Error Assessment of Robot Soccer Imaging System

Donald Bailey, Gourab Sen Gupta

Institute of Information Sciences and Technology,
Massey University, Palmerston North, New Zealand
D.G.Bailey@massey.ac.nz, g.sengupta@massey.ac.nz

Abstract

Position and angle errors are an obstacle to accurate, high-speed control of micro-robots in a robot soccer system. Systematic errors resulting from constraints on the global imaging system are identified, and their magnitude estimated. A full calibration procedure is proposed to minimise these errors that only uses information available from the soccer field. This calibration method has a residual error of 1.75 mm, or 0.3 pixel, and may be applied in real-time with negligible additional computational expense.

Keywords: camera calibration, lens distortion, perspective

1 Introduction

Micro robots are used in education and entertainment in many ways [1]. There are robots which navigate a maze, climb a wall, play a game of soccer, wrestle with another robot, run round a track, balance a pole, mow the lawn and vacuum clean the house.

Robot Soccer has become increasingly popular over the last decade not only as a platform for education and entertainment but as a test bed for adaptive control of dynamic systems in a multi-agent collaborative environment [2]. It is a powerful vehicle for dissemination of scientific knowledge in a fun and exciting manner. It encompasses several technologies—embedded micro-controller based hardware, wireless radio-frequency data transmission, dynamics and kinematics of motion, motion control algorithms, real-time image capture and processing and multi-agent collaboration. Because of the dynamics and high complexity of the robot soccer system as well as manoeuvrability and high speed of its robots, the accurate and real-time detection of position and orientation of objects has gained special importance as it greatly affects path planning, prediction of moving targets and obstacle avoidance.

The vision system is an integral component of modern autonomous mobile robots. One area that remains a challenge is the development of a reliable and accurate real-time global vision system for robot soccer. In this system, all position data is obtained from a global camera mounted above the centre of the table (see Figure 1). In ideal lighting conditions and large setup times, good progress has been made, but the robustness of these vision systems is questionable when the lighting requirements are not met [3].

Within the robot soccer environment, the position and orientation of the robots is determined by analysing the global image. Each robot is identified by a "jacket" which has coloured patches on it. The location of these coloured patches within the image is used to estimate the location and orientation of the robot within the playing area.

To manage complexity in collaborative robot systems, a hierarchical state transition based supervisory control (STBS) system has been proposed and successfully implemented [4]. However, the performance of such a system deteriorates substantially if the objects are not detected accurately because the generic control functions to position and orient the agents are no more reliable.

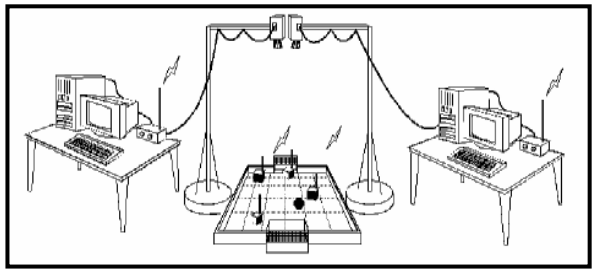

Figure 1: Setup of the robot soccer platform.

1.1 Constraints

A number of practical constraints limit the imaging system. As robot soccer is a high-speed game, with the robots and ball moving at speeds of up to 2 metres per second, high temporal resolution is essential. For this the separate fields of an interlaced-scan video camera are used, giving 60 images per second. Using each field as a separate image reduces the spatial resolution, typically to 240x320 pixels. These images must be processed, and instructions sent to the robots within the 16.67 ms before the next field arrives.

The low resolution of the images facilitates their processing within the available time, but at approximately 5 mm per pixel resolution, severely limits the ability to precisely locate the objects. Each robot occupies approximately 15x15 pixels, with each coloured patch occupying 35-50 pixels. Using the centre of gravity of the patch enables it to be located

to sub-pixel accuracy. A 1 pixel error on the boundary of the patch will move the centre of gravity by 0.1 pixels or 0.5 mm, so this is the absolute detection limit. In practise the uncertainty in measurement is about 1mm. In a similar way, the detection limit on the robot angle is about 1º, with a typical uncertainty in the angle measurement of 2 to 3º.

A second constraint is that in a robot soccer tournament, each team is responsible for providing their own camera. Both cameras cannot be mounted exactly over the centre of the table. Therefore the camera may be looking down on the table at a slight angle, introducing mild perspective distortion. Barrel distortion is another source of error in the vision processing which requires attention.

1.2 Importance of errors

In a game of robot soccer, position errors are not really an issue. If the robot is manoeuvring to manipulate the ball or avoid another robot they will be in similar positions on the field. Therefore they will be subject to similar errors, and the relative error is insignificant. When multiple robots are collaborating on a task and are required to move in formation, position errors are more important.

On the other hand, angle errors are more critical both in playing soccer and in a collaborative environment when multiple robots are working together to accomplish a task. Small angle errors become amplified with distance as the robot moves. While such errors may be detected and compensated for in subsequent frames, when moving at high speed there is limited data and limited time to make any corrections. This may affect issues such as being able to track the ball to determine where to intercept it (for example by the goalie), or being able to shoot a goal from a distance.

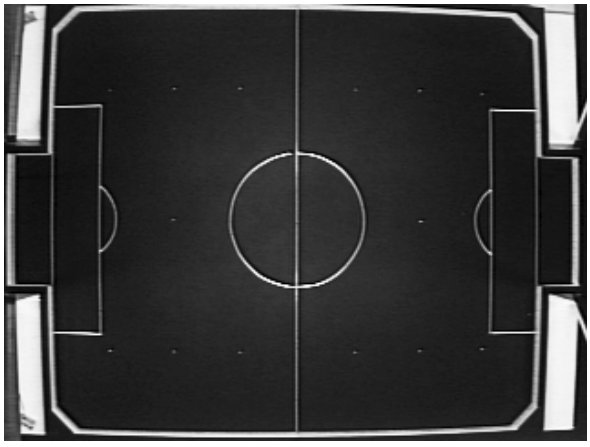

Figure 2: The system's view of the robot soccer table.

2 Error Sources

Figure 2 shows a typical image of the field obtained from the camera. Both barrel (lens) distortion and perspective distortion are clearly visible, although their effect is relatively mild.

A simple calibration may be performed by assuming that the distortion is negligible. This locates the column positions of the goal mouths (C_{Left} and C_{Right}), and the row positions of the field edges at the centreline (R_{Top} and R_{Bottom}). An object positioned at row R and column C within the image may be determined relative to the centre of the field as

$$x = 75\frac{2C - (C_{Left} + C_{Right})}{C_{Right} - C_{Left}}$$
$$y = 65\frac{2R - (R_{Top} + R_{Bottom})}{R_{Bottom} - R_{Top}} \quad (1)$$

where 75 is half the field length in cm, and 65 is half the field width in cm.

While this calibration is simple to perform, it will only be accurate if there is no distortion within the image. The rest of this paper considers a more complete calibration that takes into account the distortion, and uses this to estimate the errors associated with this simple calibration approach.

2.1 Lens distortion

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. Barrel distortion is particularly noticeable with wide-angle lenses such as those used with robot soccer.

A simple radial distortion model is usually sufficient to account for most of the distortion [5]:

$$r_u = r_d(1 + \kappa r_d^2) \quad (2)$$

where κ is the distortion parameter, and r_u and r_d are the distances from the centre of distortion in the undistorted and distorted images respectively.

2.1.1 Effect on Position

If the simple calibration was based on distances in the centre of the image, the position error would increase with radius according to the radially dependent magnification. However, since the calibration of equation (1) sets the positions at the edges and ends of the field, the position error there will be minimal. The absolute position errors should also be zero near the centre of the image, increase with radius to a local maximum between the centre and edges of the playing area, and decrease to zero again on an ellipse through the table edge and goal points. Outside this ellipse, the errors will increase rapidly with distance, affecting the corners of the playing field.

2.1.2 Effect on Angle

Determining the effect of radial distortion on the angle is more complicated. Consider a test point in the undistorted image using radial coordinates (r_u, φ). At this point the lens distortion results in a magnification, *M*. Next consider another point offset from this by a small distance at an angle θ_u. If the magnification is constant at the two points then everything is scaled equally, and by considering similar triangles, there is no distortion. Therefore angle distortion is the result of a change in magnification with position (radius in the case of lens distortion).

$$\tan(\varphi - \theta_d) \approx \frac{M}{M + 2r_u^2 \frac{dM}{dr_u^2}} \tan(\varphi - \theta_u) \quad (3)$$

There is no angle error in the radial direction because the offset point is in line with the test point ($\varphi - \theta = 0$). There is also no error in the tangential direction because the offset point and test point are subject to the same magnification and the angle is unchanged ($\tan(\varphi - \theta) = \infty$).

Since the magnification changes faster with increasing radius, the angle error will also be larger further from the centre of distortion.

2.2 Perspective distortion

Perspective distortion results when the line of sight of the camera is not perpendicular to the plane of the playing area. This will occur when the camera is not directly over the centre of the playing area, and it must be tilted to fit the complete playing area within the field of view. Perspective distortion is often modelled using a homogenous coordinate system [6]:

$$\begin{bmatrix} kx_d \\ ky_d \\ k \end{bmatrix} = \mathbf{H} \begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix} \quad (4)$$

where (x_u,y_u) and (x_d,y_d) are the coordinates of an undistorted and distorted point, and **H** is a 3x3 transformation matrix. The matrix **H** incorporates rotation, translation, scaling, skew, and stretch as well as perspective distortion. Just considering perspective distortion, **H** simplifies to

$$\mathbf{H}_{perspective} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ p_x & p_y & 1 \end{bmatrix} \quad (5)$$

2.2.1 Effect on Position

From equation (4), this corresponds to changing the scaling factor, *k*, giving a position dependent magnification. One side (or corner) will have a larger magnification than the opposite corner. Since the simple calibration sets four points on the edges of the playing area, these points will have no error. On the side (or end) where the magnification is larger, the position errors will be positive, and on the side where the magnification is smaller the position error will be negative. A line approximately across the middle of the table will have no error. The angle of the line, and the severity of the errors will depend on the angle and extent of the camera tilt respectively.

2.2.2 Effect on Angle

Again the distortion will depend on the change of magnification with position. As the content becomes more compressed (closer to the perspective vanishing line) angle distortion will increase. This is because as $p_x x_u + p_y y_u + 1$ approaches 0, the slope of the magnification becomes steeper and the angle errors will be greater.

With the mild perspective distortion of the robot soccer, this effect will be less pronounced. However, angle errors will be larger on the side of the field where the image appears compressed.

3 Full Calibration Procedure

Traditional camera calibration procedures require a dense set of points scattered throughout the image [7,8]. Without providing a custom target, relatively few data points are available from the robot soccer platform. However the distortion is obvious from the edges of the playing area, so the method presented by Bailey [9] that uses a perpendicular set of parallel lines has been adapted. The two sides and the two ends of the playing area provide a minimum set of lines for characterising both lens and perspective distortion.

3.1 Edge Segmentation

The white boundaries of the playing area are actually the inside walls of the table. There is a distinct step edge with high contrast that may be located to sub-pixel accuracy [10]. The field lines (and also any robots) on the table will also have a response to the filter. Therefore it is necessary to create a mask to select only the edges of interest. This mask was constructed as follows:

- A morphological filter removes vertical lines. This removes the field line across the mouth of the goal, and makes the whole playing area a single connected component.
- A threshold operation detects the white borders of the playing area (and other remaining light features in the image).
- A flood fill in the centre of the image selects the connected component belonging to the table.
- A morphological filter is used to remove the remaining horizontal field lines, and any holes (corresponding to robots on the field, and other larger objects) are filled.

- The edges of this structure are detected giving the mask shown in Figure 3.

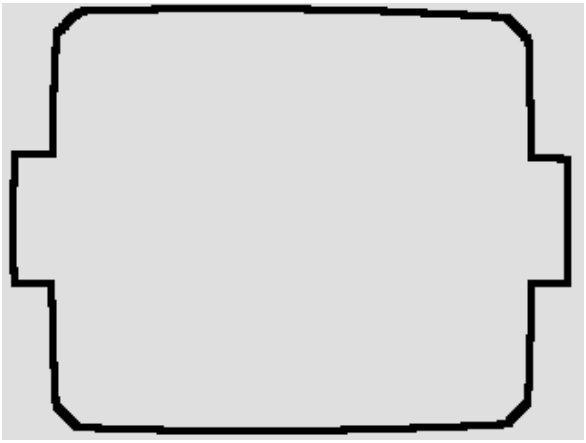

Figure 3: Edge mask constructed from Figure 2.

0	0	0
1	0	-1
0	0	0

0	0	0
-1	0	1
0	0	0

0	1	0
0	0	0
0	-1	0

0	-1	0
0	0	0
0	1	0

Figure 4: Local linear filter masks for detecting the left, right, top and bottom edges of the soccer field.

A set of local edge detection filters (see Figure 4) is applied to the image to detect the playing area edges. The position of the local maximum is found, and a parabola is fitted to the maximum and adjacent pixels perpendicular to the edge. The subpixel offset of the edge is then given by

$$offset = \frac{p_+ - p_-}{4p_0 - 2(p_+ + p_-)} \tag{6}$$

where p_0 is the local edge maximum, and p_+ and p_- are the adjacent filter responses in the positive and negative directions respectively [10].

As the detected edge points have noise, and may contain errors (particularly where field lines run to the edge or there are robots adjacent to the edge) it is necessary to eliminate outliers while fitting a smooth curve to the edge. Along each of the 4 sides the median edge position is determined. A parabola is fitted to the points within 1 pixel of the median. Finally the pixels that are within 0.5 pixels of this parabola are selected, and the parabola refined using these points. The residual error from this fitting process is 0.12 pixel or 0.6 mm. The 4 parabolas are:

$$\begin{aligned} x_{Left} &= a_L y^2 + b_L y + c_L \\ x_{Right} &= a_R y^2 + b_R y + c_R \\ y_{Top} &= a_T x^2 + b_T x + c_T \\ y_{Bottom} &= a_B x^2 + b_B x + c_B \end{aligned} \tag{7}$$

3.2 Distortion Characterisation

The curvature of the parabolas conveys information about the lens distortion [9] because the edges should be straight. The centre of radial distortion is determined by

$$(x_c, y_c) = \left(\frac{a_R c_L - a_L c_R}{a_R - a_L}, \frac{a_B c_T - a_T c_B}{a_B - a_T} \right) \tag{8}$$

and distortion aspect ratio is

$$AR = \sqrt{\frac{(a_R - a_L)(c_B - c_T)}{(c_R - c_L)(a_B - a_T)}} \tag{9}$$

The parabola equations are adjusted to make the centre of distortion the origin, and the x axis is scaled by AR to make the distortion radially symmetric.

The distortion parameter for equation (2) is then calculated as

$$\kappa = \underset{i \in \{L,R,T,B\}}{mean} \left\{ \frac{-a_i}{c_i(3a_i c_i + 3b_i^2 + 1)} \right\} \tag{10}$$

In correcting for the distortion, the parabolas in equation (7) become straight lines:

$$x_{Left} = \frac{b_L(3\kappa c_L^2 + 1)}{\kappa c_L^2 + 1} y + c_L(\kappa c_L^2 + 1) = m_L y + d_L \tag{11}$$

and similarly for the other parabolas. These four lines then provide sufficient data to solve for the 8 unknowns of **H** in the perspective transformation of equation (4):

$$\begin{bmatrix} m_T & -1 & d_T & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -65m_T & 65 & 65d_T & m_T & -1 \\ m_B & -1 & d_B & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 65m_B & -65 & -65d_B & m_B & -1 \\ 0 & 0 & 0 & -1 & m_L & d_L & 0 & 0 \\ 75 & -75m_L & -75d_L & 0 & 0 & 0 & -1 & m_L \\ 0 & 0 & 0 & -1 & m_R & d_R & 0 & 0 \\ -75 & 75m_R & 75d_R & 0 & 0 & 0 & -1 & m_R \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \end{bmatrix} = \begin{bmatrix} 0 \\ -d_T \\ 0 \\ -d_B \\ 0 \\ -d_L \\ 0 \\ -d_R \end{bmatrix} \tag{12}$$

where

$$\mathbf{H} = \begin{bmatrix} h_1 & h_4 & h_7 \\ h_2 & h_5 & h_8 \\ h_3 & h_6 & 1 \end{bmatrix} \tag{13}$$

The calibration model therefore has 12 parameters: x_c, y_c, AR, κ, and h_1–h_8 for the perspective transformation.

3.3 Distortion Correction

The basic approach to correct for distortions in the image is to apply equations (2) and (4) to the distorted coordinates of each point in the image. However, rather than correct the whole image before detecting the robots, it is computationally less intensive to perform the detection and segmentation of the robot jackets within the distorted image, and transform the centre of gravity of each detected blob before determining the final position and orientation.

First the distorted pixel is adjusted for the centre and aspect ratio, and then the radial and perspective corrections applied:

$$(x_o, y_o) = \left((x_d - x_c)/AR, (y_d - y_c)\right) \quad (14)$$

$$\begin{bmatrix} kx_u \\ ky_y \\ k \end{bmatrix} = \mathbf{H}^{-1} \begin{bmatrix} x_o(1+\kappa(x_o^2+y_o^2)) \\ y_o(1+\kappa(x_o^2+y_o^2)) \\ 1 \end{bmatrix} \quad (15)$$

Finally the point is normalised by dividing by k to get the true undistorted location.

3.4 Model Validation

The full calibration has been performed with limited data, taken only from the edges of the playing area. It is expected therefore that the data used would transform to the target location on the edges of the playing area, and indeed this is the case – the RMS residual error is again 0.6 mm.

However, to have confidence that the model is actually correcting points anywhere in the playing area, it is necessary to check the transformation at a number of points scattered throughout the image. For this purpose, a set of 76 points was extracted from throughout the playing area using the field lines and free kick markers as input. The residual error from the validation points was 1.75 mm, which corresponds to about 30% of the width of a pixel.

4 Calibration Assessment

The simple calibration of equation (1) will introduce systematic errors as a result of the assumptions made. These systematic errors may be estimated by comparing the simple calibration with the full calibration.

4.1 Position Error

The results from equation (1) are subtracted from equation (15) to give the position error. This was over a dense grid 1cm apart covering the whole playing area. The results are shown in Figure 5 to Figure **7**.

As expected, the x error at the goal mouth is small and the y error at the sides of the playing area is small. In between, the perspective error resulting from the camera tilt results in a positive and negative peak. This is most pronounced in the x direction where the tilt and hence perspective distortion is greater. The effect of lens distortion is most noticeable in the corners where the error starts to increase rapidly with increasing radius.

The fact that the total error is least at the goals is not surprising because these points were used to set the simple calibration. What is interesting is that the left side of the table (in these views) has significantly less error that the right. In the upper right corner the error is the worst at over 3 cm.

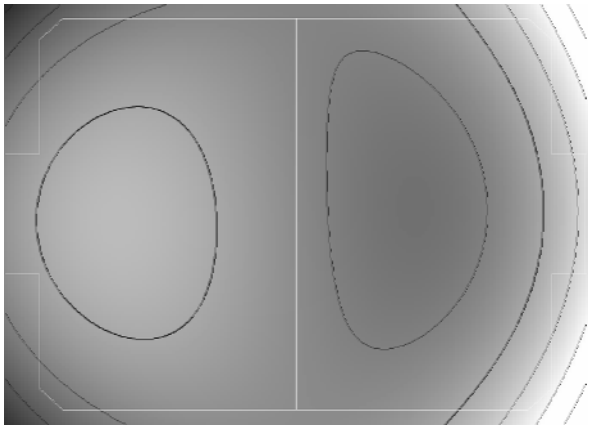

Figure 5: Error in the *x* direction. Contours are drawn every 1 cm of error, with the 0 error contour bold.

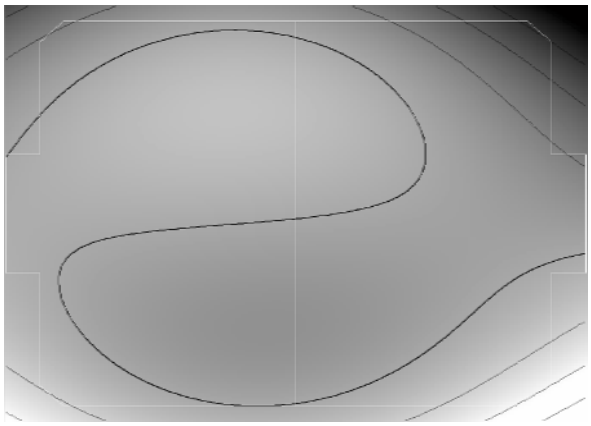

Figure 6: Error in the *y* direction. Contours are drawn every 1 cm of error, with the 0 error contour bold.

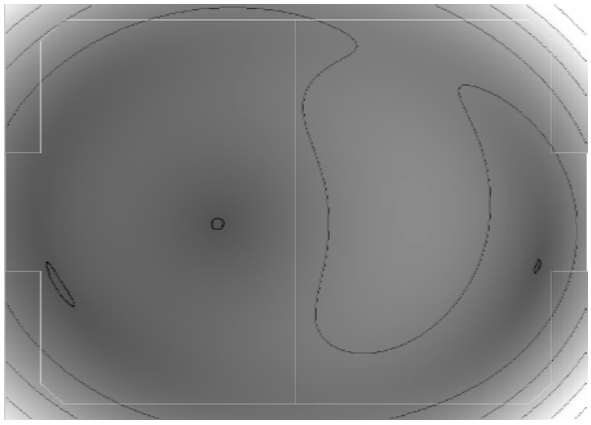

Figure 7: Absolute error. Contours are drawn every 1 cm of error, with the 0.1 mm error contour bold.

4.2 Angle Error

To assess the angle error, at each point a test point was offset by 0.1 mm for a range of angles. The angle of the test point was measured using both the simple calibration and full calibration methods. The difference in angle from the two methods is the angle error. The maximum error (over all angles) is shown in Figure 8 with the angular dependence of the error shown in Figure 9.

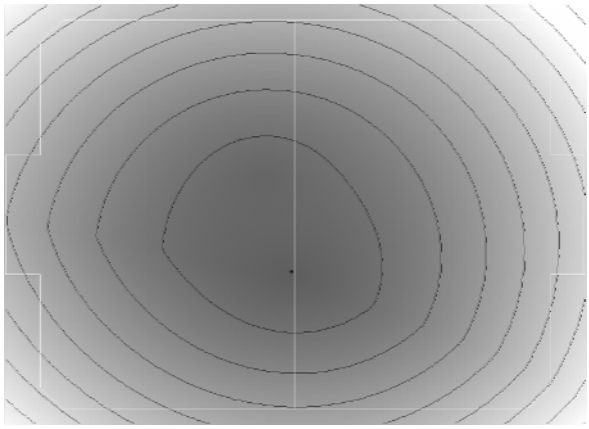

Figure 8: Maximum angle error as a function of position. Contours are drawn every 0.5°.

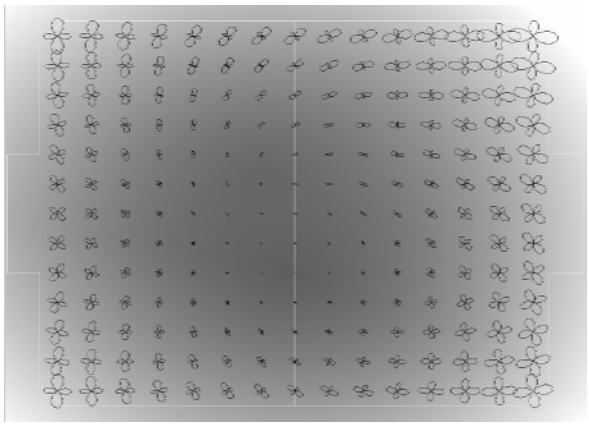

Figure 9: The angle dependence of the angle error measured at 10 cm spacing. The angle error is plotted radially as a function of angle at the test points.

As expected, the angle error is minimum near the centre of the table, and generally increases with radius. The angle error is also 0 in the general direction of the centre of the table, and tangential to this. These effects result from the dominance of the radial lens distortion component on the angle error. The angle error is worse on the right side of the table because of the perspective distortion, with the maximum error being almost 4° in the top right corner.

What is unexpected is the double minimum near the centre of the table. This is the result of interaction between the angle errors from the perspective and lens distortions. At the second minimum (slightly above the centre of the table), the angle error is constant. The asymmetry in the error as a function of angle seen in Figure 9 is also a result of the perspective distortion.

5 Summary

The position and angle errors resulting from lens and perspective distortions are significant. A simple calibration is able to ensure that the position errors are minimal at the critical goal areas, but the errors in the corners may be up to 3 cm – almost half the width of a robot. Angle errors are worst at the ends of the playing field, and are almost 4° in one corner.

Implementing a full calibration may significantly reduce these systematic errors. The proposed calibration method uses the easily detectible edges of the playing field to correct for both lens and perspective distortions. The residual error after correcting for distortion is 1.75 mm – significantly less than a pixel.

The calibration parameters may be measured offline, and applied to the detected blob centres in real-time, almost completely eliminating the systematic errors introduced by lens and perspective distortion.

6 References

[1] Kim, J.H., Jung, M.J., "Micro-Robots for Education and Entertainment", *Proceedings of International Symposium on Autonomous Robots and Agents, ISARA 2000*, 1-7 (2000).

[2] Messom, C.H., "Robot Soccer – Sensing, Planning, Strategy and Control, a distributed real time intelligent system approach", *International Symposium on Artificial Life and Robotics, AROB,* 422-426 (1998).

[3] Rong, X., Jian, C., "'A Global Vision System: using Hue Thresholds to Extract feature and recognize", *Journal of Harbin Institute of Technology*, 8(3): 233-238, (2001).

[4] Sen Gupta, G., Messom, C.H., Sng, H.L., "State Transition Based Supervisory Control for Robot Soccer System", *International Workshop on Electronic Design, Test and Applications*, 338-342 (2002).

[5] Li, M., Lavest, J.M., "Some aspects of Zoom lens camera calibration", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18(11): 1105-1110, (1996).

[6] Hartley, R., Zisserman, A., *Multiple View Geometry in Computer Vision*. Cambridge University Press, (2000).

[7] Tsai, R.Y., "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 (1987).

[8] Heikkila, J., Silven, O., "Calibration procedure for short focal length off-the-shelf CCD cameras. *Proceedings International Conference on Pattern Recognition*, 1: 166-170 (1996).

[9] Bailey, D.G., "A new approach to Lens Disortion Correction", *Proceedings of Image and Vision Computing N Z Conference*, 59-64 (2002).

[10] Bailey, D.G., "Sub-pixel estimation of local extrema", *Proceedings of Image and Vision Computing NZ Conference*, 414-419 (2003).