

An Automated Colour Calibration System using Multivariate Gaussian Mixtures to Segment HSI Colour Space

Naomi Henderson, Robert King, Stephan K. Chalup

The University of Newcastle

Australia

{naomi.henderson, robert.king, stephan.chalup}@newcastle.edu.au

Abstract

This paper presents a system for automating the time consuming task of manual colour calibration for a mobile robot. By converting a series of YUV images to HSI format and analysing histogram data it can be seen that there are distinct regions of colour space for each object colour class and that one dimension, hue, can be used to uniquely identify each colour class. Using an expectation maximisation (EM) algorithm to estimate the parameters of a Gaussian mixture model, it is proposed that the HSI colour space can be segmented and automatically labeled for the purpose of automatic colour calibration. This method is applied to a Aldebaran Nao robot vision system that uses a ‘soft’ colour classification method to classify non-unique colour space. By reducing the colour labeling dimension to one and implementing soft classification principles, a reliable automatic calibration system was achieved.

1 Introduction

In many RoboCup Leagues robot vision is used as the primary source of information about the environment. When a robot operates in a colour coded environment, such as the Standard Platforms League, colour classification is used to extract information from images. Colour classification is the process of mapping input image pixels to a colour code stored in a colour table. Manual calibration of this table involves hours of manual pixel mapping from test images to colour classes. A change of lighting conditions shifts image values, hence at each competition venue the colour table must be recalibrated.

There is an additional problem with calibration of how to classify pixel values that are non unique to an object. That is, shades of colour that occur elsewhere in the environment. To address this problem our vision system

uses a soft colour classification system [Quinlan *et al.*, 2005] with additional classes to classify these shades. An overview of the soft colour classification system is given in Section 2.1.

To automate colour calibration a source of additional information is required to label colour classes. Although there are methods for automation (Section 2.2), most are unreliable for use in competition and manual calibration is most commonly used. We propose that by transforming selected images from YUV to HSI format (Section 2.3) we can use the characteristics of the HSI colour space as the additional information required to automatically label colour classes. Using HSI sample data we applied an expectation maximisation (EM) algorithm [Bishop, 2006] [Dempster *et al.*, 1977] to estimate the parameters of a multivariate Gaussian mixture model to automatically segment the HSI colour space into colour regions. These regions can be automatically labeled using the known order of hue colour values. The advantage of using the HSI colour space is that there is only one dimension that the boundaries are required to be highly accurate. The advantage of using soft colour classification is that it reduces the accuracy required of these colour class boundaries. This method offers a practical solution to the time consuming task of manual calibration and is detailed in Section 3.

The resulting multivariate Gaussian mixture model and colour table are outlined in Section 4 and future work is discussed in Section 5.

2 Background

2.1 Colour Classification and Soft Colour Classification

The first process of the Nao robot vision system is to colour classify the image. This involves creating a colour classified image by mapping the input YUV image pixels to a colour class stored in a pre-defined colour table. Colour classes are based on the colour coded environment. Red, orange, yellow, green, blue and white are all colour codes used in the environment. When a change in

lighting conditions occurs image pixel values shift breaking the map. Hence a new colour table is required at each venue.

There is the additional problem of how to classify overlapping colour classes. That is, colour space that is non unique to an object. It can be seen in any colour space (eg. YUV, HSI, RGB) that certain colours are close and hence more difficult to separate. Due to the fixed position of the camera, the mobile nature of the robot and the dynamic objects on the field, varied lighting conditions occur that cause these close colours to overlap. A main concern is the separation of red robot uniforms and the orange ball, the orange ball and yellow goal, and the blue goal and blue robot uniforms.

The vision system uses a soft colour classification system to account for this overlap. Additional ‘soft colour’ classes were added to the colour table to classify the non unique, shared pixel values. These soft colours delay the decision making process of which colour class the pixel belongs to until the entire image is processed. Based on the spatial relationship to neighboring colour classes a decision can be made as to what object the soft colour class belongs to.

2.2 Previous Methods of Automation

The main problem with automating colour calibration is finding the additional information required to automatically apply a colour label. Previous methods use information such as the known geometry of the environment or a manually calibrated base table. Using the geometry a knowledge based classification system can be used to dynamically build a probabilistic colour model by growing and matching colour regions to the expected size and shape of objects [Cameron and Barnes, 2004][Gunnarsson *et al.*, 2006]. Another similar approach uses colour learning with a known walk routine to collect information from the known geometry and colour codes of the environment to self calibrate a colour table [Sridharan, 2005]. Other machine learning techniques that have been used include neural networks to recognise the valleys of hue histograms and determine meaningful hue ranges [Amoroso *et al.*, 2000] and Support Vector Machines (SVMs) to fill a manually calibrated base colour table [Quinlan, *et al.*, 2003].

However many of these systems are designed for the future purpose of an on board self calibrating system and for use in dynamic lighting situations. Most systems are unreliable for use in competition and manual calibration is the most common method used.

2.3 YUV and HSI Colour Space

Input images are from a Nao robot 160*120 CMOS camera and are of YUV format. The YUV colour space is represented by intensity (Y) and colour components (U

and V) which represent the blue chrominance and red chrominance respectively. Whilst the YUV colour space is convenient format for the transmission of video data, for classification and segmentation it suffers from being non-intuitive and sensitive to changes in lighting.

The HSI colour space (Figure 1a) is represented by; hue (H) an angular value that represents the predominant wavelength of colour (Figure 1b), saturation (S) a value for the amount of colour present and intensity (I) the darkness/brightness of the colour. By transforming the YUV image to HSI we can use hue to uniquely identify colour segments and reduce the critical colour dimension from two to one.

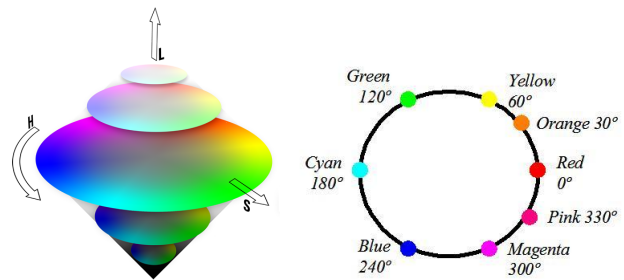


Figure 1: (a)The HSI colour space [Sande, 2005] (b) Hue and colour locations of the HSI colour space

2.4 Gaussian Mixture Models

Gaussian Mixture Models are a statistical clustering method consisting of a multivariate probability density function composed of a number of multivariate Gaussian components. Such models are used for a wide variety of classification purposes. Image processing applications include face colour recognition [Greenspan *et al.*, 2001] and texture classification [Raghu and Yegnanarayana, 1998]. Applications in robot vision include segmenting in YUV [Cohen *et al.*, 2003], colour learning in HSI [Soest *et al.*, 2006] and probabilistic visual learning for object recognition [Moghaddam *et al.*, 1997].

3 Method

3.1 HSI Data

Manually selected images were used to collect HSI pixel data. An image of each coloured object was selected in ideal lighting conditions. Only one image of each object is used to minimise values that overlap and assist fitting the multivariate Gaussian mixture. Images that fill the majority of the frame were chosen to reduce noise (Figure 2). The multivariate Gaussian mixture was applied to the collected images, rather than individual images for ease of automation. The aim is that in future the robot

will use random images of the environment rather than selected images.

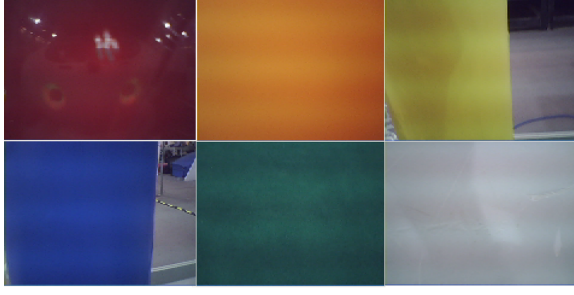


Figure 2: Input images used for HSI histogram data

3.2 Multivariate Gaussian Mixtures

Applying multivariate Gaussian mixtures to HSI data is an extension of previous work that used Gaussian mixtures in one dimension to segment hue for a Sony Aibo vision system [Henderson *et al.*, 2008]. The Aibo vision system used a very poor quality 208*160 CMOS camera. This method automated the separation of hue values but relied upon manually tuned upper and lower boundaries for saturation and intensity channels. The method proved to be effective but not optimal. For optimal classified regions and the purpose of full automation the system was extended to multivariate Gaussian mixtures.

Using the HSI colour space we gain additional colour label knowledge and reduce the ‘colour decision’ dimension to one. When applying the multivariate Gaussian mixture the characteristics of the HSI colour space can be used. Although colour temperature and brightness effects the perception of colours, the theoretical range of hues for each colour and the sequence of hue colour labels (Figure 1b) can be used to identify colour classes.

Hue is the only dimension that must be unique and the colour labels are constrained to the order: red, orange, yellow, green and blue. White and black are constrained to low saturation segments. White is constrained to high intensity and black is constrained to low intensity. Any order and overlap of the other colour classes is allowed for saturation and intensity segments.

The Gaussian Model Applied to HSI data components is shown in Equation 1.

$$\phi \left(\begin{bmatrix} h \\ s \\ i \end{bmatrix} \right) = (2\pi)^{\frac{n}{2}} |V|^{\frac{1}{2}} e^{-(HSI_{diff})V^{-1}(HSI_{diff})^T} \quad (1)$$

Where:

ϕ is probability density function of a multivariate normal distribution

n is 3, the number of dimensions

μ_c is the mean of the colour space component c

c is the colour space component hue, saturation or intensity.

HSI_{diff} is given by:

$$HSI_{diff} = \begin{bmatrix} h \\ s \\ i \end{bmatrix} - \begin{bmatrix} \mu_h \\ \mu_s \\ \mu_i \end{bmatrix} \quad (2)$$

V is the variance-covariance matrix given by:

$$V = \begin{bmatrix} \sigma_h^2 & \rho_{hs}\sigma_h\sigma_s & \rho_{hi}\sigma_h\sigma_i \\ \rho_{hs}\sigma_h\sigma_s & \sigma_s^2 & \rho_{si}\sigma_s\sigma_i \\ \rho_{hi}\sigma_h\sigma_i & \rho_{si}\sigma_s\sigma_i & \sigma_i^2 \end{bmatrix} \quad (3)$$

3.3 An Adapted Expectation Maximisation (EM) Algorithm

We adapted the EM algorithm to estimate a multivariate Gaussian distribution in the HSI colour space for each colour class. The algorithm was initialised at mean values known to correspond to the field colours, with moderate to large variances. In the first iteration of the algorithm the expectation step was replaced with the use of initial values.

The algorithm involved an expectation step and a maximisation step. The expectation step (Equation 4) calculated a weighted mean and variance-covariance matrix for each colour class. The maximisation step (Equation 5) allocated sample pixels to colour classes proportional to the ratio of the multivariate Gaussian probability density function (*mvnpdf*).

The process was iterated until $\Delta\mu$ converged to a threshold value. An *mvnpdf* of colour classes was created in the HSI colour space. Rather than creating one *mvnpdf* for the maximum probable colour class, a *mvnpdf* was created for each colour and weighted proportionally. This was for the purpose of applying soft colour principles to pixel values of close probability which is discussed in Section 3.4.

$j = 1 \rightarrow J$
 $k = 1 \rightarrow K$

$$\Pi_j = \phi \left(\begin{bmatrix} h_j \\ s_j \\ i_j \end{bmatrix} \right) \quad (4)$$

$$\Pi_{kj} = \frac{\phi_{kj}}{\sum_{k=1}^K \phi_{kj}} \quad (5)$$

Where:

Π is the weighting of colour k at j

j = sample pixel number

J = number of sample pixels

k = colour class

K = number of colour classes

c = h, s, i value

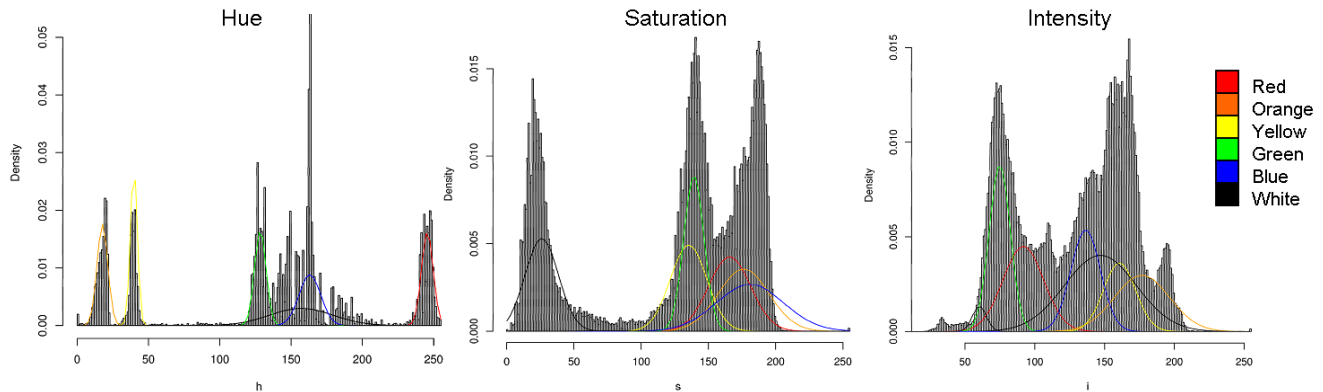


Figure 3: Histogram of H, S and I input data and resulting univariate Gaussian mixture model for each colour class.

3.4 Building the Colour Table

For each pixel value in HSI colour space the multivariate normal probability density function (*mvnpdf*) for each colour class is compared. A classification rule based on one or two standard deviations from the mean is used to label the most probable colour class. The number of standard deviations is dependent on the characteristics of the histogram. Two standard deviations was chosen to classify green and white, one standard deviation is used for more important object colours red, orange, yellow and blue. This can be varied depending on lighting conditions and image quality. To apply soft colour classification principles, overlapping probable colours within two standard deviations of colour classes are classified as soft colour.

Soft colour classes were added to the table by comparing probabilities within two standard deviations of neighboring colour means. If the probability of both colour classes are within a probability region the colour is classed as the corresponding soft colour class. For example if the probability for red and orange at a particular point is within two standard deviations of each class, the point will be classified as the soft colour class ‘red-orange’. This principal is applied to colour regions that tend to overlap. Blue is an exception, HSI pixels between one and two standard deviations from the mean of blue is classed as the soft colour ‘shadow blue’.

Once the HSI table is calibrated, a YUV table must be generated. All coordinates of the YUV colour space are transformed to a HSI value and used to reference the HSI colour table. The colour class found in the HSI colour table is stored for each YUV coordinate.

For each new set of HSI sample data and colour table, the weightings and cut off probabilities are recalculated. Therefore the images and number of images can vary.

By building a YUV colour table offline no additional processing was required on board the robot, making it

practical for use in competition. However, given sufficient processing power this system could be fully implemented on board a robot for a self calibrating system.

3.5 Implementation

The implementation of the system involved programing the *mvnpdf* and adapted EM algorithm using the package mvtnorm [Genz *et al.*, 2007] in R [R, 2008]. The weighted mean and variance-covariance matrix calculated in R was used to calculate the *mvnpdf* for each colour class in Matlab. The HSI colour table was built in Matlab using the classification rule and soft colour classification principles described in Section 3.4. The generation of the YUV colour table was programmed in C++ into the vision debug application. When prompted, the vision application uses the HSI colour table to generate a YUV colour table for use on board the robot. The colour table is modifiable using the vision debug application.

4 Results

The histogram data and resulting univariate Gaussian mixtures for the hue, saturation and intensity channels are displayed in Figure 3. For each channel, the Gaussian components for each colour class are displayed. Manual calibration involves classifying pixel by pixel, this leads to sparse colour tables and possible ‘holes’. Applying the multivariate Gaussian mixtures to segment the HSI colour space resulted in 3D solid regions and ‘solid’ classification. Whilst manual classification is reliable, solid classification has the advantage of classifying additional information that is lost with holes. This improves size information and reliability of object recognition.

The automated colour table was compared to a manual colour table using a stream of 275 images taken at RoboCup 2008. All 275 images were used to calibrate the manual colour table. The manual colour table used in competition is used here for comparison. 6 selective

	Auto			Manual		
	Blb	Obj	F+ve	Blb	Obj	F+ve
Orange ball %	100	83	0	100	93	0
Yellow goal %	100	94	0	96	94	0
Blue goal%	100	90	0	99	95	0
Blue robot%	100	0	100	100	0	50

Table 1: Comparison of performance between the HSI automated colour table and manual colour table. Blb represents the blob formation rate in images with the object present, Obj represents the percentage of objects recognised, F+ve represents false positive objects recognised.

images from the image stream (Figure 2) were used to calibrate the automated table. Images from the stream and the resulting classified images and blobs formed for both methods are shown in Figure 4.

Table 1 shows a quantitative measurement of the results. It is difficult to measure how well an image is classified. Blob formation is used as a measure of classification. If pixels are classified as a colour class it will be grouped into a blob, therefore the formation of blobs is our main measure of classification.

There are many variables in blob formation such as size, the number of pixels included etc. The object recognition process analyses blob variables when making a decision about the reliability of the colour information. Therefore object recognition is used as an additional measurement of the reliability of blobs formed. However, object recognition is manually tunable and was written for the manual lut. The rate of objects recognised and false positives can be altered by modifying the object recognition code.

4.1 Robust Classification

It can be seen from the images of Figure 4 that the classified images are of similar reliability to the manual calibrated table. This is confirmed in Table 1. The test images show a variety of distances of objects that cause the object to reflect light differently. The limitations of the automated colour table can be seen in the fifth set of images. The boundary of orange classification does not include high enough intensities to fully classify the more

distant ball.

Classifying in solid HSI regions resulted in a larger number of classified pixels and ‘solid’ classified images. Limiting the hue range allows for greater ranges of saturation and intensity to be classified, making the system more robust in varied brightness levels. The most improvement came from the classification of a greater range of intensity values for yellow. Table 1 shows an improvement of yellow blob formation, these were particularly in images of distant goals. It can be seen from the set of images that the automated classification and blob formation for yellow is much ‘cleaner’ than manual.

4.2 Colour Separation

Colours were successfully separated. However the automated table classified a greater area of blue than the manual table. The pixel values of the blue goal and blue robot greatly overlap and the additional classified blue resulted in more false positive goals than the manual table (Table 1). However this is more due to the fact that the object recognition system that was calibrated for the manual table and has not accounted blue robot recognition yet.

4.3 Time Improvement

Calibration time was greatly reduced from approx 4 hours to approx 10 minutes. The R program to generate the *mvnpdf* and EM algorithm took approximately 5 minutes to compute on an Intel Core 2 Duo E6600 (@2.4GHz) desktop computer with 2GB RAM. The Matlab code to compute the HSI classified regions took approximately 2 minutes and the C++ code to generate the YUV lut took approximately 2 minutes. The number of required images was greatly reduced from approx 275 random test images to 6 selected images.

5 Conclusion and Future Work

Results have shown that colour classes can be successfully separated and labeled in HSI colour space. The extension from one dimensional to three dimensional segmentation and the improvement in camera quality from the Aibo to the Nao robot produces a system that is more amenable to automation and provides an obvious method to separate colours and detect black and white regions. The EM algorithm provides a quick and stable estimate of colour component weighted mean and variance-covariance matrices. The addition of soft colour principles allows for flexibility when separating colour classes and results in a practical automated colour table.

While the reliability of the automated table was very similar to that of the manual table, the time improvement was significantly greater. Classifying in solid HSI regions resulted in a larger number of classified pixels

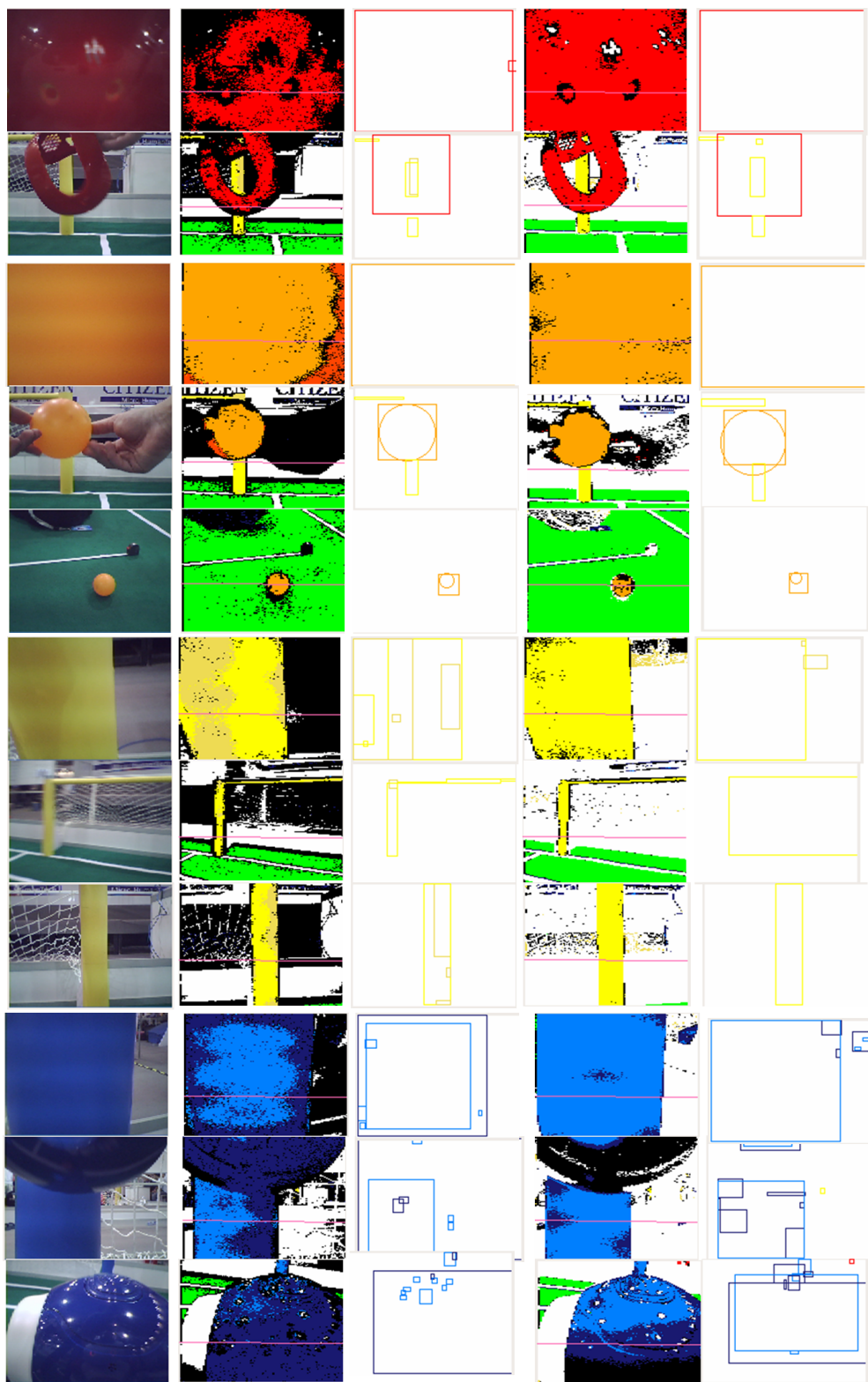


Figure 4: (First Column) Input image taken at RoboCup 2008, (Second Column) Classified image using manual colour table, (Third Column) Blobs formed with manual colour table, (Fourth Column) Classified image using automated HSI colour table, (Fifth Column) Blobs formed with automated colour table.

and 'solid' classified images. This is an advantage and disadvantage. Classifying more colour space, particularly a larger range of intensity values, improves the vision systems robustness in dynamic lighting conditions. However, manual classification is more sparse but improves reliability in a non dynamic situation. Future work involves integrating the automated HSI classification system into the competition vision system. For this, improved robot and goal recognition is required to account for the overlap of blue. Additional checks of an objects geometry and texture are to be accounted for using hue edge detection.

References

- [Bishop, 2006] Bishop, C.M. *Pattern Recognition and Machine Learning*. Springer, New York, 2006.
- [Dempster *et al.*, 1977] Dempster, A. P. and Laird, N. M. and Rubin, D.B. *Maximum likelihood from incomplete data via the EM algorithm*. Journal of the Royal Statistical Society B, vol 39, pages 1-38, 1977.
- [Cameron and Barnes , 2004] Cameron, D., and Barnes, N. *Knowledge-Based Autonomous Dynamic Colour Calibration*. D. Polani et al. (Eds.): RoboCup 2003: Robot Soccer World Cup VII, LNAI 3020, pp. 226-237, 2004.
- [Gunnarsson *et al.*, 2006] Gunnarsson, K., Wiesel, F., Rojas, R. *The Color and the Shape: Automatic On-Line Color Calibration for Autonomous Robots..* RoboCup 2005: Robot Soccer World Cup IX, pages 347-358, 2006.
- [Abelson *et al.*, 1985] Harold Abelson, Gerald Jay Sussman, and Julie Sussman. *Structure and Interpretation of Computer Programs*. MIT Press, Cambridge, Massachusetts, 1985.
- [Sridharan , 2005] Sridharan, M. and Stone, P. *Color Learning on a Mobile Robot: Towards Full Autonomy under Changing Illumination*. In The International Joint Conference on Artificial Intelligence (IJCAI 07), 2007.
- [Amoroso *et al.*, 2000] Amoroso, C., Chella, A., Morreale, V., Storniolo P. *A Segmentation System for Soccer Robot Based on Neural Networks*. RoboCup-99: Robot Soccer World Cup III, pages 136-147, 2000.
- [Quinlan, *et al.*, 2003] Quinlan, M., Chalup, S., Middleton, R. *Application of SVMs for colour classification and collision detection with aibo robots..* In Proceedings of the Advances in Neural Information Processing Systems Conference, pages 635-642, 2003.
- [Quinlan *et al.*, 2005] Quinlan, M. J., Nicklin, S. P., Hong, K., Henderson, Young, S. R., Moore, T. G., Fisher, R., Douangboupah, P., Chalup, S. K., Middleton, R. H. and King, R. *The 2005 NUBots Team Report*. University of Newcastle (EECS) Technical Report, 2005.
- [Sande, 2005] van de Sande, A. http://en.wikipedia.org/wiki/Image:Color_cones.png.
- [Greenspan *et al.*, 2001] Greenspan, H., Goldberger, J., Eshet, I. *Mixture model for face-color modeling and segmentation*. Pattern Recognition Letters Volume 22, Issue 14, Pages 1525-1536, 2001.
- [Raghu and Yegnanarayana, 1998] Raghu, P.P., Yegnanarayana, B. *Supervised texture classification using a probabilistic neural network and constraint satisfaction model*. Neural Networks, IEEE Transactions on , vol.9, no.3, pages 516-522, 1998.
- [Cohen *et al.*, 2003] Cohen, D., Hua Ooi, Y., Vernaza, P. and Lee, D. D. *The University of Pennsylvania Robocup 2003 Legged Soccer Team*. Technical Report University of Pennsylvania, 2003.
- [Soest *et al.*, 2006] D.A. van Soest, M. de Greef, J. Sturm and A. Visser. *Autonomous Color Learning in an Artificial Environment*. Proc. 18th Dutch-Belgian Artificial Intelligence Conference, BNAIC'06, Namur, Belgium (pp. 299-306), 2006.
- [Moghaddam *et al.*, 1997] Harold Abelson, Gerald Jay Sussman, and Julie Sussman. *Probabilistic visual learning for object representation*. Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.19, no.7, pp.696-710, 1997
- [Henderson *et al.*, 2008] Naomi Henderson, Robert King, and Richard H. Middleton. *An Application of Gaussian Mixtures: Colour Segmenting for the Four Legged League Using HSI Colour Space*. RoboCup 2007: Robot Soccer World Cup XI (Lecture Notes in Computer Science), Springer Berlin / Heidelberg, p254261, 2008.
- [Genz *et al.*, 2007] Genz, A., Bretz, F., port by Torsten Hothorn. *R.: mvtnorm: Multivariate Normal and T Distribution*. R package version 0.8-1, 2007.
- [R, 2008] R Development Core Team *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, 2008.