Automatic Estimation of Camera Position in Robot Soccer

Donald Bailey, Gourab Sen Gupta
School of Engineering and Advanced Technology,
Massey University, Palmerston North, New Zealand
D.G.Bailey@massey.ac.nz, g.sengupta@massey.ac.nz

Abstract

On the robot soccer field it is important to know where the camera is positioned to correct for parallax offsets resulting from the height of the robots. This paper demonstrates a novel method of estimating the camera position purely from the image of the robot soccer field. The principle is to back project the parallax resulting from the walls of the field. The lateral position of the camera may be measured with reasonable accuracy (approximately 1 cm error in our case), although the camera height is more sensitive to measurement errors, especially those resulting from specular reflection. The height of the camera was under-estimated by 11% on one field; this corresponds to a position error of less than 1 pixel when correcting a robot position's parallax anywhere on the field.

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 the house.

Robot Soccer has become popular 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-controllers, 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 of the robot soccer system as well as the 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 accurate high speed calibration of the vision system for robot soccer. In this system, all position data is obtained from a global camera (one camera per team) mounted above the centre of the table (see Figure 1). 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 patches within the image is used to estimate the location and orientation of the robot within the playing area.

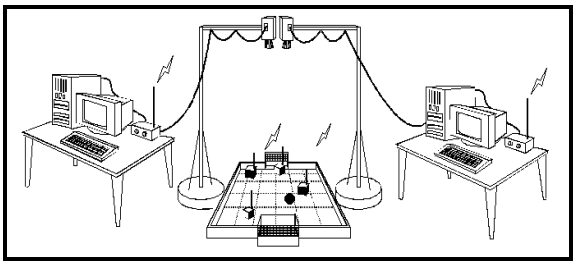

Figure 1. Setup of the robot soccer platform

To manage complexity in collaborative robot systems, a hierarchical state transition based supervisory control system has been proposed and successfully implemented [3]. 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 then no longer reliable. For this reason, accurate calibration of the imaging system is essential for accurate location and control [4].

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. The wide angle of view requires use of a short focal length camera which can introduce significant barrel distortion, as is apparent in Figure 2.

978-0-7695-3332-2/08 $25.00 © 2008 IEEE
DOI 10.1109/IMVIP.2008.7

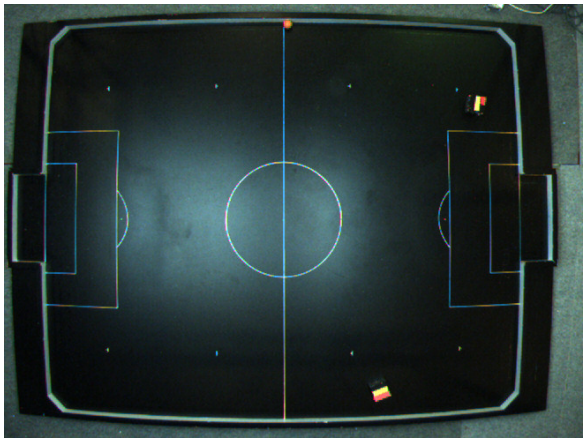

Figure 2. A typical view of a robot soccer field

Traditional camera calibration procedures require a dense set of points scattered throughout the image [5,6]. Without providing a custom target, relatively few data points are available from the robot soccer platform. The calibration procedure commonly used is to manually position the robots at a set of predefined positions around the field, and use the detected locations of these to provide the necessary calibration data. Such manual setup for calibration is both time consuming and prone to human errors.

The distortion is most obvious from the curvature of the edges of the playing area, as illustrated in Figure 2. Therefore a method based on detecting and straightening the edges of the playing area directly can be used [4]. The basic principle of this method is to detect the two ends and two sides of the playing field. Making these apparently curved edges straight will remove most of the lens distortion, and transforming the image to make these parallel will correct for perspective distortion [7].

The perspective correction transform can be used to determine the orientation of the camera (pan, tilt and roll), but not the actual camera position. Knowledge of the camera position is important to correct for parallax errors introduced by the physical height of the robots.

To the author's knowledge, no teams are currently obtaining the height and position of the camera by calibration. Most existing calibration methods are performed manually, either by manually positioning the robots, or by manually selecting the edges of the playing area within the captured images. In terms of camera location, the standard height is specified in the rules [8], and the position is either measured manually, or assumed to be over the centre of the field. In this paper, we extend our automatic calibration procedure [4] to obtain the camera position.

2. Estimation principle

In Figure 2, the white wall around the playing area is actually the inside of the wall. The field itself and the top of the wall are both painted with a matt black paint as shown in Figure 3. The fact that the wall appears white is because the camera is positioned over the field, and the inside wall is visible from the camera. As a result of parallax (the camera is not positioned directly over the wall) the top and bottom edges of each wall appear in slightly different locations within the image. These can be used to estimate the camera location.

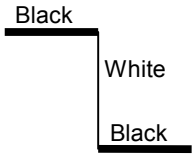

Figure 3. Inside wall is painted white

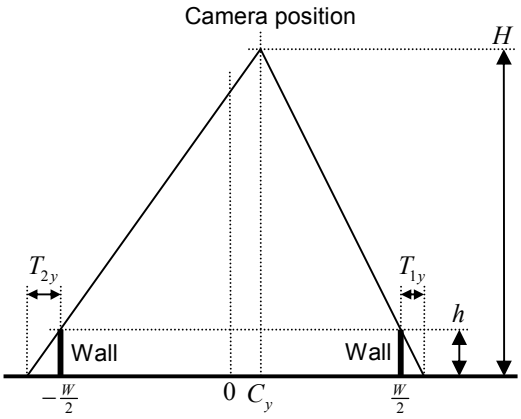

Figure 4. Geometry and definition of variables for estimating camera position

The basic principle is to back project the apparent positions of the top edges of the walls on two sides. These will intersect at the camera location, giving both the height and lateral position, as shown in Figure 4. If the surface of the field is defined as the reference height, then the image from the camera can be considered a projection of every object onto this plane. The bottom edges of the walls will appear in their true locations (neglecting image distortion for the moment), and the top edges of the walls are offset by parallax.

Let the width of the playing area be W and wall height be h. Also let the projection of the face of the two side walls onto the reference plane be T_{1y} and T_{2y}. Then the height, H, and lateral offset of the camera from the centre of the field, C_y, may be determined. From similar triangles:

$$\frac{H}{\frac{W}{2}-C_y+T_{1y}}=\frac{h}{T_{1y}} \tag{1}$$

Rearranging gives:

$$T_{1y}=\frac{h\left(\frac{W}{2}-C_y\right)}{H-h} \tag{2}$$

and similarly for the other wall

$$T_{2y}=\frac{h\left(\frac{W}{2}+C_y\right)}{H-h} \tag{3}$$

Equations (2) and (3) can be solved to give the camera location

$$C_y=\left(\frac{T_{2y}-T_{1y}}{T_{2y}+T_{1y}}\right)\frac{W}{2} \tag{4}$$

$$H=\frac{hW}{T_{1y}+T_{2y}}+h \tag{5}$$

Similar geometrical considerations may be applied along the length of the field to give

$$C_x=\left(\frac{T_{2x}-T_{1x}}{T_{2x}+T_{1x}}\right)\frac{L}{2} \tag{6}$$

$$H=\frac{hL}{T_{1x}+T_{2x}}+h \tag{7}$$

where L is the length of the playing field.

Equations (4) to (7) give four independent equations for three unknowns. Measurement limitations and noise usually result in equations (5) and (7) giving different estimates of the camera height. In such situations, it is usual to determine the output values that are most consistent with the input data. In this case, the outputs are the camera location: C_x, C_y, and H, and the measured inputs are the projected wall parallax offsets: T_{1x}, T_{2x}, T_{1y} and T_{2y}. For a given camera location, the error between the corresponding input and measurement can be obtained from (2) as

$$E_{1y}=\frac{h\left(\frac{W}{2}-C_y\right)}{H-h}-T_{1y} \tag{8}$$

and similarly for each of the other inputs. The camera location can then be chosen that minimises the total squared error

$$E^2=E_{1y}^2+E_{2y}^2+E_{1x}^2+E_{2x}^2 \tag{9}$$

This can be found by taking partial derivatives of (9) with respect to each of the camera location variables and setting the result to 0:

$$\frac{\partial E^2}{\partial C_y}=\frac{4h^2C_y}{\left(H-h\right)^2}-\frac{2h\left(T_{2y}-T_{1y}\right)}{H-h}=0 \tag{10}$$

or

$$\frac{2hC_y}{H-h}=T_{2y}-T_{1y} \tag{11}$$

Similarly

$$\frac{\partial E^2}{\partial C_x}\Rightarrow\frac{2hC_x}{H-h}=T_{2x}-T_{1x} \tag{12}$$

Taking partial derivatives with respect to the camera height is a little more complex because the H appears in the denominator of each of the terms. For example the partial derivative of the term from (8) gives

$$\begin{aligned}\frac{\partial E_{1y}^2}{\partial H}&=\frac{-2h\left(\frac{W}{2}-C_y\right)}{\left(H-h\right)^2}\left(\frac{h\left(\frac{W}{2}-C_y\right)}{H-h}-T_{1y}\right)\\&=\frac{-2h}{\left(H-h\right)^2}\left(\frac{h\left(\frac{W}{2}-C_y\right)^2}{H-h}-T_{1y}\left(\frac{W}{2}-C_y\right)\right)\end{aligned} \tag{13}$$

Combining the partial derivatives across the width of the field

$$\begin{aligned}&\frac{\partial E_y^2}{\partial H}=\frac{\partial E_{1y}^2}{\partial H}+\frac{\partial E_{2y}^2}{\partial H}=\frac{-2h}{\left(H-h\right)^2}\cdot\\&\left(\frac{2h\left(\frac{W^2}{4}+C_y^2\right)}{H-h}-\frac{W}{2}\left(T_{1y}+T_{2y}\right)-C_y\left(T_{2y}-T_{1y}\right)\right)\end{aligned} \tag{14}$$

This can be simplified by substituting (11) in the first term within the brackets:

$$\begin{aligned}&\frac{\partial E_y^2}{\partial H}=\frac{-2h}{\left(H-h\right)^2}\cdot\\&\left(\frac{\left(T_{2y}-T_{1y}\right)}{C_y}\left(\frac{W^2}{4}+C_y^2\right)-\frac{W}{2}\left(T_{1y}+T_{2y}\right)-C_y\left(T_{2y}-T_{1y}\right)\right)\\&=\frac{-2h}{\left(H-h\right)^2}\left(\frac{\left(T_{2y}-T_{1y}\right)\frac{W^2}{4}}{C_y}-\frac{W}{2}\left(T_{1y}+T_{2y}\right)\right)\\&=\frac{-h}{\left(H-h\right)^2}\left(\frac{hW^2}{H-h}-W\left(T_{1y}+T_{2y}\right)\right)\end{aligned}$$

Finally, combining the partial derivatives along the length of the field with those across the width of the field gives:

$$\begin{aligned}&\frac{\partial E^2}{\partial H}=\frac{\partial E_y^2}{\partial H}+\frac{\partial E_x^2}{\partial H}=\frac{-h}{\left(H-h\right)^2}\cdot\\&\left(\frac{hW^2}{H-h}-W\left(T_{1y}+T_{2y}\right)+\frac{hL^2}{H-h}-L\left(T_{1x}+T_{2x}\right)\right)\end{aligned} \tag{15}$$

Setting this to 0 and solving for H gives

$$H=\frac{\left(W^2+L^2\right)h}{W\left(T_{1y}+T_{2y}\right)+L\left(T_{1x}+T_{2x}\right)}+h \tag{16}$$

Finally, the result from (16) can be substituted into equations (11) and (12) to give the lateral position of the camera:

$$C_y = \frac{\frac{1}{2}\left(W^2 + L^2\right)\left(T_{2y} - T_{1y}\right)}{W\left(T_{1y} + T_{2y}\right) + L\left(T_{1x} + T_{2x}\right)} \quad (17)$$

and

$$C_x = \frac{\frac{1}{2}\left(W^2 + L^2\right)\left(T_{2x} - T_{1x}\right)}{W\left(T_{1y} + T_{2y}\right) + L\left(T_{1x} + T_{2x}\right)} \quad (18)$$

3. Measuring the wall projections

All of the equations in the previous section assume that there is no distortion within the image. In practise, the wall edges are detected within the distorted image. These are then used to first calibrate the image for lens and perspective distortion, and then using the corrected positions to determine the camera position.

3.1. Calibration

Following the general procedure of [4], the edges are detected using a 3x3 Prewitt filter. Since the projected wall only occupies 4 or 5 pixels within the image, to estimate the camera position to any reasonable accuracy it is necessary to measure the edge positions to sub-pixel accuracy. The positions of the edges are determined by finding the local minima and maxima of the filtered outputs. A parabola is fitted to the local extreme and the responses of the adjacent pixels perpendicular to the edge. The sub-pixel offset [9] of the edge is then given by

$$offset = \frac{p_+ - p_-}{4p_0 - 2(p_+ + p_-)} \quad (19)$$

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

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. A robust fitting method is used to fit parabolas to both the top and bottom edges of each wall, with the results shown in Figure 5.

The curvature of the parabolas is used to estimate the lens distortion. The effect of this is to straighten each parabola into a line [4,7]. The perspective transformation is then estimated by mapping the bottom wall edges to their known locations, and making the top wall edges parallel to their counterparts [4,7]. Figure 6 shows the corrected image.

3.2. Wall locations

In practise, it is not necessary to transform the complete image as in Figure 6. The lens distortion correction transforms each parabola into a line, and the perspective distortion correction transforms the lines to their undistorted locations.

If the origin is taken as the centre of the playing area then the projected position of the top of each wall may be determined directly from the intercept of the line equations. The initial parabola fitting process means that the results are effectively averaging the sub-pixel data along the length of the wall rather than taking a single point measurement. This will significantly improve the accuracy in spite of the limited resolution.

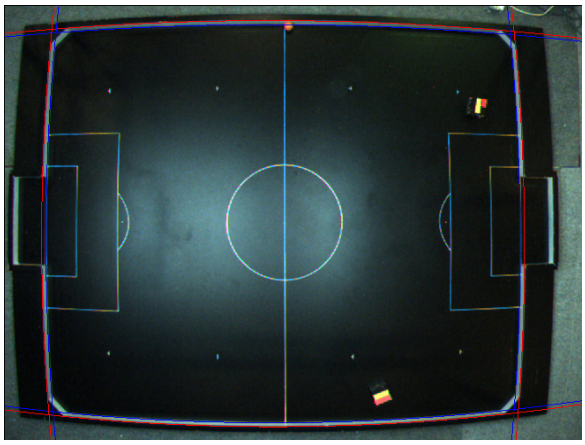

Figure 5. Detected top (red) and bottom (blue) edges of the wall in the distorted image

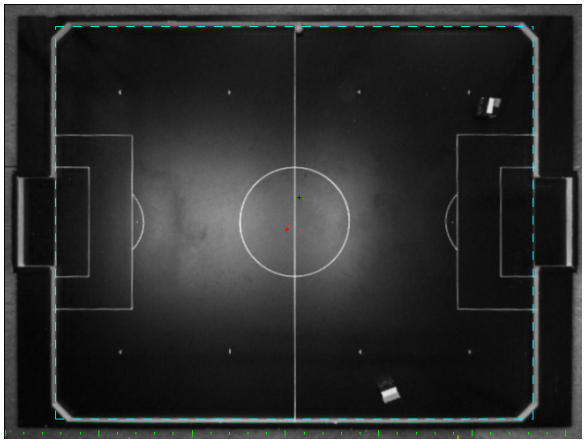

Figure 6. After correcting for distortion; the dashed line shows the calibrated locations of the bottom edge of the wall

4. Results and discussion

The camera was set up over a Mirosot (Micro-robot Soccer Tournament) middle league (with five robots

per team) playing field. For this league, the playing area measures 220 x 180 cm, and the wall height is 5 cm [8]. The true position of the camera was measured using a plumb-line suspended below the centre of the camera lens to get the lateral position. The actual camera position is given in the first column of Table 1. Results were obtained for two completely different fields: that from Figure 2, and from Figure 7.

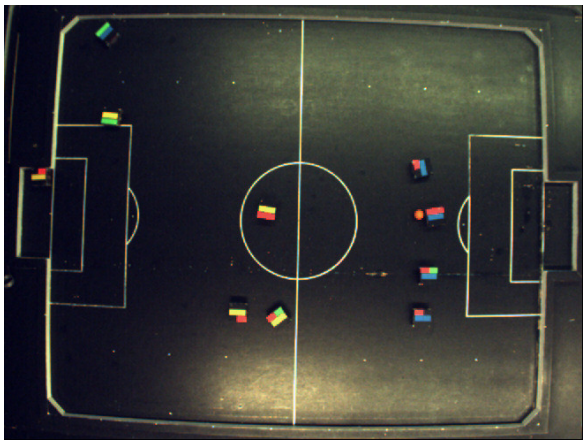

Figure 7. Image from the second field

A Basler A311FC colour firewire camera was used to capture images with a resolution of 656 x 492 pixels. This gave an image resolution of approximately 4 mm per pixel.

Applying the calibration and measurement procedure described in the previous section takes less than a second, giving fully automatic camera calibration in a fraction of the time required for manual calibration. The calculated results listed in Table 1 were obtained from the calibration. The lateral position of the camera was measured with considerable accuracy, with a total lateral error of 0.85 cm ($\sqrt{0.70^2+0.49^2}$) and 1.03 cm ($\sqrt{0.70^2+0.75^2}$) respectively for the two fields. This corresponds to an error of between 2 and 3 pixels. The error in the height, however, was significantly larger, at almost 33 cm for the image from Figure 7.

Table 1. Camera position

Field		Actual	Calculated	Error
1 (Fig. 2)	H	253.5 cm	244.86 cm	8.64 cm
	C_x	3.0 cm	3.70 cm	0.70 cm
	C_y	2.5 cm	2.01 cm	0.49 cm
2 (Fig. 7)	H	298.5 cm	265.77 cm	32.73 cm
	C_x	7.5 cm	8.20 cm	0.70 cm
	C_y	3.5 cm	2.75 cm	0.75 cm

A larger height error is not completely unexpected, because the height is estimated by back projecting the relatively small parallax resulting from a low wall. Any small errors in measuring the wall parallax are amplified to give a large error in the estimate of the camera height. The lateral position is not affected by measurement errors to the same extent, because it is based on the relative difference in parallax between the two sides.

The cause of the height errors associated with Figure 7 warrants further investigation. Table 2 compares the measured parallax associated with each of the 4 walls around the playing area. These measurements are compared with the expected parallax based on the known camera positions using equations (2) and (3), and the corresponding similar equations along the length of the field.

Table 2. Raw measurements of projected wall from Figure 7

	Expected	Calculated	Difference
T_{1x}	20.0 mm	21.6 mm	1.6 mm
T_{2x}	17.5 mm	18.4 mm	0.9 mm
T_{1y}	15.9 mm	19.1 mm	3.2 mm
T_{2y}	14.7 mm	18.1 mm	3.4 mm

It is observed that the measured parallaxes all overestimate the expected measurements, although these overestimates are all less than 1 pixel. However, because the parallaxes all appear in the denominator of equation (16), any overestimate of the parallax will result in underestimating the height of the camera.

It is interesting that the sides of the playing field (T_{1y} and T_{2y}) have a larger error than the ends of the playing field (T_{1x} and T_{2x}). This would imply that the height estimated using the sides alone, from equation (5), would be lower than the estimate using the ends alone, from equation (7). This is indeed the case, with the height estimated from the sides as 247.0 cm, compared with 280.1 cm from the ends.

Further investigation showed that one of the reasons for the overestimate of the wall parallax, particularly in Figure 7, is a slight rounding of the wall profile at the top (see Figure 8). This introduces a specular reflection component which extends over the top of the wall, increasing the apparent width of the white wall region within the image. A similar effect could also occur at the bottom edge of the wall if the join between the wall and playing area was also curved.

The narrower width of the playing area, combined with the position of the lights, meant that the specular component extended further onto the top of the wall for

the sides of the playing area than it did for the ends. This accounts for the differences observed in Table 2. The asymmetry of the lighting, as can be observed in Figure 7, also accounts for why the right end of the playing area, T_{1x}, was more affected than the other.

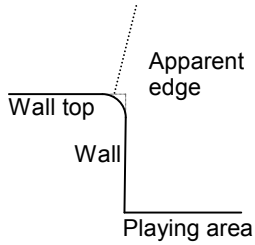

Figure 8. Movement of apparent edge position resulting from specular reflection

The first field (Figure 2) was less affected for the following reasons. Firstly, the lights were positioned over the field rather than outside it. The different light angle means that this field was less prone to the specular reflection effects on the top corner of the wall. Secondly, the field was in a better condition, with less rounding of the top edge of the walls, and a fresher coat of paint. As a result, the errors were less significant.

As the camera height is used to correct for parallax errors of the robots on the field, we can analyse the effects of both the lateral and height errors on the calculated object positions. Assume the robot has height R and is positioned at a lateral distance d from the camera. An error in estimating the height of the camera of ΔH will result in a robot position error of

$$E_R = \frac{-dR\Delta H}{(H-R)(H+\Delta H)} \qquad (20)$$

This is proportional to the lateral distance, which will be the greatest in the corners of the playing field. For a robot that is 5 cm high, the 32.7 cm error in height of the camera will result in a 3.0 mm error in correcting the parallax. This corresponds to less than one pixel error. Therefore, even though the height is significantly underestimated on the second field, it is still sufficiently accurate to correct parallax errors within the image to better than 1 pixel.

5. Conclusions

This paper has shown that it is possible to effectively estimate the position of the camera relative to a robot soccer field purely by imaging the field. Back projecting the parallax offsets of the walls enables the camera location to be determined. The lateral position may be determined with good accuracy. However, specular reflection from the top corner of the walls results in the wall parallax being over-estimated and consequently the height is being under-estimated by as much as 11%. It is shown that even such large errors are not that significant when correcting for parallax of objects on the soccer field.

6. Acknowledgements

This research was performed within the Advanced Robotics and Intelligent Control Centre (ARICC). The authors would like to acknowledge the financial support of ARICC and the Department of Electrical and Electronic Engineering at Singapore Polytechnic.

7. References

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

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

[3] G. Sen Gupta, C.H. Messom, and H.L. Sng, "State Transition Based Supervisory Control for Robot Soccer System", in *International Workshop on Electronic Design, Test and Applications (DELTA)*, Christchurch, NZ, 338-342 (29-31 January, 2002).

[4] D. Bailey and G. Sen Gupta, "Error Assessment of Robot Soccer Imaging System", in *Image and Vision Computing New Zealand (IVCNZ'04)*, Akaroa, New Zealand, 119-124 (21-23 November, 2004).

[5] R.Y. Tsai, "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).

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

[7] D.G. Bailey, "A New Approach to Lens Distortion Correction", in *Image and Vision Computing New Zealand 2002*, Auckland, New Zealand, 59-64 (26-28 November, 2002).

[8] FIRA, *FIRA MiroSot Game Rules*. Federation of International Robosoccer Association: Available from: http://www.fira.net/soccer/mirosot/MiroSot.pdf [cited 8 January, 2008].

[9] D.G. Bailey, "Sub-pixel Estimation of Local Extrema", in *Image and Vision Computing New Zealand 2003*, Palmerston North, New Zealand, 414-419 (26-28 November, 2003).