# An Illumination Identification System for the AIBO Robot

Michael S. Zehmeister, Yang Sok Kim and Byeong Ho Kang
*School of Computing, University of Tasmania*
*Sandy Bay, Tasmania 7001, Australia*
*{yangsokk, bhkang}@utas.edu.au*

## Abstract

*The Four Legged League is a division of the RoboCup initiative that uses Sony AIBO$^{TM}$ robots to further robotics research. Most participants implement vision systems that use the color of objects to perform identification. Calibration of the color classification system must be done and any changes to the lighting of the environment after calibration reduces the accuracy of the system, often to a point at which the robot is effectively blind. This study investigates the relationships in the color data of image pixels between lighting conditions in an effort to identify trends that can be used as the basis of a rule-based system. The aim of the system is to identify the current lighting level as one of a set of known conditions. The proposed systems uses the color data of image pixels and information about the AIBO's location and orientation to identify lighting levels, allowing a vision system to switch to an appropriate pre-configured calibration.*

## 1. Introduction

This research focuses on color-based vision systems using low-quality cameras, specifically in the domain of the Four-Legged League of the AIBO robot. The primary aim of this research is to determine if it is possible to develop a system that can identify lighting conditions based on the YUV data of image pixels. The first step involves creating a system that takes in an image or sequence of images as input and processes them with the intent of providing meaningful information about how lighting conditions affect the color classification process and the YUV data of pixels. The desired result was to figure out a way to use the raw color data of the image pixels to reliably indicate a change in lighting conditions. While the area of direct application is on the AIBO in the RoboCup domain, the work is applicable in any area that uses color-vision in a situation in which lighting changes may occur.

This paper consists of following contents: Section 2 provides a summary of prior research results. Section 3 describes our methodology in detail, including used system, experimental environment, image acquisition process, and data analysis approach. Experiment results and evaluation of rule-based illumination classification approach are presented in Section 4. Conclusions and further work is discussed in Section 5.

## 2. Related Work

Murch and Chalup [1] implemented a vision system that combined color classification with edge detection to improve the recognition of objects. While the technique proved effective in that it improved the performance of sub-optimally calibrated lookup tables, it had a serious problem in that motion blur on images taken when the AIBO was moving its head rendered the edge detection algorithm ineffective. There are also issues with partially obscured objects and the problem of false edges being found.

Jungel and colleagues [2] and Jungel [3] used a technique for updating a dynamic color cube as a lookup table. Essentially the color space is divided up into rectangles by setting thresholds for each color class. The unique part is that instead of static thresholds all of the color classes are defined by their relation to the color class 'green'. It was found that the shifts were consistent enough among the different color classes to make this process yield fairly accurate results in changing light. While the reduced accuracy is made up for by the simplicity of the system and the removal of the need for manual calibration there are problems with colors that have comparatively small differences between each other such as orange and yellow.

Sridharan and Stone [4] presented a promising technique that involves using a set of three color cubes; one calibrated for light conditions, one for dark conditions and one for conditions of intermediate illumination. By using the KL-divergence algorithm,

the system is able to detect which lookup table is best suited for the current conditions and the robot is able to switch between the three calibrations accordingly.

Lenser and Veloso [5] use a mathematical algorithm to detect changes in illumination over time. They implement a system that uses this detection algorithm to switch between several sets of pre-calibrated vision thresholds. This method is similar to that implemented by [4, 6] but using a different method of detecting the illumination and thus which set of color table data to use.

## 3. Methodology

The image processing software was implemented in Java. Some of the code was based on offline color calibration tools that form part of the rUNSWift 2004 RoboCup system. [7]. The nature of the software as an offline analysis tool also means that the system does not need to run on the AIBO platform itself. Two methods of image processing were tested. The first is to process only the pixels that the color classification process classifies as *green*. The reason green is used is because it is the only color that can be relied on to be present in the vast majority of images. The expectation is that in the normal condition for which the lookup table has been calibrated the most pixels will be correctly classified and that in other lighting conditions fewer pixels will be classified, as the calibration has been rendered inaccurate by the unfamiliar conditions. By comparing the data for the pixels classified green in each condition it should be possible to see variations between lighting conditions in the data of what pixels are classified. The second method will be to simply look at the YUV data of *all pixels* in the image regardless of the results of the color classification process. This method is less reliable as there are no constraints (other than those inherent in the domain) to limit what pixels are processed. That said there is a good chance that patterns may be visible between lighting conditions in the Y channel data due to its very nature as a representation of the brightness of a color.

### 3.1. System Overview

When the user opens an image in the software it is displayed in a window. Upon opening a calibration file the user can toggle the display between the original image and the color-classified image. The software also allows the user to save the classified version of the image. The software reads and saves images in the portable pixmap (P6) format (.ppm extension). This format represents the RGB color information of an image as non human-readable binary data. Portable pixmap images do not compress the image at all and are very quick to read and process. As the portable pixmap format stores images in RGB (the color space of the acquired images to be used) conversion into the YUV color space must be performed before the image is processed. This is done using the following standard conversion equations [8].

$$Y = (0.299*R) + (0.587*G) + (0.114*B)$$
$$U = (0.500*R) - (0.419*G) - (0.081*B)$$
$$V = -(0.169*R) - (0.331*G) + (0.500*B)$$
, where R, G, B & Y are in the range 0 to 1 and U & V are in the range –0.5 to 0.5

The classification process used is based on that of the rUNSWift 2004 RoboCup entry system. The specific implementation used as a reference is that of the offline calibration tool MC (Manual Classifier) that is used in the process of creating and maintaining color classification lookup tables. The set of images is manually classified one by one and this manually generated classification data for each image is fed into another tool called 'calibrate' that runs the data through a nearest-neighbor algorithm to map the YUV color space and generate a lookup table that is used for color classification. Once a calibration is created it acts as a simple lookup table for the vision system. It is done this way simply for reasons of speed. The calibration table is loaded into an array that has a length equal to the number of unique YUV values. Normally the values for the Y, U & V channels range from 0 to 255, but to reduce the file size the rUNSWift system creates a calibration based on the channels having a range of 0 to 128. The lookup table used for this study was calibrated for the 'normal' lighting condition as described in Section 3.2. The calibration table is loaded into an array, whose length is 2097152 ($128 \times 128 \times 128$) like rUNSWift. To classify a color the YUV values are used to refer to one of the elements in the array, the element will contain a number representing the color that the YUV value is classified as.

### 3.2. The Environment

This section details the conditions in which the images for the data study were taken.

**The Field** The field meets the RoboCup specifications for dimensions and color scheme, but is only half of a whole field due to limitations in the availability of resources and physical space [9].

**Lighting Conditions** The study compares the YUV information of processed images in three distinct lighting conditions: bright, normal and dim. The lights

in the laboratory are standard overhead fluorescent lights. As such, there is no control in terms of dimming the lights but the room does have two light switches, each controlling half of the lights. The bright condition has both sets of fluorescent lights on. The normal lighting condition has only the lights in the half of the room in which the field is located on so that the light is directly overhead. The dim condition has only the lights in the other half of the room on. The lookup table used for color classification was calibrated for the normal condition.

**The Camera** The device used was a Canon A75 digital camera. The main reason for using this particular model was that it has full manual control over all settings, which allowed for the closest approximation possible of the AIBO camera. The smallest pictures that could be taken were 640×480 so they had to be shrunk to 176×132 before processing. Sequences of images were taken using the camera's drive mode, which continuously takes pictures while the button is held down. The camera settings used were 400 ISO speed, F2.8 aperture, and 1/25 shutter speed.

### 3.3. Image Acquisition

For the purposes of the study, three main types of image sequences were taken. The first type is sequences that consist of a large number of random images, all in the same lighting condition (Figure 1(a)). The purpose of these sequences is to get a good range of images displaying various compositions so that the analysis of the sequence as a whole would give an indication of broad trends in the YUV data regardless of what the individual pictures were of. Most images contained at least some of the green field within the frame. The only object on the field was the orange ball, which appeared in some but not all images. The second type of sequence was smaller sets of images with each picture recreated in each lighting condition (Figure 1(b)). Positioning the camera on a stand and taking the same picture in each lighting condition created these sequences. The result was a series of images in each lighting condition with matching compositions for each picture. This allowed for an analysis of how YUV data changes between lighting conditions given identical image compositions. The aim was to see if it was possible to identify a pattern in how the data of lighting conditions related to each other to create rules that would hold true in all situations. The third and final type of image sequence acts as an intermediate between the first two types. Large sequences were taken in each lighting condition but all from the same general location and in the same general direction (Figure 1(c)). As such each sequence has a reasonable

amount of variation in regards to the image composition, but the content from image to image is fairly consistent. The purpose of these sequences is to allow for adding an extra dimension to the possible rule system, that being a constraint of location and orientation.
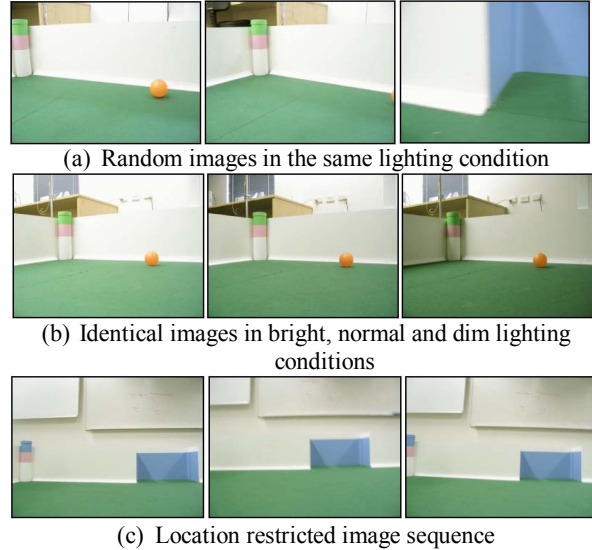

(a) Random images in the same lighting condition


(b) Identical images in bright, normal and dim lighting conditions


(c) Location restricted image sequence

**Figure 1. Three types of image sequences**

The proposed method for restricting rules to specific locations and orientations would require knowledge of the AIBO's position and facing on the field. This information is easily obtainable from the localization module of the system. Basically, the field would need to be virtually divided up into sections with each section having several possible orientations. In this example the field is divided into a 4×4 grid with eight directions for each location.

### 3.4. Data Analysis

The first step was the analysis of the general image sequences in each lighting condition to perceive the broad trends in the data to give an indication as to what types of information would most likely prove fruitful in the later steps of the study. This initial step is important as it can tell us what areas are not likely to yield useful information upon closer inspection. The second step involves comparing the data of the identical images in each lighting condition. The aim of this step is to identify consistent relationships between the data types in the distinct lighting conditions. This stage should reveal if it is possible to create rules that apply in all situations or if the extra information about image composition is needed. Finally, an analysis of images constrained to a single location and orientation will show if there is a consistent enough relationship

between the YUV data and the lighting conditions if the image contents are known to be generally uniform. If it is possible to create rules constrained by location and orientation then the proposed system will incorporate both color and physical localization information to produce rules for lighting identification.

# 4. Results

## 4.1. Trends

Figure 2(a) illustrates that the U & V channel variation in data from one lighting condition to another is only slight. In Y channel, several areas showed reasonable levels of variation between lighting conditions. In particular, the mean, minimum and inter-quartile statistics appeared to show significant levels of variance (Figure 2(b)).
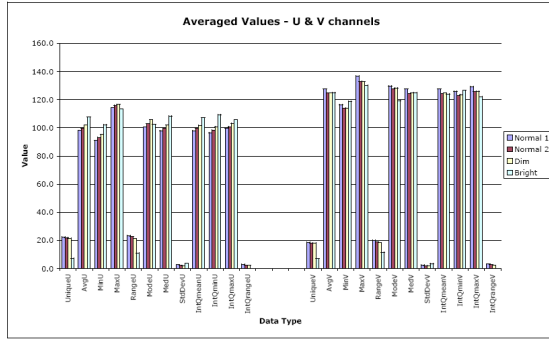


(a)   U&V Channels
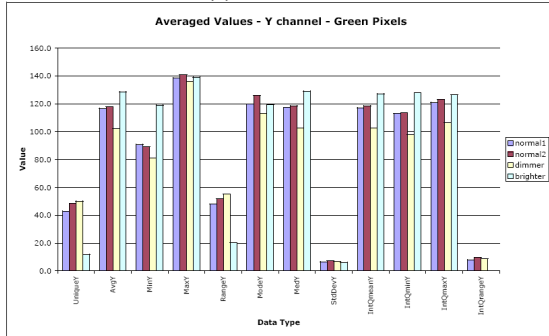


(b)   Y Channel

**Figure 2. U, V, and Y Trends**

Looking closer at the separate values in the sequences, it can be seen that while the trends are definitely discernable the consistency is far from ideal. The mean Y channel values have a large range for the green pixel results but there is definitely a trend for the dim values to be lower and the bright values higher with normal values in the middle. That said there is a large overlapping range between the higher normal values and the lower bright values, there is also a large overlapping range between the higher dim values and

the lower normal values. The inter-quartile minimum values for the all pixel results showed the most consistency for the sequences though there was still significant overlap.

## 4.2. Direct Image Comparisons

Looking at the values for the green pixel results, it is apparent that without restriction on image contents there is a large variation in values from image to image while on a per-image basis there are significant differences in values between lighting conditions. For example in the mean Y channel values (Figure 3(a)) it can be seen that for most images the mean Y channel value is lowest in the dim condition, highest in the bright and in-between in the normal.
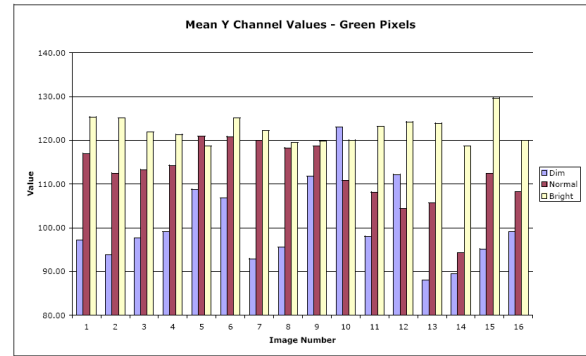


(a)   Mean Y Channel Value
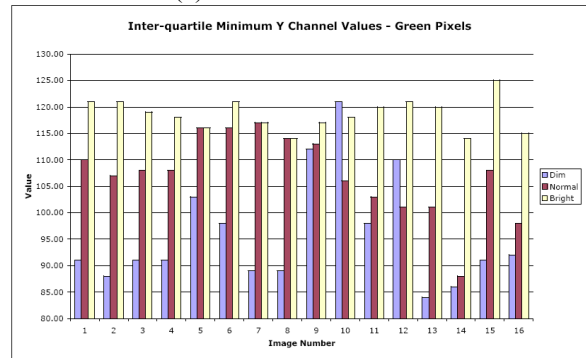


(b)   Inter-Quartile Minimum Y Channel Value

**Figure 3. Direct image comparison**

That said there are cases for which this does not hold true which makes this aspect of the data unsuitable for rule making. The same circumstance is evidenced in the inter-quartile minimum Y channel values (Figure 3(b)). While there is a general trend for the bright condition to have a higher value it is not consistent enough for rule creation. This stage of the study showed that there is definitely too much inconsistency in the data to make rules that apply in all situations without a restriction on image contents.

Undertaking the direct image comparison has shown that there are certain relationships in the data that remain reasonably constant.

## 4.3. Location Restricted Images

This analysis focuses on the sequences of images that are all taken from the same general location and facing the same general direction. The result is a sequence of images that has very little variation in image contents compared to the sequences studied early in the analysis. For each of the two locations tested at this stage sequences for both training and testing were taken in each lighting condition.

**Location 1** The first location is left of the centre facing the blue goal. From this angle, the images have very little variation in terms of contents with most of the image being taken up by the ground, wall and goal. As such, the results are far more consistent from image to image in the sequences. Looking at the number of unique Y channel values of the green classified pixels for the training sequences in each lighting condition it can be seen that there is a much more consistent pattern in the data than with the unrestricted images (Figure 4(a)). The amount of variation is far less and as such, there is no overlap between the values in different lighting conditions. Using this data the following rule-set was composed:

```
If [Unique Y channel values < 15] then Lighting = Dim
If [Unique Y channel values > 40] then Lighting = Normal
If [20 < Unique Y channel values < 40] then Lighting = Bright
```

The above listed rule-set successfully detected the correct lighting condition for 100% of the images in the test sequences for the same location and orientation.

Using the data for all pixels in the images it can again be seen that by restricting the image contents by applying rules only for a certain location and orientation much more consistency is evident in the values. The inter-quartile maximum Y channel values (Figure 4(b)) show no overlap in values between lighting conditions. Using this data the following rule-set can be made:

```
If [IQM Y channel value < 140] then Lighting = Dim
If [160 < IQM Y channel value < 220] then Lighting = Normal
If [IQM Y channel value > 220] then Lighting = Bright
*IQM: Inter-quartile Maximum
```
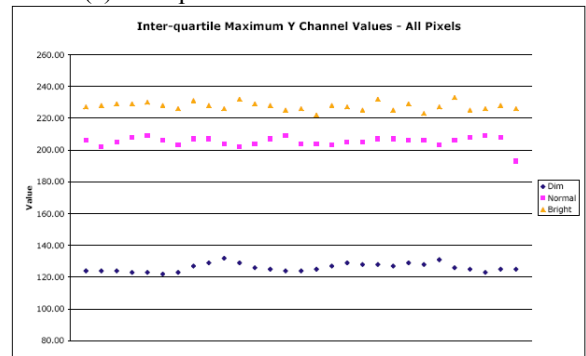
Once again, this rule-set successfully detected the correct lighting condition for 100% of the images in the test sequences. This result shows that restricting the

image contents to a reasonable extent allows rules based on the YUV data to be crafted that accurately detect lighting conditions.
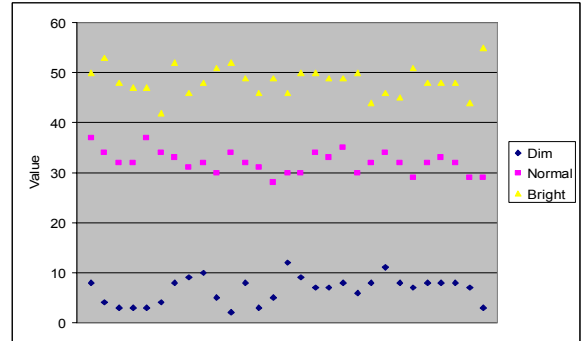


(a) Unique Y Channel Values – Green Pixel



(b) Inter-Quartile Maximum

**Figure 4. Comparison of Location Restricted Images**

**Location 2** This second location provided more variation in the image sequence than location 1. This location faced diagonally towards the corner to the left of the blue goal from a position near the right wall close to the centre line of the field. The image contents varied more in this second set of sequences due to both the larger range of shades of green visible by looking across a significant portion of the field diagonally and the larger amount of visible background. The result was noisier data than that from location 1. The data for green classified pixels again shows that the best results appear to be gained from the number of unique Y channel values. While there is a larger range of values for each condition than in the data from location 1 there is still no overlap of values between lighting conditions. Using this data the following rule-set is made:

```
If [Unique Y channel values < 20] then Lighting = Dim
If [Unique Y channel values > 50] then Lighting = Normal
If [20 < Unique Y channel values < 50] then Lighting = Bright
```

This rule set correctly identified the dim and normal lighting conditions for 100% of the images from the test sequences but only correctly identified 96.3% of the bright condition images with the remaining 3.7% incorrectly identified as normal. While this result is not perfect, it is still within an acceptable range.

Looking at the results for all pixels from this second location the best consistency of values is found in the inter-quartile minimum Y channel values. The difference in values between the normal condition and the bright is not as great as would be ideal but the consistency appears to be suitable. Using these values the following rule-set was tested:

```
If [IQM Y channel value < 130] then Lighting = Dim
If [115 < IQM  Y channel value < 140] then Lighting = Normal
If [IQM Y channel value > 140] then Lighting = Bright
*IQM: Inter-quartile Maximum
```

This rule-set again correctly identified the dim condition in 100% of the images of the test sequence. 100% of the images in the bright condition were correctly identified but only 92.6% of the images in the normal condition were correctly identified. The remaining 7.4% of images in the normal condition were misclassified as bright. The results for this second location were not as clear as for the first though they still show a significant level of accuracy in assessing the lighting conditions despite a noisier range of image compositions. Together the results for the two locations show that a system can be created by applying different rules for each combination of location and orientation on the field to identify lighting conditions with reasonable accuracy by using the YUV data of the image pixels.

## 5. Conclusions

While it is not possible to create a system of rules to determine the current lighting condition that holds true in all cases, a system can be created that has distinct rules for each location and orientation, information about which is easily obtainable from the localization module. Based on Y channel pixel data the proposed system can accurately estimate the lighting condition and allow the vision system to switch to a color lookup table calibrated for the current situation. Due to the fact that the system is reliant on variable conditions such as the physical environment and the color classification calibrations used (in the case of the green pixel method) the development of a system to partially automate the construction of the rule set would be advantageous. This could reduce the time and effort it would take to create a rule set. Further testing would also be beneficial to establish the robustness of an implemented system.

## 7. References

1. Murch, C.L. and S.K. Chalup. Combining Edge Detection and Colour Segmentation in the Four-Legged League. in 2004 Australasian Conference on Robotics & Automation. 2004. Canberra, Australia.
2. Jungel, M., J. Hoffmann, and M. Lotzsch. A Real-Time Auto-Adjusting Vision System for Robotic Soccer. in RoboCup 2003: Robot World Cup VII, Lecture Notes in Artificial Intelligence. 2004. Padova, Italy: Springer.
3. Jungel, M. Using Layered Color Precision for a Self-Calibrating Vision System. in ROBOCUP2004 SYMPOSIUM. 2004. Lisboa, Portugal.
4. Sridharan, M. and P. Stone. Towards Illumination Invariance in the Legged League. in RoboCup-2004: Robot Soccer World Cup VIII. 2005. Berlin, Germerny: Springer Verlag.
5. Lenser, S. and M. Veloso. Automatic detection and response to environmental change. in 2003 IEEE International Conference on Robotics and Automation (ICRA '03). 2003.
6. Sridharan, M. and P. Stone, Real-Time Vision on a Mobile Robot Platform. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005.
7. Xu, J., 2004 rUNSWift Thesis - Low Level Vision. 2004, University of New South Wales.
8. Li, Z.-N. and M.S. Drew, Fundamentals of Multimedia. 2004, Upper Saddle River, NJ: Pearson Prentice Hall, Pearson Education, Inc.
9. Sony Four Legged Robot Football League Rule Book. 2004, RoboCup Technical Committee.