A New Approach to Lens Distortion Correction

Donald G Bailey
Institute of Information Sciences and Technology, Massey University, Palmerston North
D.G.Bailey@massey.ac.nz

Abstract

When using an imaging system to make measurements, it is important to correct for lens distortion. An image of a grid is captured, and parabolas are fitted to each of the gridlines. The curvature of the parabolas is used to estimate the radial distortion component. Once this is removed, the image can then be corrected for perspective distortion. The approach presented is a simple, direct approach that provides an accurate first approximation to characterising the distortion.

Keywords: barrel distortion, camera calibration, least squares fit, sub-pixel fitting, aspect ratio, perspective distortion.

1. Introduction

Image processing is widely used in measurement applications to provide rapid non-contact measurements. The accuracy of the results obtained depends not only on the image analysis algorithm, but also the quality of the images obtained. While image resolution is the ultimate limiting factor in terms precision, lens distortion can introduce significant errors. The field of photogrammetry has developed techniques for characterising and compensating for any distortions present, enabling accurate measurements to be made.

One of the most prevalent forms of lens distortion is barrel distortion. It results from the lens having a slightly higher magnification in the centre of the image than at the periphery. The reduction in magnification at the edge causes the edges to shrink around the centre, giving the characteristic barrel shape (see Figure 1). Barrel distortion is particularly noticeable with both wide-angle lenses and inexpensive lenses. Such lenses are commonly found on web-cameras because the wide angle gives better depth of field, making focussing less critical, or even unnecessary. Although web-cameras are not normally used in measurement applications, the distortion can be objectionable, especially if the scene contains straight lines. As a result of their low cost, web-cameras are increasingly being used in robotics and even image analysis applications where the effects of distortion become even more critical. Even when higher quality cameras and lenses are used, the constraints of the imaging environment may dictate the use of wide-angle lenses with their associated distortions.

1.1 Camera Calibration

The procedure for modelling the imaging processes and characterising the distortions introduced by an imaging system is known as camera calibration. Clarke and Fryer [1998] review the history of camera calibration, in particular looking at the developments in characterising lens distortion.

In computer vision, the approach of Tsai [1987] or some derivation is commonly used. These approaches estimate the position and orientation of the camera relative to a target, as well as estimating the lens distortion parameters, and the intrinsic imaging parameters. A full camera calibration requires purpose built three-dimensional models, or multiple images of planar models taken in different poses [Heikkila and Silven, 1996].

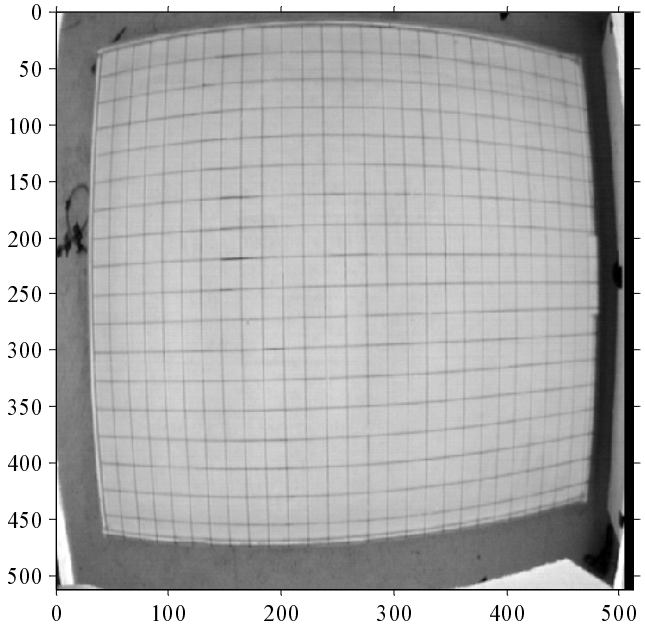

Figure 1: An image with severe barrel distortion. The grid region measures 2.7 m x 1.75 m, and is viewed with a wide angle lens from a distance of 2.4 m. Note there is also significant aspect ratio distortion, and some perspective distortion present.

Generally the 'target' consists of an array of spots, a grid, or a checkerboard pattern. From the construction of the target, the relative positions of the 'target' points are well known. The corresponding points are then located within the image, and the imaging model is adjusted to try and make the target points match their measured image points.

The known location of model enables target points to be measured in 3D world coordinates. This coordinate system is used as the frame of reference. A rigid body transformation (rotation and translation) is applied to the target points. This uses an estimate of the camera pose (position and orientation in world coordinates) to transform the points into a camera centred coordinate system. Then a projective transformation is performed, based on the estimated lens focal length, giving 2D coordinates on the image plane. Next, these are adjusted using the distortion model to account for distortions introduced by the lens. Finally, the sensing element size and aspect ratio are used to determine where the control points should appear in pixel coordinates. The coordinates obtained from the model are compared with the coordinates measured from the image, giving an error. The imaging parameters are then adjusted to minimise the error, resulting in a full characterisation of the imaging model.

The camera and lens model is sufficiently non-linear to preclude a simple, direct calculation of all of the parameters of the imaging model. Correcting imaging systems for lens distortion therefore requires an iterative approach, for example using the Levenberg-Marquardt method of minimising the mean squared error [Press et al., 1996]. One complication of this approach is that for convergence, the initial estimates of the model parameters

must be reasonably close to the final values. This is particularly so with the 3D rotation and perspective transformation parameters.

Planar objects are simpler to construct accurately than full 3D objects. Unfortunately, only knowing the location of points on a single plane is insufficient to determine a full imaging model [Sturm and Maybank, 1999]. Therefore if a planar target is used, several images must be taken of the target in a variety of poses to obtain full 3D information. Alternatively, a reduced model with one or two free parameters may be obtained from a single image. A typical 2D target consists of a grid or array of dots. With a grid target, the intersections of the gridlines provide the control points.

An alternate, the "plumbline" method, goes the other way around and is based on the observation that straight lines are invariant under a perspective (or projective) transformation. Therefore any deviation from straightness must be due to lens distortion [Brown, 1971; Fryer et al., 1994; Park and Hong, 2001]. In these approaches, several points are taken on each line, and used to estimate the lens distortion parameters.

2. Analysis Procedure

The approach taken in this paper works from the lines of a regular square grid. Each gridline is detected, and a parabola is fitted to the detected points. The quadratic term of the parabola provides an estimate of the curvature. The centre of distortion is found by analysing how the curvature changes across the image (both vertically and horizontally), based on the observation that any line through the centre of distortion will remain straight. Next the distortion parameter is determined that will transform the measured parabolas into straight lines. Finally, if necessary, the straight gridlines may be analysed for perspective distortion.

2.1 Grid Detection

The grid is arranged such that the lines are approximately parallel with the edges of this image. This enables the gridlines to be easily detected and processed automatically.

First the image is preprocessed to keep only the line features. It is assumed that the grid consists of dark lines against a lighter background. A 5x5 greyscale opening filter is used to remove the line features. The grid edges can then be detected by finding the largest light object in the centre of the image. This effectively creates a mask, which is used to suppress the irrelevant information outside the grid region (see Figure 2(a)).

The grid region is masked, and 21x1 and 1x21 local minimum filters are used to detect the vertical and horizontal lines respectively. The local minimum filter works by determining the minimum pixel value within the window centred on the pixel being considered. If the centre pixel has the same value as the minimum, then it is classified as a local minimum. Sometimes the detected pixel lines are broken or offset at the grid intersections, so a 3x13 (or 13x3 for the horizontal lines) maximum filter is used to link the grid segments (see Figures 2(b) and (c)).

Any segments that do not span at least half of the image vertically or horizontally are discarded as noise. A parabola is fitted to the detected local minima associated with each gridline region, using minimum least squares. The horizontal gridlines are represented by

$$y = a_i x^2 + b_i x + c_i \tag{1}$$

and the vertical gridlines by

$$x = a_j y^2 + b_j y + c_j \tag{2}$$

where i and j are the gridline indices. The fitted parabolas are shown superimposed on the original image in Figure 2(d). It has been found that even for quite severe distortion, a parabola provides an excellent fit to the data.

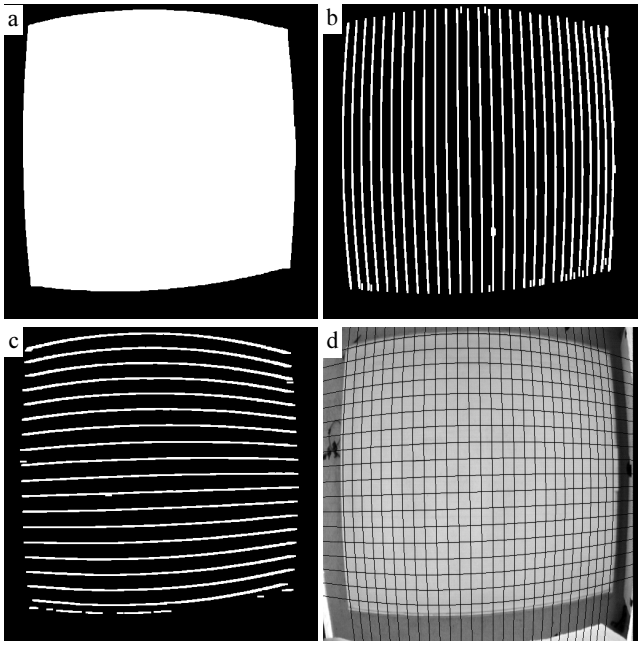

Figure 2: Steps in gridline detection: (a) mask for suppressing the background; (b) detected vertical gridlines; (c) detected horizontal gridlines; (d) fitted parabolas overlaid on original image.

Although the gridlines were only detected to the nearest pixel, the fitted parabola provides a sub-pixel estimate of the grid location. If however, the gridline is straight and well aligned with the edges of the image, the parameterised gridline may be offset by up to half of a pixel from its true position, and the curvature estimate can be wildly incorrect. This problem is illustrated in Figure 3. The error can be significantly improved by using the perpendicular pixels adjacent to the detected grid pixel to estimate the grid position to sub-pixel accuracy.

If p_0 is the value of the local minimum, and p_- and p_+ are the pixel values on either side, then a parabola can be fitted to these three values. The local minimum of this parabola is offset from the central minimum pixel by

$$x_o = \frac{p_+ - p_-}{4p_0 - 2(p_+ + p_-)} \tag{3}$$

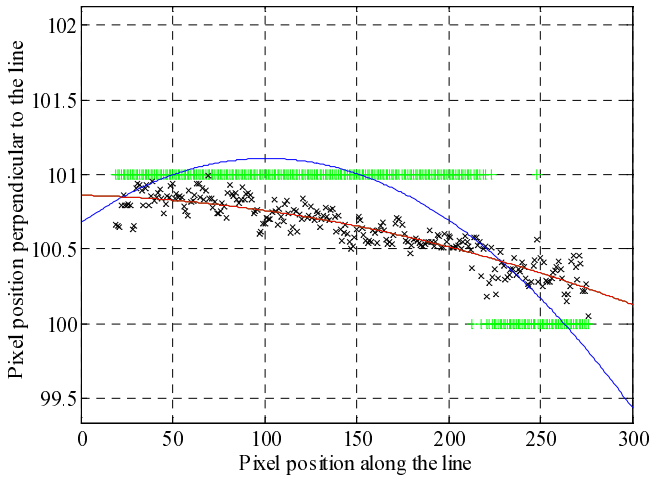

Figure 3: The effect of fitting with sub-pixel accuracy when a line is almost aligned with the grid. The + represent the pixel accuracy lines, and the x represent the sub-pixel estimated positions.

It should be emphasised that if the gridline is not almost perfectly horizontal (or vertical), or it has significant curvature, then the pixel level fit is very close. However since the sub-pixel fit does not incur significant computational cost, there is no harm in using equation (3) to refine the detected grid pixel positions before fitting the grid parabolas.

2.2 Centre Location

The first approximation of the centre of distortion is taken as the centre of the grid. This is estimated as

$$(x_0, y_0) = (\bar{c}_j, \bar{c}_i) \tag{4}$$

Each parabola is offset to make this the new origin. For example equation (1) is transformed to:

$$\begin{aligned} y' &= y - y_0 \\ &= a_i(x' + x_0)^2 + b_i(x' + x_0) + c_i - y_0 \\ &= a_i x'^2 + (2a_i x_0 + b_i)x' + (a_i x_0^2 + b_i x_0 + c_i - y_0) \\ &= a_i' x'^2 + b_i' x' + c_i' \end{aligned} \tag{5}$$

The reason for using this offset is that the intercepts, c_i, are used to estimate the position of the gridlines. The values at the edge of the image (from equation (1)) are not as significant as the values through the centre of curvature (estimated from equation (5)). Adjusting the origin in this way allows a more accurate estimate of the centre of distortion.

Assuming that perspective and radial lens distortion area the only causes of distortion present in the image, the line through the centre of distortion should be straight. With radial distortion, the curvature of the lines will increase with distance from the centre of distortion. The a coefficients represent the curvature and the c coefficients now represent the distances from the centre. The relationship between a and c for all of the parabolas is shown in Figure 4.

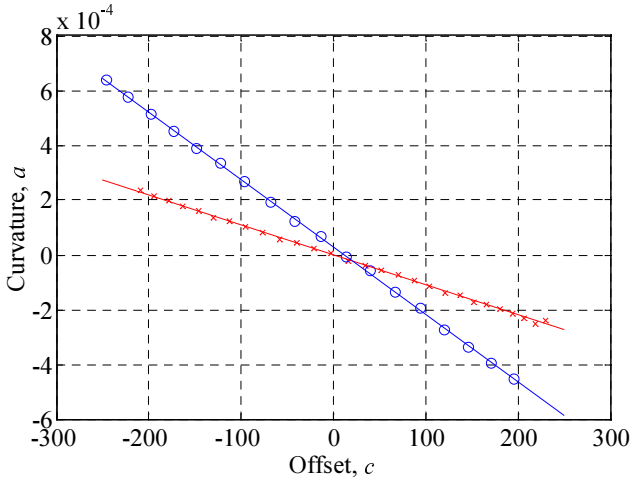

Figure 4: Linear relationship between curvature and offset. The x correspond to the vertical gridlines, and the o to the horizontal.

A straight line is fitted to the curvature data, and solved for $a_i = 0$, giving y_c and $a_j = 0$, giving x_c. The original parabolas are adjusted using equation (5) to make the centre of curvature correspond to the origin.

2.3 Aspect Ratio

Radial distortion is circularly symmetric, so the slopes of the two lines in Figure 4 should be identical. Any difference in slope can be attributed to the aspect ratio of the pixels. Even with nominally square pixels, there can be some aspect ratio distortion present [Bailey, 1995], which must be removed before correcting for radial distortion.

Let s_i be the slope of a_i vs c_i and s_j be the slope of a_j vs c_j. The aspect ratio is approximately

$$AR = \sqrt{s_i / s_j} \ . \tag{6}$$

Scaling the x-axis by this ratio corrects for the aspect ratio distortion:

$$\begin{aligned} a_i &= a_i AR^2 & a_j &= a_j / AR \\ b_i &= b_i AR & b_j &= b_j / AR \\ & & c_j &= c_j / AR \end{aligned} \tag{7}$$

Figure 5 shows the detected gridline parabolas after centring and correcting for aspect ratio distortion.

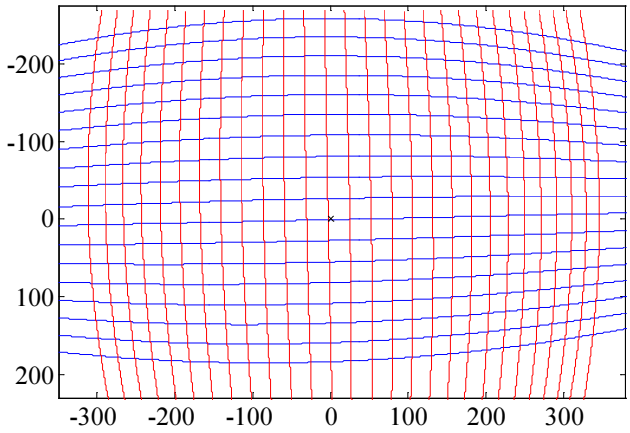

Figure 5: Gridline parabolas after correcting for aspect ratio.

2.4 Lens Distortion

The curvature information is contained in the a coefficients of the parabolas. From these, we are able to estimate the radial distortion. Using the distortion model

$$r_u = r_d(1 + \kappa r_d^2) \tag{8}$$

where r_d and r_u are the distance from the origin in the distorted and undistorted images respectively. Substituting the parabola data (equation (1) with coefficients adjusted according to equations (5) and (7)) for r_d for each of the gridlines gives

$$\begin{aligned} y_u &= y_d\left(1 + \kappa\left(x^2 + y_d^2\right)\right) \\ &= (ax^2 + bx + c)\left(1 + \kappa\left(x^2 + (ax^2 + bx + c)^2\right)\right) \\ &= c(1 + \kappa c^2) + b(1 + 3\kappa c^2)x + \left(a + c\kappa(3ac + 3b^2 + 1)\right)x^2 \\ &\quad + \text{higher order terms} \end{aligned} \tag{9}$$

This equation assumes small offsets from the centre for each parabola, so that

$$x_u \approx x_d \tag{10}$$

and that the contribution of the higher powers of x are negligible. While not strictly true, this will select the value of κ that is maximally flat at the tangent at x=0. Since the undistorted equation should be a straight line, the coefficient of the quadratic term in equation (9) is set to 0, and the higher order terms ignored as insignificant. Solving for κ :

$$a + c\kappa(3ac + 3b^2 + 1) = 0 \tag{11}$$

gives

$$\kappa = \frac{-a}{c(3ac + 3b^2 + 1)} \approx \frac{-a}{c} = -s \tag{12}$$

κ was for determined for each parabola. These values were all similar (apart from noise) with an increase with the distance of the parabola from the centre of distortion (see Figure 6). The outliers apparent for small offsets result from the division by c in equation (12). The curvature indicates that to model the distortion more accurately, it will be necessary to use higher order terms in equation (8). To obtain a single value of κ a weighted average of the measured values was taken, weighted by the c values (distance from origin) after discarding the extreme outliers. This transforms the parabolas to the lines:

$$\begin{aligned} y &= b_i(3\kappa c_i^2+1)x + c_i(\kappa c_i^2+1) \\ &= m_i x + d_i \end{aligned} \tag{13}$$

and similarly for the vertical gridlines.

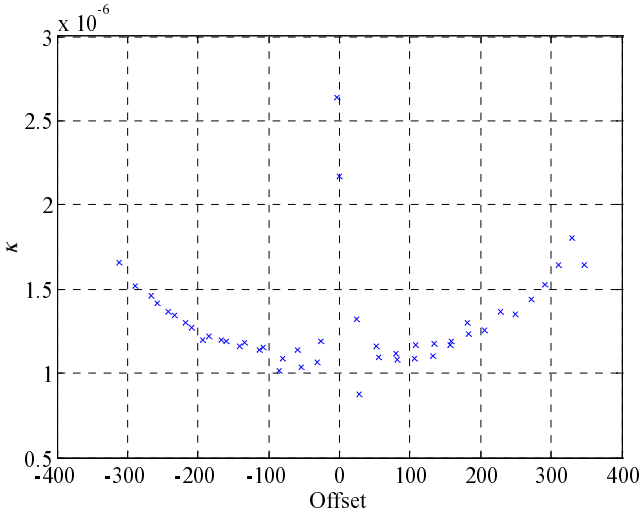

Figure 6: Distortion parameter calculated from each gridline.

At this stage, there is sufficient information to characterise and remove the lens distortion from the image. The results of such a correction are shown in Figure 7. Note that there is still some deviation from linear in the top corners of the image. This is because simple model used here is unable to completely account for such severe distortion.

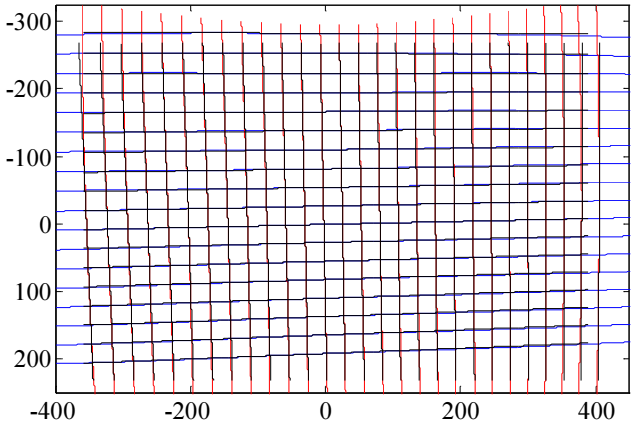

Figure 7: Results of correcting for lens distortion. Both the warped parabola, and the linear approximations are shown.

2.5 Perspective Distortion

A further step allows a 2D perspective transformation to be determined from the gridlines. While this is insufficient to completely determine the position of the camera relative to the target grid [Sturm and Maybank, 1999], it does give the direction to the camera (and consequently, the orientation of the camera). Determining the distance from the camera requires either 2 (or more) targets at different ranges, or knowledge of the camera focal length and sensor size.

In homogenous coordinates, a point **P** at (x, y) is represented by the 3 element column vector

$$\mathbf{P} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{14}$$

A 2D perspective transformation can the be represented as multiplying **P** by a **H**, a 3x3 transformation matrix, to give the transformed point **P′** at (x', y'):

$$k\mathbf{P}' = \begin{bmatrix} kx' \\ ky' \\ k \end{bmatrix} = \mathbf{HP} \tag{15}$$

where k is a scaling factor required to normalise the point. The transform matrix **H** has only with 8 degrees of freedom since scaling **H** by a constant will only change the value of k but will leave the transformed point unchanged.

A line **L** can be represented in homogenous coordinates as a row vector:

$$ax + by + c = 0 \Rightarrow \mathbf{L} = \begin{bmatrix} a & b & c \end{bmatrix} \tag{16}$$

A point will be on a line if

$$\mathbf{LP} = 0 \tag{17}$$

Equations (15) and (17) can be combined to give the transformation of lines in the homogenous coordinate system [Hartley and Zisserman, 2000]:

$$\mathbf{L}' = \mathbf{L}\mathbf{H}^{-1} \tag{18}$$

Correcting for perspective distortion transforms the horizontal gridlines from equation (13) as follows:

$$\mathbf{L}_i' = \begin{bmatrix} 0 & -1 & i \end{bmatrix} = \mathbf{L}_i\mathbf{H}^{-1} = \begin{bmatrix} m_i & -1 & d_i \end{bmatrix}\mathbf{H}^{-1} \tag{19}$$

and similarly for the vertical gridlines:

$$\mathbf{L}_j' = \begin{bmatrix} -1 & 0 & j \end{bmatrix} = \mathbf{L}_j\mathbf{H}^{-1} = \begin{bmatrix} -1 & m_j & d_j \end{bmatrix}\mathbf{H}^{-1} \tag{20}$$

The use of i and j as the transformed line offsets relies on the assumption that the gridlines are equally spaced. This will be the case if a square grid is used. A rectangular grid also works, because there is no requirement that the horizontal and vertical spacing be identical. The grid aspect ratio will be absorbed within the perspective transformation.

Each gridline provides 3 equations in the coefficients of $\mathbf{H}^{-1}$. So, with a total of N total gridlines (both horizontal and vertical), equations (19) and (20) provide $3N$ equations. However since the equation of line has only 2 independent variables (scaling **L** by a constant, k**L**, represents the same line as **L**) this reduces the number of independent equations to $2N$. If the coefficients of $\mathbf{H}^{-1}$ are

$$\mathbf{H}^{-1} = \begin{bmatrix} h_1 & h_4 & h_7 \\ h_2 & h_5 & h_8 \\ h_3 & h_6 & h_9 \end{bmatrix} \tag{21}$$

then eliminating the arbitrary scaling constant for each line gives the following system of equations:

$$\left[\begin{array}{cccccccccc} m_i & -1 & d_i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & im_i & -i & id_i & m_i & -1 & d_i \\ \hline 0 & 0 & 0 & -1 & m_j & d_j & 0 & 0 & 0 \\ -j & jm_j & jd_j & 0 & 0 & 0 & -1 & m_j & d_j \end{array}\right] \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_9 \end{bmatrix} = \mathbf{0} \tag{22}$$

Finding a nontrivial solution to this requires determining the nullspace of the $2N$x9 matrix. This can be found through singular value decomposition, and selecting the vector corresponding to the smallest singular value [Press et al., 1993]. Performing the

singular value decomposition usually involves an iterative process, so an alternative approach is to set one of the coefficients, and solve for the rest using least squares. If the perspective distortion is small, which will be the case if the plane containing the 2D grid is approximately perpendicular to the camera, then setting h_9 to 1 will work. Alternatively, since the grid is approximately aligned with the axes, setting either h_1 or h_5 to 1 will also work.

3.0 Results and Discussion

The results of solving equation (22) and applying the resulting perspective transformation to the data in Figure 7 are shown in Figure 8. The axes are scaled in grid units because of the coordinates used for the representation of $\mathbf{L}'$ in equations (19) and (20). Most of the distortion has been removed through this process.

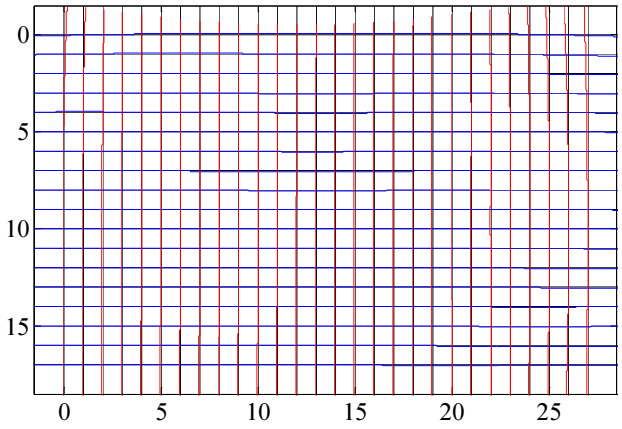

Figure 8: After correcting for 2D perspective distortion. The image axes are in grid units.

To evaluate the performance of the distortion characterisation, the intersections of the gridlines in the original image are located by finding the intersections of the corresponding horizontal and vertical parabolas from equations (1) and (2). These are then offset by the centre of distortion, and the x value adjusted for the aspect ratio. Equation (8) is then used to correct for lens distortion, and equation (14) for perspective distortion. The RMS residual of the error between the corrected position and the known grid coordinates is 0.038 grid units (approximately 1 pixel in the original image). Some of this residual will result from measurement error and uncertainty in fitting the parabolas to the gridlines in the original image, but most of it is from the limitations of the distortion model used.

While the direct approach described in this paper is able to model much of the distortion, the results may be refined further by incorporating additional radial and tangential distortion terms. These additional terms cannot be calculated directly because of the non-linear nature of the model. Levenberg-Marquardt minimization, using the calculated parameters as a starting point, converges very quickly because the initial estimate is close to the final result. With the addition of a tangential component and a second radial distortion component, Levenberg-Marquardt minimization converged with 5 iterations to an RMS residual of 0.017 grid units (approximately 0.5 pixel). While this is better than the direct estimate, it indicates that the original estimate was very close.

To correct for distortion present in the image, the model is applied using the reverse procedure to the analysis described in section 2.

First, the desired mapping between the grid and the output image is determined. To align the grid with the pixel coordinates, this is simply a scaling and an offset. If a rectangular grid is used, the horizontal and vertical scale factors will be different to reflect the grid aspect ratio.

For each pixel in the output image, the coordinates are transformed using inverse of the perspective mapping defined in equation (15). The required matrix, $\mathbf{H}^{-1}$, is that calculated directly from equation (22). This corrects the coordinates for the perspective distortion observed in the grid.

Next, the lens distortion is corrected by applying the inverse of equation (8) to the coordinates. Unfortunately, this provides the transformation from distorted to undistorted coordinates, and not the other way around as desired. If equation (8) is represented the other way around, the analysis procedure described in section 2.4 does not provide a convenient closed form solution. To overcome this problem, the normalised magnification factor is calculated as a function of undistorted coordinates:

$$r_u \Rightarrow F(r_u) = \frac{r_d}{r_u} = \frac{1}{(1+\kappa r_d^2)} \tag{23}$$

This mapping was approximated by a fifth order polynomial, with a maximum error of 2% of a pixel for typical image sizes. If higher accuracy is desired, a higher order approximation may be used. The approximation enabled the correction to be made:

$$r_d \approx r_u F(r_u) \tag{24}$$

or in terms or image coordinates:

$$\begin{aligned} x_d &= x_u F\left(\sqrt{x_u^2 + y_u^2}\right) \\ y_d &= y_u F\left(\sqrt{x_u^2 + y_u^2}\right) \end{aligned} \tag{25}$$

The distorted coordinates are then adjusted for aspect ratio, and offset by the centre of distortion to give the pixel locations in the distorted image. Finally, bicubic interpolation is used to estimate the pixel value in the undistorted image from the neighbouring pixels in the distorted image. The resulting corrected image is shown in Figure 9.

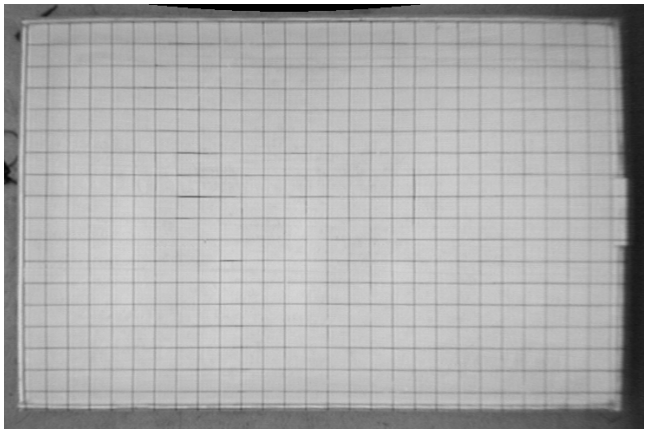

Figure 9: The corrected grid image.

Some distortion is still apparent in this image, particularly in the corners. From the curvature in Figure 6, the selection of the average value for κ implies that the distortion will be over corrected in the centre, and under corrected in the periphery. This would be corrected by using a second order term in equation (8).

If the distortion is less severe than in the example shown here, fewer gridlines could be used. In fact, a single square or rectangular target would suffice as the two pairs of parallel lines would provide sufficient data for determining all of the necessary parameters.

4. Summary

The proposed method directly estimates the pixel aspect ratio, the centre of radial distortion, the first order radial distortion parameter, and the 2D perspective transformation present. The approach of works form an image of a 2D rectangular grid, and accomplishes its goal by approximating the distorted gridlines by parabolas. The resulting model can then be used to correct for distortion present in the image, either by transforming the image, or by correcting the locations of detected feature points.

Even in cases of severe distortion, this approach gives excellent results. If further refinement of the model is necessary, the parameters obtained may be used as the starting point for more conventional iterative refinement methods. This allows additional higher order terms to incorporated into the model.

References:

BAILEY, D.G. 1995. Pixel calibration techniques. *Proceedings of the New Zealand Image and Vision Computing '95 Workshop,* 37-42.

BROWN, D.C. 1971 Close range camera calibration *Photogrammetric Engineering* 37(8): 855-866.

CLARKE, T.A., FRYER, J.F. 1998. The development of camera calibration methods and models, *Photogrammetric Record*, 16(91): 51-66.

FRYER, J.G., CLARKE, T.A., CHEN, J. 1994. Lens distortion for simple C-mount lenses. *International Archives of Photogrammetry and Remote Sensing*, 30(5): 97-101.

HARTLEY, R., ZISSERMAN A. 2000. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

HEIKKILA, J., SILVEN, O. 1996. Calibration procedure for short focal length off-the-shelf CCD cameras. *Proceedings International Conference on Pattern Recognition*, 1: 166-170.

PARK, S.W., HONG, K.S. 2001. Practical ways to calculate camera lens distortion for real-time camera calibration. *Pattern Recognition*, 34: 1199-1206.

PRESS, W.H., FLANNERY, B.P., TEUKOLSKY, S.A., VETTERLING, W.T. 1993. *Numerical recipes in C: The art of scientific computing*. Cambridge University Press.

STURM, P.F., MAYBANK, S.J. 1999. On plane-based camera calibration: a general algorithm, singularities, applications. *IEEE Conference on Computer Vision and Pattern Recognition*, 1: 432-437.

TSAI, R.Y. 1987. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation* 3(4): 323-344.