Jones, P. Denny, "Wide-angle camera technology for automotive applications: a review", *IET Intelligent Transportation Systems* 3(1):19-31, 2009.

[14] C. Hughes, R. McFeely, P. Denny, M. Glavin, E. Jones,"Equidistant fish-eye perspective with application in distortion centre estimation", *Image and Vision Computing* 28(3):538-551, 2010.

[15] C. Hughes, P. Denny, M. Glavin, E. Jones, "Fish-eye radial distortion compensation and rectification by vanishing point extraction", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32(12):2289–2296, 2010.

[16] R. Hartley, S. B. Kang, "Parameter-free radial distortion correction with center of distortion estimation", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29:(8):1309–1321, 2007.

[17] C. Slama, *Manual of photogrammetry (4th ed.)*, American Society of Photogrammetry, 1980.

[18] K. Miyamoto, "Fish eye lens", *Journal of the Optical Society of America* 54(8):1060–1061, 1964.

[19] T.H. Hopp, *Technical Report NISTIR 5501: The sensitivity of three-point circle fitting*, National Institute of Standards and Technology, 1994.

[20] I. Sobel, G. Feldman, "A 3x3 isotropic gradient operator for image processing", presented (unpublished) at *Stanford Artificial Intelligence Project*, 1968. Originally published in: R. Duda, P. Hart, *Pattern classification and scene analysis*, John Wiley and Sons, 1973.

[21] Q. Wang, R. Ward, "A new orientation-adaptive interpolation method", *IEEE Transactions on Image Processing* 16(4):889–900, 2007.

[22] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical recipes: The art of scientific computing (3rd ed.)*, Cambridge University Press, 2007.

[23] C. Hughes, P. Denny, E. Jones, M. Glavin, "Accuracy of fisheye lens models", *Applied Optics* 49(17):3338-3347, 2010.