Robust Adaptive Vision for Robot Soccer

Gordon Wyeth and Ben Brown
Computer Science and Electrical Engineering,
University of Queensland, Brisbane, Australia 4072

Abstract

Robot soccer requires vision to be fast, accurate and reliable for a team to perform well. To win competitions, the vision system must work every day under all contest conditions. In practice, the field conditions for robot soccer can vary enormously from field to field, and even from day to day throughout a contest. This paper presents a system that can adapt to rapidly varying light and colouring conditions, while maintaining speed and accuracy. Quantitative results of performance are presented, and applications outside the robot soccer domain illustrated.

Keywords:
Robot soccer, vision, colour histogram.

1 Introduction

The Robot World Cup Initiative (RoboCup) is an international research initiative that uses the domain of robot soccer to foster AI and intelligent robotics research [3]. The robot soccer game is a standard problem where a wide range of technologies can be integrated and examined. International contests allow clearly measurable comparisons between technological approaches to research problems. This paper examines the global vision module employed for a team of robots, the RoboRoos [6], developed at the University of Queensland. The module was used in competition at RoboCup '99 held in Stockholm during August, 1999. The global vision module provides the robotic soccer team with:

- the position and orientation of the team robots,
- the position of the opposition robots, and
- the position and velocity of the golf ball.

The system uses a single CCD camera mounted 3m above the field which is processed on a standard PC at 25 frames per second (fps). The information gleaned from the image is analysed to develop strategic plays on the PC, and also transmitted directly to the players for navigation purposes. From a systems perspective, it is important that the information delivered to the robots from the camera is:

- reliable,
- accurate,
- delivered regularly, and
- delivered with minimal latency.

These alone would be sufficient conditions for effective operation in the laboratory, but the system must be able to travel to contests around the world and still function as it does in the lab. From this perspective, it is also important that the system is robust to:

- changes in lighting,
- changes in the colour of the playing field, and
- the many different colouring schemes used on the opposition robots.

While the rules of the contest provide some boundaries on the lighting conditions, in practice the lighting levels can vary between 400 - 1500 lux average illumination for different competitions. In addition there may be variations of up to 500 lux for different parts of the field. External factors may also causes variations to the mean illumination of up to 400 lux during the day at a given competition.

The colourings of the field and the opposition robots are also ill-defined. The field may be supplied in various shades of green. The opposition robots may be composed of many colours, provided that they do not clash with the blue and yellow marker colours (described later), or with the orange of the golf ball that is used in the contest. Again, the extent to which teams avoid the key colours varies, and the total spectrum of colours used can be wide.

This paper describes how the RoboRoos vision system deals with such large variations in colouration and lighting without resorting to complex and often fragile calibration schemes. The following section describes the complete RoboRoos soccer system in further detail, and describes the problematic approach to vision that we adopted in 1998. Section 3 explains the new system developed in 1999. Section 4 shows results that illustrate the system's accuracy, reliability, frame rate and robustness. Section 5 draws conclusions that illustrate the application of this system in other robotic domains.

2 The Problem in Detail

The RoboRoos global vision module must supply information for two other modules in the soccer system. MAPS, the Multi-Agent Planning System [4], is the strategic core of the RoboRoos system. It is here that all of the long-term tactical decisions are made, such as where each robot should be on the field, the number of attackers and defenders, and the ideal methods of attacking the goal. Once the instructions for each robot have been determined by MAPS, these instructions are passed on to the individual robots. The robots use these instructions to provide goal and parameter information for the on-board navigation system.

The global vision module also provides information directly to the navigation system for the robots. The robots use the information given for position and orientation to maintain their own internal position reference. The robots integrate their own position and orientation based on internal encoders. The information from the vision system re-calibrates the integrated position and orientation information. Given that the robots have an internal reference, it is more important the information from the vision system is accurate and reliable rather than frequent. For this reason, the vision module will only transmit position and orientation data that is certainly reliable.

The robots use the information given with regard to the positions of the opposition robots for obstacle avoidance. Failing to see the opposition can lead to fouls being awarded against the robots for charging. Serious charges can be penalised with red cards, leading to the robot being removed from the field. It is imperative, therefore, to maintain good information with regard to the opposition so that effective obstacle avoidance is achieved. Similarly, obstacle avoidance is performed with regard to robots on the same team.

The robots require the position of the ball and its velocity to perform interception of the moving ball. As the match is placed on a hard painted surface, the ball is rarely at rest. It is essential to be able to play at a moving ball to achieve good results. Prediction of the ball's future position is best achieved based on accumulated position and velocity information.

2.1 Input Specifications

The global vision system obtains information about the world through a Pulnix TMC-6 CCD colour camera. The camera has a variable length lens that is set to the field length on the horizontal axis. The PAL video signal is translated to RGB values by a Sensoray 611 frame grabber card. This card performs a 2:1 sub-sampling of the image and places the RGB values in the main memory of the PC over the PCI bus. The card also crops the image to the height of the field. This provides 376 pixels for the length of the field, and 217 pixels for the height of the field, translating to approximately 7mm per pixel. A sample image is shown in Figure 1.

2.2 Output Specifications

As stated earlier, the global vision module provides the robotic soccer team with:

- the position and orientation of the team robots,
- the position of the opposition robots, and
- the position and velocity of the golf ball.

The output generated from the scene shown in Figure 1 is shown in Figure 2.

2.3 Previous Approaches

The RoboRoo’s initial vision system relied extensively on colour in its attempts to track the robots. For RoboCup '98, each robot had two coloured balls mounted on top of it, in a specific orientation. This system made use of these balls to determine the orientation of the robots, and each robot was tagged at the start of play and tracked around the field.

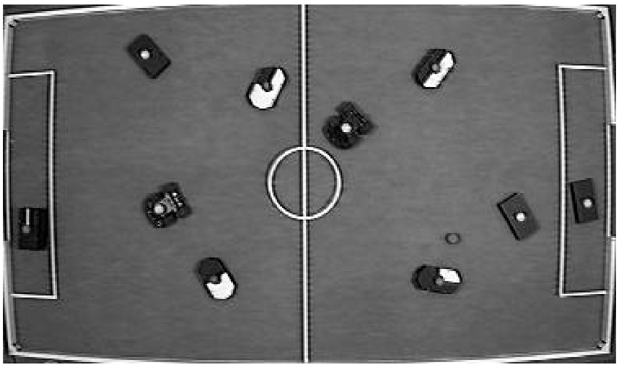

Figure 1. A typical input scene to be analysed by the global vision module. The robots with black and white markings are the RoboRoos; the plain black "robots" are simulated opposition. The robot to the far left is the RoboRoos goalkeeper which is treated separately.

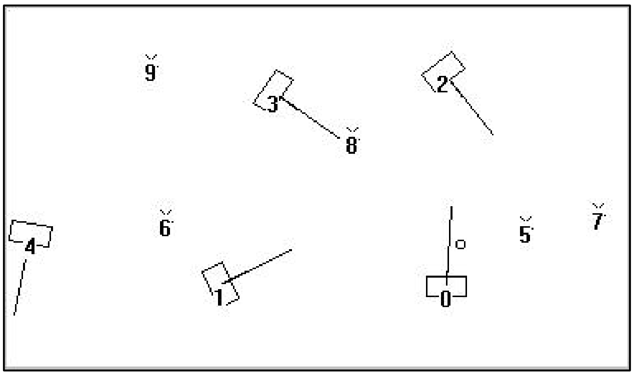

Figure 2: The output from the global vision module. This shows the five RoboRoos players with position, orientation and identification number. The five opposition robots are shown as crosses with assigned identification. The ball is shown as the small circle above robot 0.

To recognise the two balls on top of each robot, a user was required to manually enter the red, green and blue values of each colour ball to be matched. This meant that as the image's brightness changed, the values for each colour would become incorrect. Various methods of normalisation of illumination were tested, but the results were unreliable.

Based on the premise that it was too difficult to search the image for the marker balls reliably, a second system was developed that used background subtraction to initially find the robots [5]. The system was based on capturing an image of an empty field and then subtracting that base image from the current image to find the robots on the field. This approach was prone to error from lighting change or movement of the overhead camera with respect to the field.

Both of these systems were also flawed by the need to track the robots around the field after initially tagging them, rather than being able to identify them from a single frame. This meant that once a robot was 'lost' it was unlikely to be recovered until the next stop in play.

3 The Current Solution

The solution to the vision problem works in a number of stages. First the areas of interest in the image are isolated and segmented. Each of these areas is then deemed to be either one of the RoboRoos, one of the opposition, or an unknown item. Then if it is one of our robots, information regarding orientation and identification are extracted. Each of the phases in the solution are elaborated below.

3.1 Removal of the Background

The first step in the vision system is the removal of the background playing field, or conversely, the highlighting of points of interest. To do this, a method of adaptive colour histogramming was developed. This method relies on the fact that the field is, with the exception of marking lines, a consistent colour. Three separate histograms are created from an image of the field; a red, green and blue histogram corresponding to each of the component colours in the image. The histograms can be used to determine the levels of red, green and blue which are most common. The peak levels correspond to the colour of the playing field, which takes up the bulk of the image. Using these peaks as a starting point, maximum and minimum limits are grown outwards until the bin counts drop away below a preset minimum. In this way a cube is defined in RGB colour space that contains the pixel values of the field. The histograms for the image shown in Figure 1 are given in Figure 3.

The cube is considered to be the bounds within which field pixels will occur. Any pixels which fall outside the cube is then deemed to be a location of interest. By using three single variable histograms rather than a single multi-variable histogram (such as [2]), computational requirements are significantly reduced. Savings are notable in the formation of the histograms, the formation of the cluster representing the background pixels, and the determination of whether a pixel inside or outside the cluster. As the histogram and cube boundaries can be created rapidly, it can be reassessed every frame. This means that the system can respond reliably even in the presence of transient changes in lighting such as camera flashes from spectators photographing a contest.

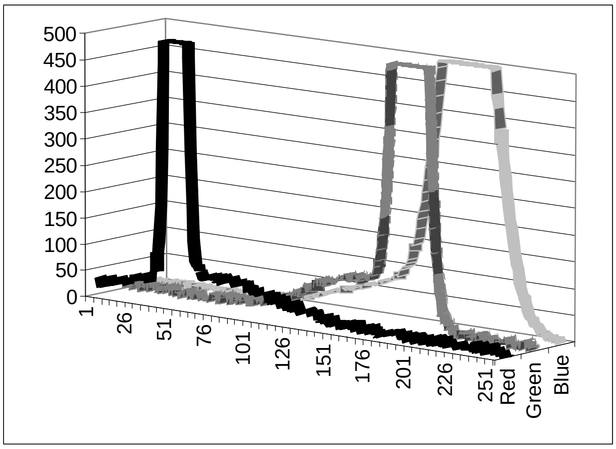

Figure 3. The histograms for the image in Figure 1.

Naturally the resulting image has some noise, including the lines on the field, shadows from field walls and inaccuracies in the determination of minimum and maximum bounds on the field colour. Noise is removed by a two pass thinning operation, where thinning is performed on pixels with less than 6 of their 8-neighbours present.

The robots have coloured markers on the viewable surface, and those markers may be similar to the field (under some forms of illumination). In such a case, the pixels from the marker may be inside the adaptive histogram cube, creating a hole in the blob of pixels representing the robot. A two pass dilation operation is run after thinning, where all 8-neighbours of the remaining pixels are replaced. This fills any holes in the blobs, and replaces pixel mass lost during the thinning operation. Figure 4 shows the final output from the background removal process.

3.2 Blob Formation and Team Determination

Pixel aggregation is straightforward, since it is operating on a binary thresholded image, and the only joining condition is proximity to another member-pixel [7]. Aggregated pixels, or blobs, are assigned indices and are checked against minimum mass constraints to eliminate any noise blobs, and against maximum mass constraints to indicate collided robots. Minimum mass blobs are tested as the ball, which is identified by its colour and accurately positioned by a template match. Its velocity is determined by filtering the position difference between frames.

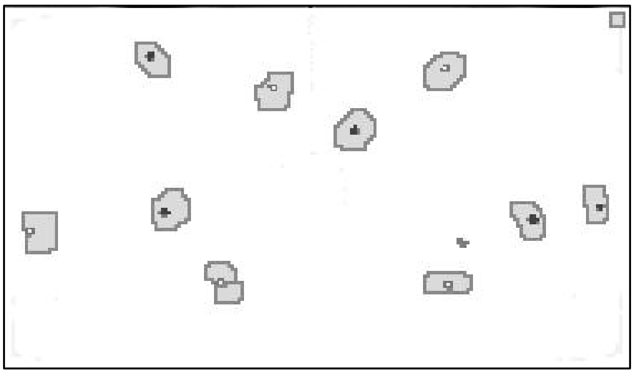

Figure 4. Result of the background removal operation. The white area was removed by the initial thresholding using the adaptive colour cube or by thinning. The dark grey areas are pixels reformed by dilation. The pale grey pixels are pixels outside the adaptive colour cube that have not undergone any morphological operation. Pixels of other shades have been highlighted as seed pixels for later processes.

For blobs of valid size, it remains to determine if the blob is one of the RoboRoos robots, or one of the opposition. This is done by using the marking which are required by the RoboCup rules, namely either a blue or a yellow ping-pong ball on the centre of each robot. To determine the identification a measure called blueness was created:

$$Blueness = B - \frac{R+G}{2} \tag{3.2.1}$$

where *R*, *G* and *B* are the red, green and blue pixel values respectively. Positive blueness indicates the presence of a blue marker, negative blueness indicates a yellow marker. A neutral blueness indicator suggests a non-marker pixel. The sign of this measure is illumination independent.

During background removal pixels that were outside the colour cube were tested for blueness and stored as seed pixels. To determine team identification the segment is scanned from the centroid out, searching for seed pixels of appropriate blueness. When a seed pixel is found, a template matching procedure is performed around the region of the seed pixel to accurately determine the position of the marker. This position is then reported as the centre of the robot. Based on the whether the matched template had positive or negative blueness, and the colour of the RoboRoos for that match, the segment is then marked as a RoboRoo or as one of the opposition. If an unsatisfactory match is performed, the segment is not reported.

In the case of segments that exceed the maximum mass, it is assumed that the segment will contain more than one robot. Such segments are searched exhaustively for markers, which are used to indicate the position of obstacles for navigation. The next step, orientation and member identification, is not performed on collided robots.

3.3 Orientation and Member Identification

Determination of the major axis is performed only on those robots which have been identified as one of the RoboRoos. Orientation is required to re-calibrate the on-board integration for navigation, and to assist MAPS in determining robot headings to produce efficient passes. It is also used later in the vision system to perform team member identification. The orientation is derived from the major axis, which may be found using moment calculations. The centroid $(\bar{x}, \bar{y})$ is pre-calculated in blob formation process. The relevant moments of the segment are then calculated using the following equation [1]:

$$m_{jk} = \sum_{x=x_{\min}}^{x_{\max}} \sum_{y=y_{\min}}^{y_{\max}} (x-\bar{x})^j (y-\bar{y})^k f(x,y) \qquad (3.3.1)$$

where $f(x, y)$ is 1 if the pixel is in the segment, and 0 if it is not. x_{min}, x_{max}, y_{min} and y_{max} are the extents of the blob determined during blob formation. The major axis of the segment, q, is calculated using the following equation:

$$\tan(2q) = \frac{2m_{11}}{m_{20} - m_{02}} \qquad (3.3.2)$$

Once the major axis of the robot has been determined, it is known that the forward direction of the robot is at 90° to the major axis, but it could be either of two ways. By using a marking scheme on top of the robots, both the orientation and identification of the robot can be determined using the centre of the robot and the alignment of the major axis. The markings used are shown below in Figure 5.

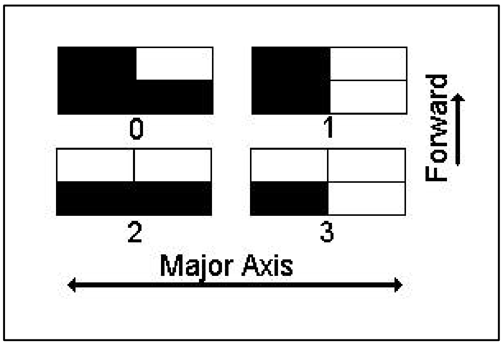

Figure 5: The robot markings provide a robust method of determining robot identification and orientation.

To determine the forward direction of the robot, the colour of the region at an angle of 45° and 225° from the major axis is inspected. One of these regions will be white, and one will be black. The white region is considered to be at the front of the robot. Once the forward orientation of the robot has been determined, the colour of regions at angles of 45° and 225° from the forward orientation are inspected. These regions can have any of four combinations of black and white, giving each a unique identity. If the system cannot determine the forward direction of the blob, or its identification conditions are not met, the blob is not reported. If all the regions are black, the robot is tested as the goalkeeper, which is distinguishable by a small white line from the marker ball. Moment analysis of this high contrast line provides excellent orientation accuracy required for effective goal keeping. After this process all information is available for reporting to the respective systems. Typical output is shown in Figure 2.

4 Results

The effectiveness of the module can be quantitatively determined with respect to the criteria stated earlier, namely:

- accuracy,
- reliability,
- frame rate and latency,
- robustness to changes in lighting, and
- robustness to opposition colouring schemes.
-

4.1 Accuracy

The module accurately determines the position of the marker in the field, as shown by Figure 6. Here, accuracy refers to the ability of the module to find the pixel that corresponds to the centre of the marker, which is fundamentally a function of the template matching of the robot marker. There are non-linear lens effects that cause the transformation from pixel coordinates to world coordinates to lose accuracy. In practice, these non-linearities have little impact on the function of the total system.

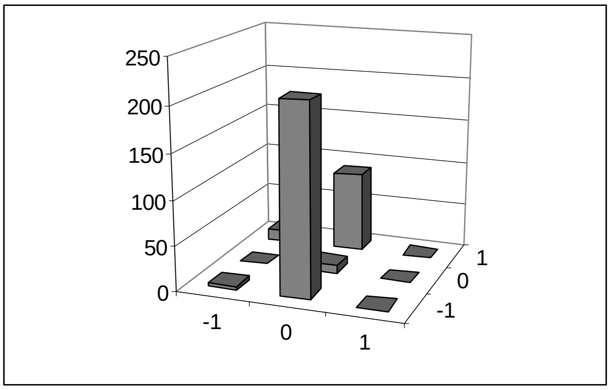

Figure 6: Graph showing typical position accuracy in the x and y directions. No pixels were had error of greater than one in this example.

The module also determines the orientation information for the RoboRoos robots based on the moments of the blob. This measure is somewhat noisier as is shown by Figure 7. Typically noise of ±7 degrees is present in orientation. In some situations, particularly when a line marking gets partially included in the robot's blob accuracy deteriorates to ±15 degrees. The orientation information is filtered by the robots to eliminate some of this orientation noise.

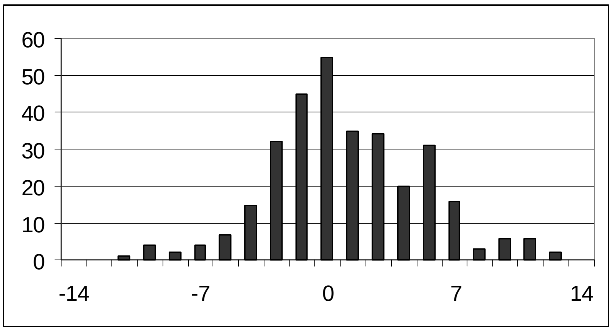

Figure 7. Graph showing typical orientation accuracy. In this example most readings are between ± 7 degrees, although in some cases the error can reach ± 15 degrees.

4.2 Reliability

The reliability of the module is best measured by the number of false-positive identifications made by the system. For all independent testing done on the module, and the many hours of development and competition, the number of false-positive identifications is negligible, if not zero. The module develops high reliability by the number of checks that are performed on each blob. A blob that fails to meet any of the checks will not be reported, effectively eliminating false positive identifications.

The effect of such stringent acceptance policies can lead to a low number of objects being reported. This becomes a function of robustness also (see below). Under typical operating conditions, the system reports 90% of the objects on the field.

4.3 Frame Rate and Latency

The adaptive histogramming algorithm used to perform the initial processing of the image is the only process that must process every pixel. After this the processing becomes local to regions of interest. The template match of the markers is the most intensive piece of operation, but it only has to operate on a small number of pixels.

The vision module runs at 25 frames per second on a 400 MHz Pentium III processor. The module processes one frame as the frame grabber grabs the next frame to be processed, and the communications system transmits the information gleaned from the last frame. This creates a total latency between the event and the reception of the information on the robots in the order of 120 ms.

4.4 Robustness

One of the primary considerations in the design of the RoboRoos global vision system was that it be robust to changes in lighting. Table 1 summarises tests in the lab involving marked changes in the lighting conditions of the field.

Spotlights	*Illumination (lux)*	*Reliability (%)*	*Effective Rate (Hz)*
2	500-800	89.4	22.4
1	200-700	73.7	18.4
0	80-120	54.5	13.6

Table 1. The performance of the system with varying lighting conditions. No false positives were reported.

In the lab, the system is set in a room with typical fluorescent lighting supplemented with two 500 W spotlights that are situated to provide even coverage across the field surface. This gives illumination levels at around 1000 lux for the "hot spots" of the lights, falling off to 800 lux near the corners of the field. With the frame grabber gains set to give a histogram distribution close to that shown in Figure 3, the system reports an average of 89.4% of objects per frame. This means that each objects position will be updated at an effective rate of 22.4 Hz.

When a single spot-light is extinguished, the illumination drops to under 500 lux at the dim end of the field. With no adjustment of frame grabber gain, the system continues to function, although the number of objects reported is now reduced to 73.7% per frame (18.4 Hz). With both lights extinguished, the field illumination drops to between 100 and 300 lux, and the number of objects reported drops to 54.5% (13.6 Hz). The grabber gains were then adjusted to similar levels to that shown in Figure 3, although the distributions were naturally somewhat different. The system reported an 80.1% recognition rate (20 Hz). In all tests, the number of false identifications was zero.

5 Discussion

The global vision module described here performs all of the functions required of it with excellent results. The only desirable improvement is on the accuracy of the orientation information generated from the moment calculation. The spread of ± 7 degrees injects some noise on to the navigation system that slows the robots. Accuracy could potentially be increased using the edges formed by boundaries between black and white on the robot covers.

The aim of the robot soccer concept is to bring the technology developments from the contest into real world applications. One can envisage applications where the complete system presented could be used; processing images from a high vantage point camera used in field robot control, for example. Alternatively, one can see applications of the components of the system. The adaptive colour histogram is a powerful thresholding tool for highlighting objects of interest in a colour image that has a large number of uninteresting background pixels that share colour attributes. As such, it has applications in object detection, region segmentation, enhancement, and target tracking. The process is highly efficient compared to 3D colour clustering, with efficiency gains in histogram formation, cluster determination and thresholding.

From a robot systems perspective, the module illustrates the importance of reliability in a sensor. As the system rejects incorrect analysis of image data, the robots can use the vision information as a trusted source for strategy analysis and navigation. A robot sensor is better to say nothing at all, than to provide false information. This was witnessed in robot soccer contests where robots from other teams would charge off after phantom balls, or balk at phantom opponents.

Further enhancements are planned for the RoboRoos vision system in 2000, with plans to implement the vision system on a "smart camera" that has a high speed DSP operating with a CMOS image sensor. This setup will reduce latency to 40 ms and increase the frame rate to 50 Hz. The smart camera will be feature further adaptive capabilities, with the ability to adjust integration periods "on-the-fly" based on the shape of the colour histogram.

References

1. K.S. Fu, R.C Gonzalez and C.S.G. Lee, *Robotics: Control, Sensing, Vision and Intelligence*. McGraw-Hill, pp. 407-408, 1987.
2. K.S. Fu and J.K. Mui "A Survey of Image Segmentation", *Pattern Recognition*. Volume 13, Number 1, pp. 3-16, 1981.
3. H. Kitano, "Research Program of RoboCup", *Applied Artificial Intelligence*, Volume 12, Number 2-3, 1998.
4. A. Tews and G. Wyeth "MAPS: A System for Multi-Agent Coordination". *Advanced Robotics*, VSP / Robotics Society of Japan, accepted for publication.
5. G. Wyeth, B. Browning and A. Tews, "UQ RoboRoos: Preliminary Design of a Robot Soccer Team". *Lecture Notes in AI: RoboCup-98 - Robot Soccer World Cup II*, Springer-Verlag, 1604, 1999.
6. G. Wyeth, B. Browning and A. Tews, "UQ RoboRoos: Ongoing Design of a Robot Soccer Team", *RoboCup-99 Team Descriptions*. Linköping University Electronic Press, pp. 77-80, 1999.
7. S. W. Zucker "Region Growing: Childhood and Adolescence," *Computer Graphics and Image Processing*, Volume 5, pp. 382-399, 1976.