



Hamburg University of Technology
Vision Systems

Prof. Dr.-Ing. R.-R. Grigat

**Field Color Detection using
Illumination Invariant Imaging in the
RoboCup Standard Platform League**

Bachelor Thesis

Justus-Frederik Basler

December 14, 2018



Statutory Declaration

I, Justus-Frederik Basler, born on 15.2.1994 in Hamburg, hereby declare on oath that I compiled this bachelor thesis, submitted to Hamburg University of Technology (TUHH), on my own. I only used the declared sources and auxiliaries.

Place and Date

Signature

Contents

List of Figures	iv
List of Tables	vi
1 Introduction	1
1.1 The RoboCup Standard Platform League (SPL)	1
1.2 Motivation and Goals	3
1.3 Thesis Structure	3
2 State of the Art	4
2.1 Existing Approaches in the RoboCup	4
3 Comparison of Colorspaces	7
3.1 Material and Methods	7
3.1.1 The NAO Robotic System	7
3.1.2 Ground Truth Images	8
3.1.3 White Balance	12
3.1.4 Comparison Metrics	14
3.2 Evaluation	15
3.3 Discussion	20
4 Classification into Field and Non-Field Pixels	21
4.1 Material and Methods	21
4.1.1 Classification Using Thresholds	21
4.1.2 Classification Metrics	23
4.2 Evaluation	25
4.3 Discussion	28
5 Discussion	30
6 Conclusion and Future Work	32

List of Figures

1.1	Two HULKs NAOs defending against an opponent at the RoboCup German Open 2018.	2
3.1	Lateral view and top view of a NAO head depicting the position and field of view of the two cameras [1].	8
3.2	Set of Ground Truth Test images from the GermanOpen 2018	9
3.3	Set of Ground Truth Test images from the IranOpen 2018	9
3.4	Set of Ground Truth Test images from the RoboCup 2018	9
3.5	Set of Ground Truth Test images from the HULKs laboratory 2018	9
3.6	Two test images taken with the NAO's top camera recorded at the RoboCup 2018 with the horizon marked by a red line. A yellow line marks the distance of 13 meters from the robots feet which is equivalent to the length of the field's diagonal.	11
3.7	The original images next to two different representations of the labeled image in the first row and a difficult image segment due to blurring and small details in the same representation in the second row	12
3.8	Two images with varying white balance settings with figure 3.3e on the left having a blue tint and 3.4e on the right with a slight yellow tint.	13
3.9	Normalized distributions of field pixels (green) and non-field pixels (red) in two different channels. On the left the distribution in Cb is shown with a large overlap area representing a high intersection value. The distribution in g_{wb} is shown on the right with a small overlap area and thus a low intersection value.	17
3.10	Comparison of the two two-dimensional combinations $V_{wb} - g_{wb}$ on the left and $I - g_{wb}$ on the right. The corresponding one-dimensional distributions are displayed alongside the two-dimensional histogram. The distributions on the left have an intersection value of 0.037600 and the combination on the right leads to a slightly higher intersection value of 0.037925.	19

4.1	Normalized field and non-field pixel distributions in g_{wb} channel and the corresponding cumulative histograms with the optimal threshold marked by dotted gray lines.	22
4.2	Normalized field and non-field pixel distributions in I channel and the corresponding cumulative histograms with both optimal minimum (left) and optimal maximum (right) threshold marked by dotted gray lines. . .	23
4.3	Two-dimensional normalized pixel distribution of field and non-field pixels in g_{wb} and I_{wb} with the corresponding one-dimensional histograms of each dimension. Optimal classification thresholds for both channels are marked by dotted gray lines.	24
4.4	Classification predictions for the ball test image taken at the GermanOpen 2018. The g_{wb} thresholding does not result in satisfying results with a balanced accuracy of 0.836. Thresholding of the intensity results in a balanced accuracy of 0.891. Combining both classifiers however results in a balanced accuracy of 0.925.	26
4.5	Classification predictions for the test image showing another NAO robot taken at HULKs laboratory with bright sunshine outside. The g_{wb} thresholding results in a near perfect classification except for the yellow object in the background with a balanced accuracy of 0.949. However, thresholding of the intensity fails completely and results in a poor balanced accuracy of 0.664. Combining both classifiers results in a balanced low balanced accuracy of 0.672.	27
4.6	Classification predictions for the penalty spot test image taken at the IranOpen 2018. Both classifiers result in good balanced accuracy scores of 0.929 and 0.997. The combination of both classifiers also leads to a good balanced accuracy of 0.993.	27
4.7	Illustration of lower classification performance at very dark and thus colorless pixels. The first row shows the ball test image taken at the HULKs test lab where the brightness varies but all field pixels still contain color information. The second row shows the ball test image taken at the RoboCup which contains field pixels that do not contain color information and are therefore miss-classified. The difference is apparent in the center column showing the chromaticity images for both test images.	28

List of Tables

3.1	Intersection values of channels from different colorspace that describe the brightness of a pixel. The intersection value provides an indication of the separability of field and non-field pixels by measuring the two distributions' overlap. An intersection of 0 would be best as this represents no overlap and a value of 1 would represent identical distributions with total overlap.	15
3.2	Intersection values of channels from different colorspace that describe the color of a pixel. The intersection value provides an indication of the separability of field and non-field pixels by measuring the two distributions' overlap. An intersection of 0 would be best as this represents no overlap and a value of 1 would represent identical distributions with total overlap.	16
3.3	Comparison of intersection values of selected two-channel combinations. The intersection value provides an indication of the separability of field and non-field pixels by measuring the two distributions' overlap. An intersection of 0 would be best as this represents no overlap and a value of 1 would represent identical distributions with total overlap. Only the combinations with the lowest intersection and selection of other relevant combinations is shown. The comment column provides more information about the combinations.	18
4.1	Balanced accuracy of the prediction results of all all classifiers (columns) for all images (rows) are shown in this table. A value of 1 would represent a perfect classification. The subscript wb was omitted for a more concise table.	25

Chapter 1

Introduction

Autonomously operating robotic systems like autonomous vehicles are starting to become a ubiquitous part of everyday life and represent a very active field of research. While computers surpass humans in various well defined scenarios like boardgames already [2], it is still challenging for autonomous systems to perceive the environment with different types of sensors and derive behaviors from these sensor inputs. In 1997, the same year the NASA Rover Sojourner landed on Mars as an experiment for autonomous vehicle technology [3], the RoboCup was founded. Its goal was to promote research in the field of intelligent robots and autonomous systems. The RoboCupSoccer league organizes competitions of robot teams playing soccer against each other in different sub leagues at local competitions and at the yearly Robo World Cup. TUHH joined the Standard Platform League (SPL) in 2013 with its team known as the HULKS [4].

1.1 The RoboCup Standard Platform League (SPL)

In the RoboCup Standard Platform League (SPL) two teams of five humanoid robots play soccer against each other under increasingly realistic conditions, as described subsequently. To put the emphasis of this league on the software development, the NAO robotic system from Softbank Robotics Corp. (Tokio, Japan) [5] as seen in figure 1.1 is used as a standard hardware platform. Hardware alterations and the usage of external computing power is prohibited to ensure equal conditions for all teams. Other leagues with different research focusses exist within the RoboCupSoccer league such as the Humanoid League with a emphasis on mechatronics or the Simulation League with a focus on team behavior.



Figure 1.1: *Two HULKs NAOs defending against an opponent at the RoboCup German Open 2018.*

The original RoboCup mission from 1997 is to beat a human world champion team in 2050 under realistic conditions [6]. To pursue this goal, the rules of The Robo World Cup are changed every year to incrementally move towards the conditions of human soccer matches. Recently, the rules of the SPL regarding the illumination of the playing field became less restricted. Direct sunlight is now possible and changing lighting conditions in between matches are encouraged. Furthermore, the color-coding of important objects, i.e. goals and the ball was reduced. The red ball was replaced with one that resembles a typical soccer ball with the characteristic white and black pentagons and similarly the blue and yellow color coded goals were changed to plain white. The color of the playing field is supposed to be green but is not strictly defined. No additional landmarks for self-localization are allowed [7].

1.2 Motivation and Goals

Robust self localization on the field is essential to compete successfully in the Robo Cup. The accurate knowledge of its own position is important for the autonomous robot to build a reliable world model or plan its path to its destination, for example the opponent's goal. The visual detection of land marks like the field lines, field borders, center circle and the penalty spots is crucial for a successful self localization. Furthermore, moving objects such as the ball and other NAOs have to be detected visually with the NAO's cameras as the sonar sensors only provide rough distance measurements which are insufficient to detect object boundaries.

In the robot's view, the aforementioned landmarks and objects are mostly surrounded by pixels depicting the green field. Therefore, the calculation of the field color and field color based classification of each pixel as either showing the field or not showing the field is a valuable preprocessing step. Currently, the field color is scarcely used as the current implementations is not very reliable, especially in challenging lighting conditions. Different lighting conditions pose a challenge to the vision system as they can alter the perceived image severely.

The goal of this study is to systematically compare different color spaces with regard of their ability to distinguish between field pixels and non-field pixels and to develop a classification algorithm from the best representations. This algorithm should be robust to different lighting conditions and should work while playing on different fields. Furthermore, the algorithms complexity should be as low as possible due to the NAOs' limited computing power.

1.3 Thesis Structure

First, existing field color detection approaches within the RoboCup are discussed and the colorspace utilized in these approaches and their calculation are explained. Next, the NAO robotic system is introduced with a focus on the cameras. It is explained how a set of test images was selected and labeled to obtain ground truth data. A simplified white balance correction algorithm is introduced as well as metrics to measure suitability for classification of single color channels and their combinations. The intermediate results are evaluated and discussed to continue with the most suitable approaches in the next chapter. There, different algorithms to classify the pixels increasing in computational cost are described, evaluated and discussed. In the main discussion the developed approach is compared to the existing ones. Finally, the overall results are discussed and an outlook for future work is provided.

Chapter 2

State of the Art

2.1 Existing Approaches in the RoboCup

The *RGB* color space, combined of the three tristimulus values Red, Green and Blue, is the native color space of a digital camera with a Bayer pattern filter matrix. The *RGB* colorspace is device-dependent mainly due to different filter characteristics that describe how *R*, *G*, and *B* are extracted from visible light. A vector of the three values describes one color in the *RGB* colorspace.

The NAO system stores the captured images in the *Y'CbCr* format, defined in the ITU601 reference standard [8], but is referred to as the "YUV22" format in the NAO documentation [1]. The term *Y'CbCr* is used in this study to avoid ambiguity to the *YUV* colorspace which is not defined precisely. *Y'* describes the luma component while *Cb* and *Cr* describe chroma components as the red difference and the blue difference, respectively. On the NAO system, two consecutive pixels are averaged together which reduces the chroma information by the factor of 2. The conversion of *RGB* to *Y'CbCr* is shown in equation 2.1. The reverse conversion from *Y'CbCr* to *RGB* is shown in equation 2.2. The equations are derived from the NAO cameras system's data-sheet [9]. All values range from 0 to 255.

$$\begin{pmatrix} Y' \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} .299 & .587 & .114 \\ -.1687 & -.3313 & .5 \\ .5 & -.4187 & -.0813 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix} \quad (2.1)$$

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1.402 \\ 1 & -0.3441 & -.7141 \\ 1 & 1.772 & 0 \end{pmatrix} \left[\begin{pmatrix} Y' \\ Cb \\ Cr \end{pmatrix} - \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix} \right] \quad (2.2)$$

In former SPL events, when a clear color coding was still present, a simple thresholding algorithm in the *CbCr* plane of the *Y'CbCr* color space led to sufficient segmenta-

tion [10]. After changing to a more realistic environment with white goals and a typical soccer ball other approaches became necessary.

Most approaches still use the $Y'CbCr$ color space directly. The Berlin United NAO Team uses thresholding in this color space with a variety of parameters which are adapted manually to the light conditions to deliver sufficient segmentation of the field [11]. The Nao-Team HTWK from Leipzig uses a machine learning algorithm trained offline on 300 images with images in the $Y'CbCr$ color space [12]. The current HULKS algorithm is using a simplified k-means clustering to detect the field color in the $CbCr$ plane in combination with thresholding in luma Y' , depending on the mean luma [13].

The B-Human team from Bremen uses a color space which they call $YHS2$ which is similar to the HSV color space [14]. The regular HSV color space, representing hue, saturation and value can be computed from RGB as seen in equation 2.3 with all values ranging from 0 to 255.

$$\begin{aligned}
 H &= \begin{cases} 0, & \text{if } \max(R, G, B) = \min(R, G, B) \\ \left(\frac{0}{3} + \frac{G - B}{\max(R, B, G) - \min(R, B, G)} \right) \% 1, & \text{if } \max(R, G, B) = R \\ \left(\frac{1}{3} + \frac{B - R}{\max(R, B, G) - \min(R, B, G)} \right) \% 1, & \text{if } \max(R, G, B) = G \\ \left(\frac{2}{3} + \frac{R - G}{\max(R, B, G) - \min(R, B, G)} \right) \% 1, & \text{if } \max(R, G, B) = B \end{cases} \\
 S &= \begin{cases} 0, & \text{if } R = G = B = 0 \\ \frac{\max(R, B, G) - \min(R, B, G)}{\max(R, G, B)} & \text{otherwise} \end{cases} \\
 V &= \frac{\max(R, G, B)}{255}
 \end{aligned} \tag{2.3}$$

As mentioned above, B-Human's $YHS2$ colorspace is similar to the HSV color space but it is derived in a less complex way directly from the $Y'CbCr$ color space to increase performance. It is calculated from $Y'CbCr$ as described in equation 2.4. To avoid ambiguity, the color channels of this color space are marked with the $YHS2$ subscript. The equations were adapted from B-Human's code [15] to span from 0 to 255. Multiple thresholds are used to segment the image into black areas, white areas, green field and other colored areas. However, only H_{YHS2} and S_{YHS2} are utilized to classify pixels as either field or non-field.

$$\begin{aligned}
Y_{YHS2} &= Y \\
H_{YHS2} &= \left(\frac{\arctan2(Cr - 128, Cb - 128)}{\pi} + 1 \right) \cdot 127 \\
S2_{YHS2} &= \frac{\sqrt{(Cb - 128)^2 + (Cr - 128)^2} \cdot 2}{Y'} \cdot 255
\end{aligned} \tag{2.4}$$

Another approach is derived from the *RGB* color space. The UPennalizers team from Philadelphia, USA utilizes the *g* chromaticity which is computed as *G* divided by the sum of *R*, *G* and *B* [16]. The area close to the robots' feet is used to find the most abundant *g* chromaticity value which is assumed to be the field color. Pixels within a defined range around this value are categorized as field. Similarly, the chromaticity values *r* and *b* can be computed as shown in equation 2.5. Due to the normalization, *r*, *g*, *b* are not independent but related as described in equation 2.6. Therefore, the whole *RGB* color space can not be described with only chromaticity values. The intensity *I*, defined in equation 2.7, can be used together with two chromaticity channels to describe the complete *RGB* color space [17]. Alternatively, one channel of the *RGB* color space can be used in addition to two chromaticity channels.

$$\begin{aligned}
r &= \frac{R}{R + G + B} \\
g &= \frac{G}{R + G + B} \\
b &= \frac{B}{R + G + B}
\end{aligned} \tag{2.5}$$

$$r + g + b = 1 \tag{2.6}$$

$$I = \frac{R + G + B}{3} \tag{2.7}$$

Chapter 3

Comparison of Colorspaces

3.1 Material and Methods

3.1.1 The NAO Robotic System

The NAO robotic system (short NAO) is a humanoid bipedal robot with a height of 57.3 cm and a weight of 5.2 kg. It is developed continually by Softbank Robotics. The current fifth generation is equipped with a 1.6 GHz Intel Atom CPU, including 1 GB RAM [18]. It has two digital color camera systems of the type MT9M114 [9]. The MT9M114 is a System-on-Chip (SoC) Digital Image Processor with a resolution of 1280 x 960 pixels, equipped with a RGB Bayer color filter array. The main functions of an image pipeline, including debayering, white balance, black level control, flicker avoidance and defect correction, are implemented in the system. The SoC works internally in the RGB and YUV color space and can output the image data in different formats. In the HULKS framework the YCbCr image format is used predominantly with a resolution of 640 x 480 pixels.

The two cameras are mounted in the robot's head vertically one above the other as shown in figure 3.1. This camera arrangement allows the NAO to see both the area directly in front of its feet and further away without moving its head up and down. However, a three-dimensional stereoscopic vision is not possible due to the small overlap of the two fields of view and due to the fact that top and bottom image can not be acquired simultaneously. In addition to the cameras the NAO is equipped with sonar sensors, microphones, pressure sensors and an Inertial Measurement Unit (IMU) to provide inputs about the environment.

The sixth generation is available since 2018 and is equipped with a more powerful processor, more RAM and an integrated GPU [19]. This new version is equipped with the OV5640 camera system which has different specifications compared to the fifth ver-

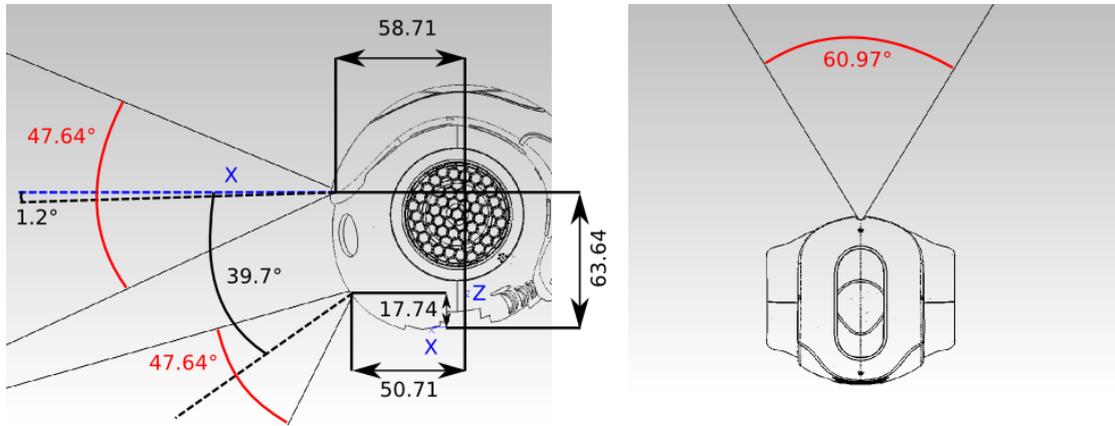


Figure 3.1: Lateral view and top view of a NAO head depicting the position and field of view of the two cameras [1].

sions cameras, including a higher resolution [20]. In combination with the increased processing power and the usage of the integrated GPU this allows for more complex image processing in the future. The programmable interface of the sixth version is not available yet, hence the fifth version of the NAO was used in this study.

3.1.2 Ground Truth Images

A visual inspection of the segmentation result can give a rough impression of the segmentation algorithm's performance. However, more precise metrics such as the accuracy, false positive rate and false negative rate can only be calculated when a ground truth is available. Here, the ground truth is defined as a classification of each pixel to either show a part of a field or not show a part of a field.

Replay Data

The replay recorder in the HULKs framework stores images of both cameras in YCbCr color space, sensor data and meta data at a fixed rate [13]. As this functionality is activated for most RoboCup events and test matches, there is a variety of datasets from different locations at different times with varying lighting conditions. The replay data is organized as follows: The top and bottom images are stored with individual timestamps. Additionally, a json file [21] is saved to store two kinds of information. First, all settings like the camera parameters and parameters used in the employed algorithms are stored. Second, all sensor data, joint angles and transformation matrices for kinematic transformation are recorded with corresponding timestamps. The cycle times of the thread running the motion of the NAO and the thread running the vision modules

are different. Thus, the additional data with the smallest time difference to the capturing of the images is matched to the image. This additional data enables a reprocessing of the collected images, yielding the same results as the ones computed by the NAO in the actual situation.

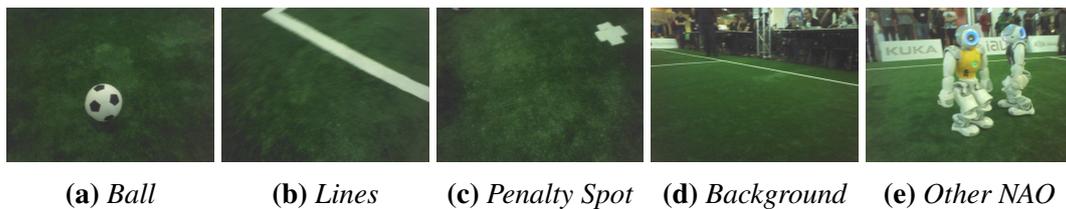


Figure 3.2: Set of Ground Truth Test images from the GermanOpen 2018

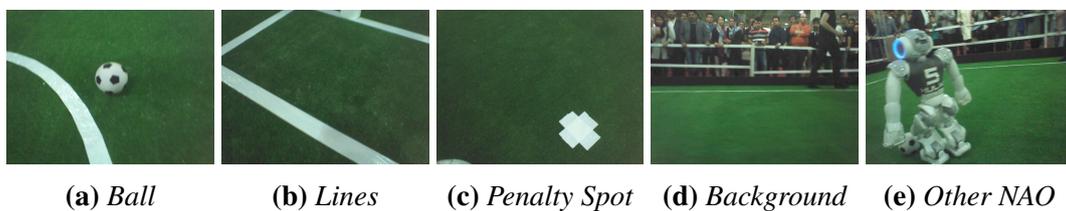


Figure 3.3: Set of Ground Truth Test images from the IranOpen 2018

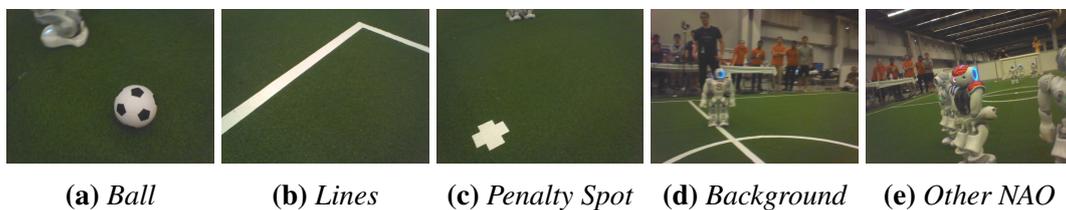


Figure 3.4: Set of Ground Truth Test images from the RoboCup 2018

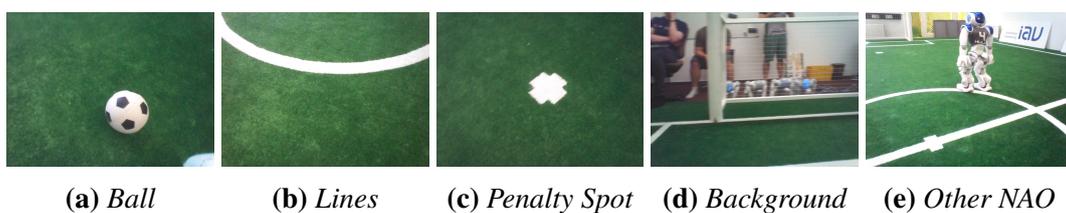


Figure 3.5: Set of Ground Truth Test images from the HULKs laboratory 2018

Selection of Test Images

It is crucial that the field color algorithm still works when applied to images captured in new settings. Color and structure of the field as well as the color and material of the field marks can vary. The lighting situation can range from dim indoor illumination to bright sunlight. There are no rules regarding the venue setup around the field except that there are no other fields visible with a distance below three meters between the turfs. This leads to a variety of different objects visible in the background with varying color, lightness and structure. [7]

As the replay recorder and thus the format of the data was changed recently, all selected settings are from 2018. Only settings with the more recent rules are present i.e. a black and white ball and white goals. The four settings are: The GermanOpen 2018 in Magdeburg (figure 3.2), the IranOpen 2018 in Tehran, Iran (figure 3.3), the RoboCup 2018 in Montreal, Canada (figure 3.4) and a test match in Hamburg on the test field of the HULKs (figure 3.5).

Five images of similar situations in each of the four settings were selected. The situations are related to important vision tasks in the HULKs framework. One image of the bottom camera with a ball present (a) is selected, as the ball detection is a crucial part of the vision pipeline. Two other images from the bottom camera with a penalty spot (b) and a variety of line segments (c) are selected, as these are important landmarks for self localization. Two images from the top camera were selected. One image contains a busy and thus challenging background (d), as this is important when looking for the field border, i.e. where the green turf ends. The other image of the top camera contains one or more other NAOs which can be opponents or HULKs team members (e). The visual detection of opponent players is not implemented in the HULKs framework at the moment. As a reliable distinction between field and non-field objects, including other NAOs, can facilitate this task, images with other NAOs are important.

Two kinds of images were not considered for the image set. Images with a strong blurring due to motion were discarded because the labelling of such images is challenging and thus does not provide a reliable ground truth for these images. Also, images with large overexposed areas are excluded as there is no information in the overexposed areas and thus the labelling is not meaningful.

Horizon

As visible in figure 3.4d and figure 3.4e, the content of the image above the field boarder i.e. were the green carpet ends, which is referred to as the background in this study, can



Figure 3.6: Two test images taken with the NAO's top camera recorded at the RoboCup 2018 with the horizon marked by a red line. A yellow line marks the distance of 13 meters from the robots feet which is equivalent to the length of the field's diagonal.

include a large variety of colors and shades. Usually, different people with clothes in different colors can be seen as well as furniture, banners with logos, different kind of barriers, walls and the ceiling. As the field is on the ground, one can assume that there are no field color pixels above the horizon. It is thus beneficial to exclude pixels above the horizon.

To calculate the horizon it is necessary to calculate the kinematic transformation matrix $ground2Camera$ as described in [13]. It represents the extrinsic camera parameters i.e. their position and rotation in relation to the point on the ground between the robot's legs. The intrinsic camera parameters are the camera center and the focal center and can be read from the camera parameters stored in the replay data. The horizon is calculated from the camera parameters as described in [22]. Similarly, the distance of a point on the field to the robot's feet can be calculated. Figure 3.6 shows the two top camera test images from the RoboCup with with the horizon marked by a red line. The yellow line marks the distance of 13 meters from the robots feet which is equivalent to the length of the field's diagonal. An array representing all pixels which are located above the horizon is stored with the image to facilitate the usage of the horizons location.

Labelling images

Adobe Photoshop CC (Adobe Systems Incorporated, San Jose, California, USA) was used to label the images. The labelling consist of selecting all pixels that show the field. The *masking* tool and the *fast selection* tool were used on a *RGB* version of the image to enable a semi-automated labelling on full resolution images. Smoothing of the edges was disabled and the mask was converted from one with transitioning edges to

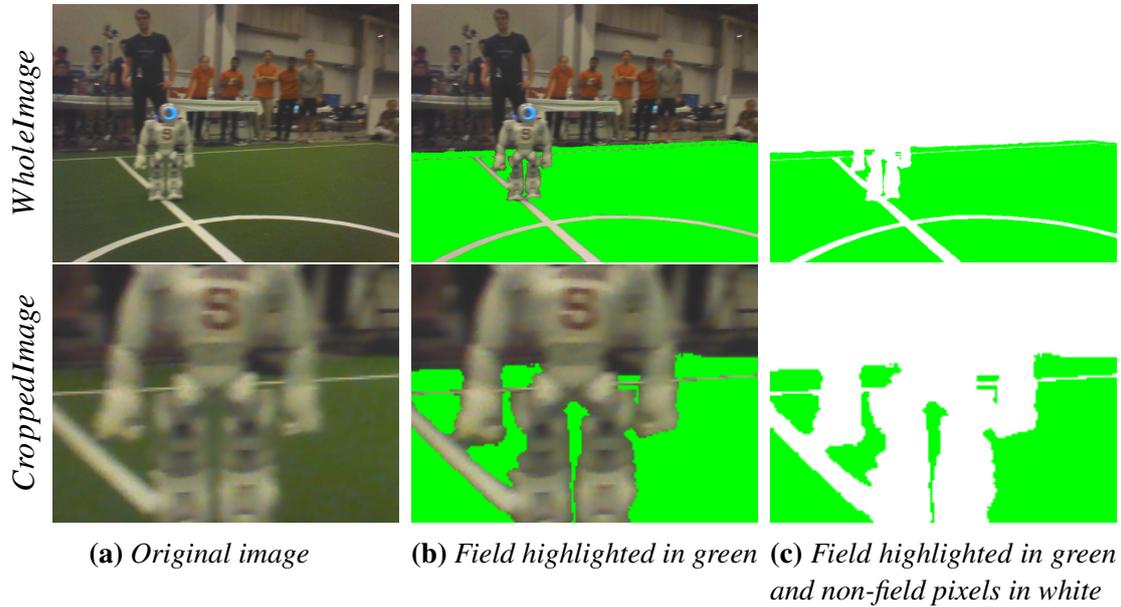


Figure 3.7: The original images next to two different representations of the labeled image in the first row and a difficult image segment due to blurring and small details in the same representation in the second row

a binary mask. Small details such as field marks in the distance or smaller parts of a NAO were labelled manually pixel by pixel. Figure 3.7 shows the busy background test image in *RGB* from the RoboCup 4.7a, the same image with field pixels highlighted in green 4.7b and the image with both field pixels highlighted in green and non field pixels represented in white 4.7c.

3.1.3 White Balance

As seen in figure 3.8, the white balance settings of the camera are not correct and vary from event to event. Parts of the images that are white or light grey, such as field marks and other NAOs, are tinted towards blue or yellow. With a constant lighting situation one could tune the white balance to the correct values before a match to counter this problem. However, this is not feasible as changing lighting conditions are encouraged in the RoboCup and thus the predominant light source can change from indoor lighting to sunlight during a match. Therefore, a simplified white balance algorithm is introduced.

The approach is similar to that approach described by Jiang et. al. that utilizes the coincidence of chromaticity histograms and their optimal relation to calculate white balance coefficients [23]. The linear diagonal transformation shown in equation 3.1 is



Figure 3.8: *Two images with varying white balance settings with figure 3.3e on the left having a blue tint and 3.4e on the right with a slight yellow tint.*

used to transform an original image to a corrected image. This represents an adjustment of the gain of each color channel. This transformation is based on the von Kries model and is computationally less expensive than other approaches [24].

$$\begin{pmatrix} R_{wb} \\ G_{wb} \\ B_{wb} \end{pmatrix} = \begin{pmatrix} k_R & 0 & 0 \\ 0 & k_G & 0 \\ 0 & 0 & k_B \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (3.1)$$

In the approach introduced in this study the calculation of the transformation matrix is simplified further. Most of the bright parts of the image seen by a NAO show neutral colors such as light grey NAOs and white field marks. Therefore, the maxima in the chromaticity histograms represent neutral areas in the image. As shown in equation 2.6, the sum of the three chromaticity values of a pixel is equal to one. The correct chromaticity values of a neutral pixel i.e. white, grey or black are identical and thus have a value of $1/3$. As this is also true for the maximum values in the histogram the transformation matrix can be calculated as shown in equation 3.2.

$$\begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} r_{bright,corrected} \\ g_{bright,corrected} \\ b_{bright,corrected} \end{pmatrix} = \begin{pmatrix} k_R & 0 & 0 \\ 0 & k_G & 0 \\ 0 & 0 & k_B \end{pmatrix} \begin{pmatrix} r_{bright,max} \\ g_{bright,max} \\ b_{bright,max} \end{pmatrix} \quad (3.2)$$

The corrected RGB_{wb} color space is computed using equation 3.1 and all other corrected color spaces are computed with the equations described in chapter 2.

3.1.4 Comparison Metrics

To evaluate how suitable different single channels are to classify pixels into field or non-field pixels, normalized histograms for both field pixels and non-field pixels are calculated in the single channel. The normalized histograms display the probability density function which is the number of pixels in each histogram bin divided by the total number of pixels. Dividing both histograms by the total number of pixels in the respective class leads to a balanced representation of the pixel distributions. This is crucial, as the number of field and non-field pixels in the images can differ by a wide margin. For example, the robot could see only field when looking down or the robot's view might be obstructed completely by another NAO in front of it so that no field pixels are seen.

The histogram intersection is defined as shown in equation 3.3 and describes how similar two histograms are by calculating the overlap between the two histograms [25]. F and N represent the normalized histograms of field pixels and non-field pixels respectively. The number of equally sized bins is given by n . If both histograms are identical, the intersection is 1 and if both histograms are mutually exclusive, i.e. any bin only contains pixels of one histogram and never pixels of both histograms, it is 0. In this application of the histogram intersection the value can be interpreted as follows: Supposed it is known which pixel class forms the majority in each histogram bin. An optimal classification would have to classify all pixels within this bin as the majority class which results in a miss-classification of all pixels in this bin belonging to the minority class. The intersection then correlates to the number of pixels that are miss-classified, despite an optimal classifier. There is no possible classifier that avoids this miss-classification. Therefore, the intersection is a good indicator of the maximum possible separability.

It is important to note however, that a low intersection of two histograms does not automatically mean that the two pixel distributions can be separated easily. For example, if the bins alternate between containing only pixels of F and only containing pixels of N , the intersection is 0 but it is very hard to separate the two. Therefore, the histograms of the combinations with the smallest intersections were verified visually.

$$\sum_{i=1}^n \min(F_i, N_i) \quad (3.3)$$

Similarly, the intersection of two-dimensional histograms is computed to compare different combinations of two channels as shown in equation 3.4. The number of bins per channel n should be constant and independent of the number of channels in a combination. This enables a sound comparison of combinations with different numbers of

channels.

$$\sum_{j=1}^n \sum_{i=1}^n \min(F_{i,j}, N_{i,j}) \quad (3.4)$$

The pixels of multiple test images are used together to form the histograms. Three histogram pairs are formed from the test images. These are F_{top} and N_{top} , F_{bottom} and N_{bottom} as well as F_{all} and N_{all} . Combining the pixels of multiple images leads to a better comparison due to the different lighting conditions present in the combined images and it also smoothes the histograms.

3.2 Evaluation

channel	all images	all images + wb
Y'	0.373	0.382
V	0.404	0.459
I	0.378	0.381
R	0.362	0.356
G	0.407	0.415
B	0.422	0.387

Table 3.1: Intersection values of channels from different colorspace that describe the brightness of a pixel. The intersection value provides an indication of the separability of field and non-field pixels by measuring the two distributions' overlap. An intersection of 0 would be best as this represents no overlap and a value of 1 would represent identical distributions with total overlap.

First, all single channels describing the brightness (Y , V , I , R , G , B) were evaluated in the uncorrected form and after adapting the white balance. The channels R , G and B were treated as channels representing the brightness of an image and not the chrominance, as they are strongly correlated to the brightness channels. Y is a weighted sum of RGB , V is the maximum and I the average as shown in chapter 2. The results are shown in table 3.1. The results show that the white balance correction does not significantly

affect the intersection value of the channels describing the brightness. The intersection in the white balanced R_{wb} channel is the lowest but not by a significant margin.

channel	all images	all images + wb
Cb	0.501	0.350
Cr	0.423	0.319
H_{YHS2}	0.280	0.179
S_{YHS2}	0.254	0.296
H	0.279	0.181
S	0.257	0.319
r	0.529	0.490
g	0.127	0.100
b	0.438	0.333

Table 3.2: Intersection values of channels from different colorspace that describe the color of a pixel. The intersection value provides an indication of the separability of field and non-field pixels by measuring the two distributions' overlap. An intersection of 0 would be best as this represents no overlap and a value of 1 would represent identical distributions with total overlap.

Second, all single channels describing the chrominance (Cb , Cr , H_{YHS2} , S_{YHS2} , r , g , b , H , S) were evaluated with the results shown in table 3.2. In contrast to the brightness channels, these results show that the intersection of all channels describing the chrominance, except for the saturation channels, decreases significantly due to the white balance correction. The intersection in the g_{wb} channel is the lowest by a large margin. Figure 3.9 shows the histograms in Cb on the left which has a high intersection of 0.501 and the histograms in g_{wb} on the right with a low intersection. This illustrates the difference of a high intersection visible by the large overlapping area on the left and a low intersection with minimal overlap on the right.

Next, the combinations of two channels were evaluated using the intersection formula for two dimensional diagrams as shown in equation 3.4. Table 3.3 shows selected combinations, sorted by the intersection value. Combination 2 is identical to combination 1 only without the white balance correction of V . The combinations ranked from 4 to 9 are not shown as they are other combinations of g_{wb} and different brightness

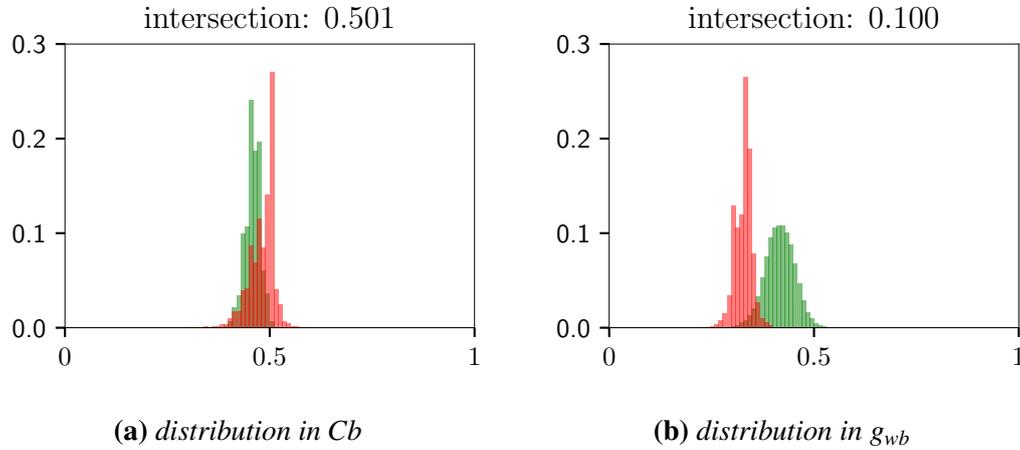


Figure 3.9: Normalized distributions of field pixels (green) and non-field pixels (red) in two different channels. On the left the distribution in Cb is shown with a large overlap area representing a high intersection value. The distribution in g_{wb} is shown on the right with a small overlap area and thus a low intersection value.

channels. Hence, the first combination which is not a combination of g_{wb} and a brightness channel is ranked 10 and is a combination of two brightness channels. The table continues with g_{wb} - brightness and brightness - brightness combination until rank 29. The table continues with a selection relevant combinations. The combination of g_{wb} - g_{wb} is equivalent to the one-dimensional intersection of g_{wb} and thus all combinations of channels below rank 110 are worse than the single channel g_{wb} . This includes all combinations utilized by other teams for the field color detection. Further information about the combinations is given in the table's "comment" column.

As the intersection alone does not necessarily mean that the pixel distributions of field and non-field pixels can be separated easily, two-dimensional histograms of the combinations 1 ($V_{wb} - g_{wb}$) and 3 ($I - g_{wb}$) are evaluated visually as seen in figure 3.10. One can see that the combination $I - g_{wb}$ on the right represents a more even two-dimensional distribution than the combination $V_{wb} - g_{wb}$. For both combinations there is almost no overlap visible.

rank	channel 1	channel 2	intersection	comment
1	g_{wb}	V_{wb}	0.037600	first pick (best combination)
2	g_{wb}	V	0.037925	identical to 1 except white balance
3	g_{wb}	I	0.038980	second pick
:	:	:	:	g_{wb} - brightness combinations
10	G_{wb}	I	0.041650	brightness - brightness combinations
:	:	:	:	only g_{wb} / g and brightness channels
29	H_{wb}	R_{wb}	0.053950	first appearance of hue
48	$H_{YHS2,wb}$	$S_{2YHS2,wb}$	0.071286	B-Human approach + wb
83	H_{wb}	S_{wb}	0.085843	standard hue and saturation + wb
110	g_{wb}	g_{wb}	0.099676	equivalent to single channel g_{wb}
118	H	S	0.102233	standard hue and saturation
128	H_{YHS2}	S_{2YHS2}	0.104507	B-Human approach
190	g	g	0.126592	equivalent to single channel g
261	Cb_{wb}	Cr_{wb}	0.162197	HULKs approach + wb
400	Cb	Cr	0.255659	HULKs approach + wb
465	r	r	0.528614	worst and last combination

Table 3.3: Comparison of intersection values of selected two-channel combinations. The intersection value provides an indication of the separability of field and non-field pixels by measuring the two distributions' overlap. An intersection of 0 would be best as this represents no overlap and a value of 1 would represent identical distributions with total overlap. Only the combinations with the lowest intersection and selection of other relevant combinations is shown. The comment column provides more information about the combinations.

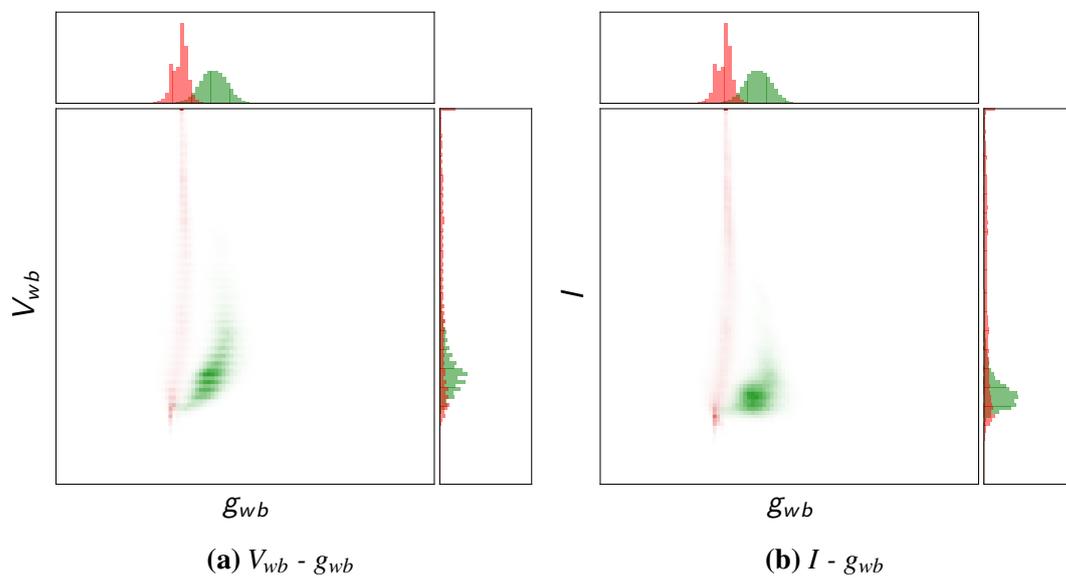


Figure 3.10: Comparison of the two two-dimensional combinations $V_{wb} - g_{wb}$ on the left and $I - g_{wb}$ on the right. The corresponding one-dimensional distributions are displayed alongside the two-dimensional histogram. The distributions on the left have an intersection value of 0.037600 and the combination on the right leads to a slightly higher intersection value of 0.037925.

3.3 Discussion

The green chromaticity g is the the single channel which is suited best for a classification algorithm. After white balance correction this channel alone represents a better color representation regarding separability of field and non-field pixels than all other approaches from other teams introduced in chapter 2. The white balance correction also improves the color representations of approaches from other teams significantly.

It should be noted however, that a better white balance could be achieved by adjusting the camera system parameters before a match starts. Nonetheless, the performance of the adaptive white balance correction algorithm introduced in this chapter demonstrates the ability to work under changing lighting conditions.

The visual comparison of the two combinations shown in figure 3.10 resulted in a preference for the pair g_{wb} and I (rank 3) despite the slightly higher intersection value of this combination. Therefore, this combination was considered in the next chapter to draft a classification algorithm.

Chapter 4

Classification into Field and Non-Field Pixels

In this chapter different classification algorithms using the channels g_{wb} and I_{wb} were explored and compared. The goal is to classify pixels into field pixels and non-field pixels under the constraint of small computational complexity.

4.1 Material and Methods

4.1.1 Classification Using Thresholds

First, a simple classifier was computed only from the g_{wb} channel, as this is the channel with the smallest intersection value and thus the most suitable channel for a classification using only a single channel. To select a threshold, the cumulative normalized histograms of field and non-field pixels and their difference were calculated. The histograms and the cumulative histograms of g_{wb} are shown in figure 4.1 with the black line describing the difference of the field pixel histogram and the non-field pixel histogram. The vertical dotted grey line marks the position of the maximum distance on the g_{wb} axis. Here, the threshold g_{min} to separate field and non-field pixels was defined as the position of the maximum difference on the g_{wb} axis. Therefore, all pixels with g_{wb} values above the threshold g_{min} were classified as field pixels while all other pixels were classified as non-field pixels.

Similarly, the difference of the two cumulative histograms of the I channel was calculated. Figure 4.2 shows the same plots as described above for I . In this channel a single threshold is not sufficient to separate the field pixels from non-field pixels, as there is a group of non-field pixels with a lower intensity and also a group with pixels with a higher intensity than the field pixels. Pixels with a lower intensity display dark

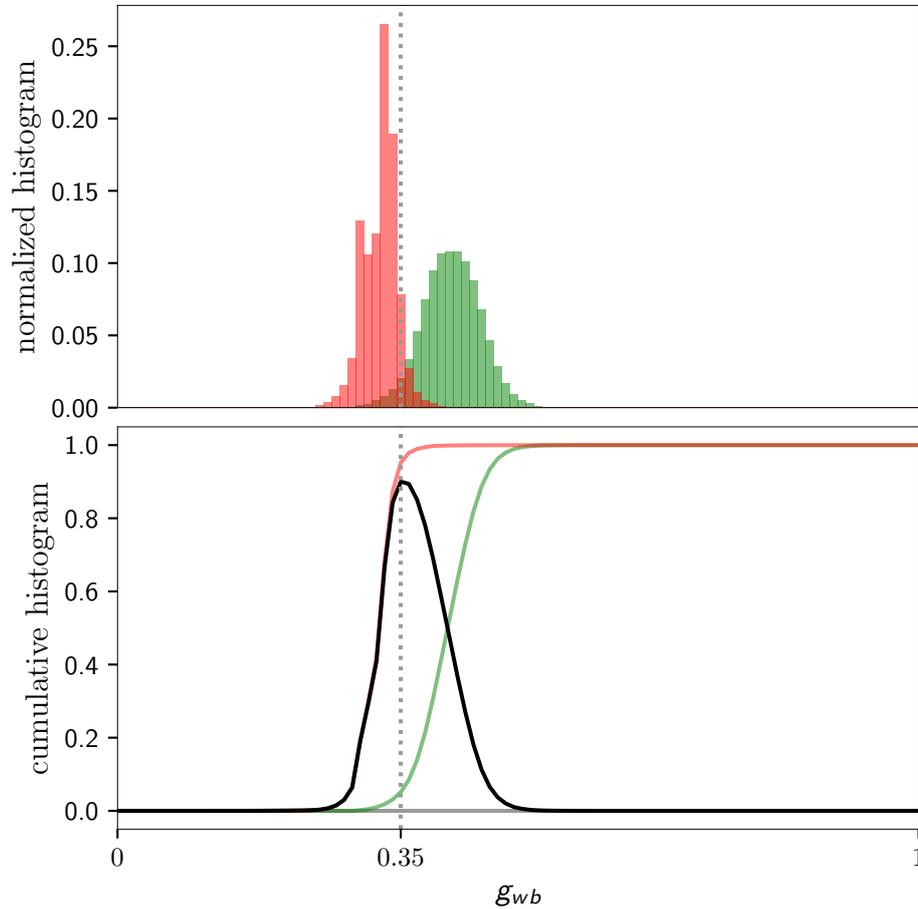


Figure 4.1: Normalized field and non-field pixel distributions in g_{wb} channel and the corresponding cumulative histograms with the optimal threshold marked by dotted gray lines.

parts of the background or black patches on the ball. Lighter pixels show field marks, white parts of the ball, other NAOs, goals, and other bright objects in the background. Hence, two thresholds are calculated with I_{min} located at the position of minimum (negative) difference and I_{min} located at the position of the maximum as shown in figure 4.2. All pixels with intensities above I_{min} and below I_{min} are classified as field pixels.

Next, both of the channels g_{wb} and I were combined. This forms a rectangle with the pixels within this rectangle classified as field pixels. Figure 4.3 shows the cumulative histograms of both channels together with the two dimensional histogram. The three thresholds are marked as dotted grey lines and form a rectangle enclosing the cluster of

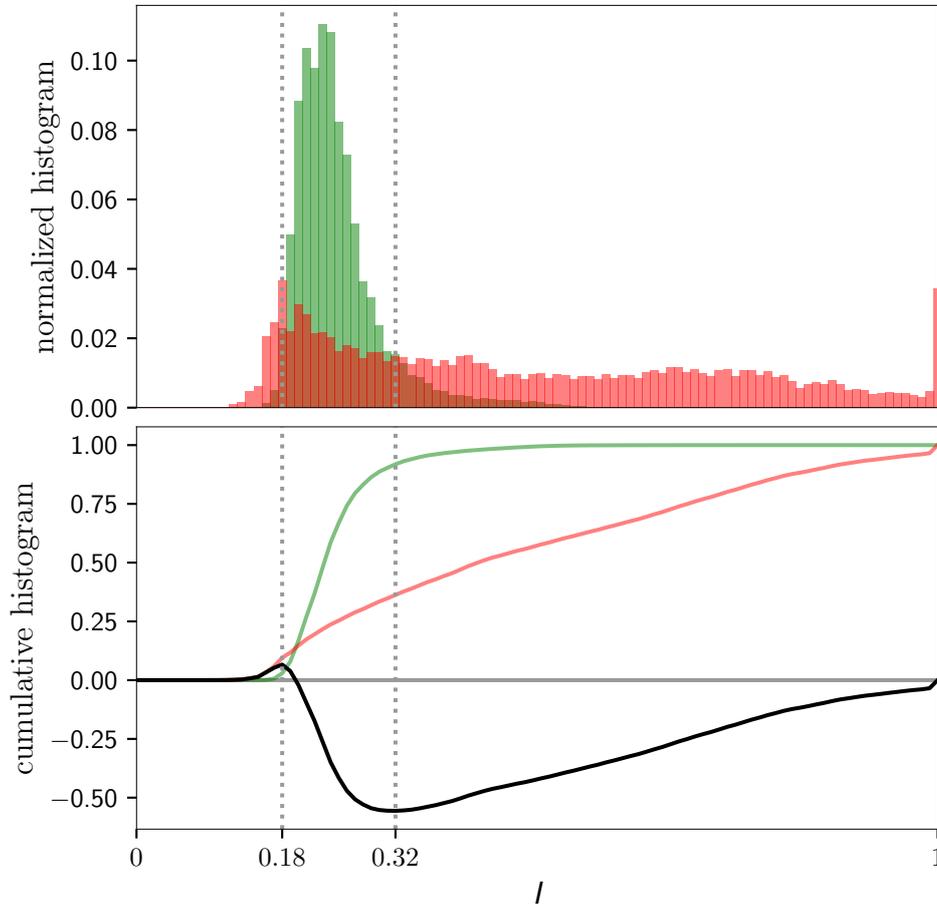


Figure 4.2: Normalized field and non-field pixel distributions in I channel and the corresponding cumulative histograms with both optimal minimum (left) and optimal maximum (right) threshold marked by dotted gray lines.

field pixels.

4.1.2 Classification Metrics

Each classifier returns an array with a binary prediction for every pixel. *True* represents field and *False* represents non-field. As the images or sets of images do not contain a balanced number of pixels in both classes, the balanced accuracy is used as a metric to evaluate the classifiers and combinations of them. A confusion matrix was calculated for selected combinations of classifiers and image sets to verify the validity of the bal-

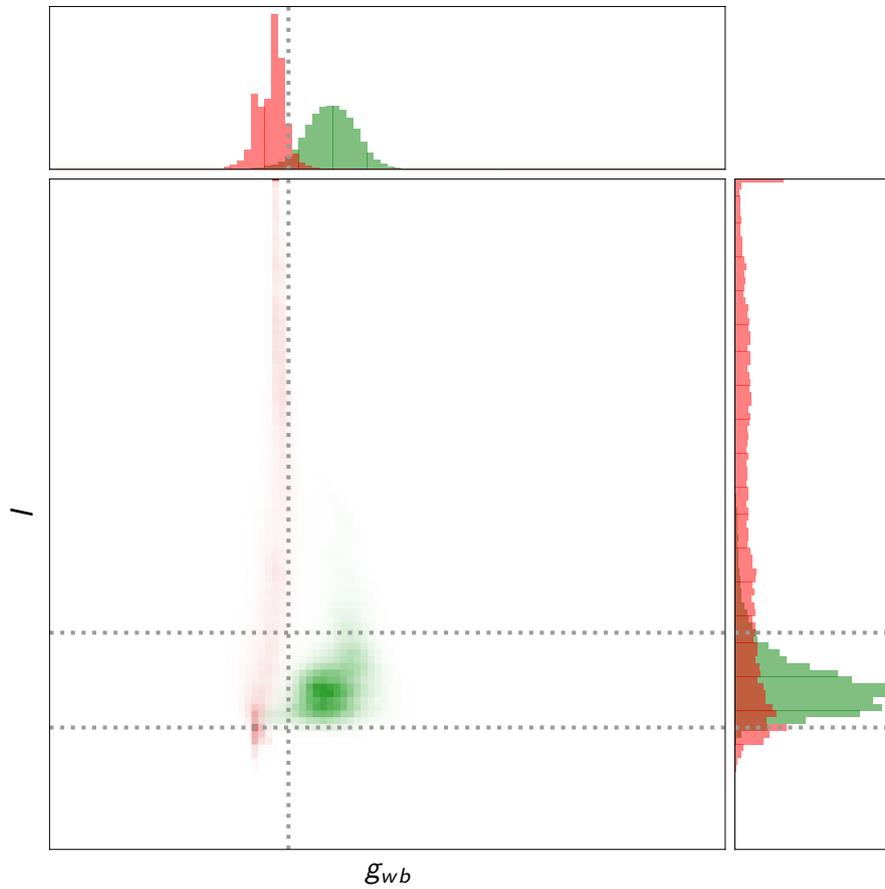


Figure 4.3: *Two-dimensional normalized pixel distribution of field and non-field pixels in g_{wb} and I_{wb} with the corresponding one-dimensional histograms of each dimension. Optimal classification thresholds for both channels are marked by dotted gray lines.*

anced accuracy. In addition, the classification results were compared visually.

4.2 Evaluation

Image	$g_{min} < g$	$I_{min} < I < I_{max}$	$g_{min} < g \wedge I_{min} < I < I_{max}$
GO-BOTTOM-BALL	0.836	0.891	0.925
GO-BOTTOM-LINES	0.886	0.991	0.937
GO-BOTTOM-PENALTY	0.862	0.955	0.924
GO-TOP-BACKGROUND	0.938	0.683	0.968
GO-TOP-NAO	0.880	0.796	0.897
IO-BOTTOM-BALL	0.942	0.971	0.988
IO-BOTTOM-LINES	0.917	0.975	0.969
IO-BOTTOM-PENALTY	0.929	0.997	0.993
IO-TOP-BACKGROUND	0.966	0.737	0.956
IO-TOP-NAO	0.901	0.777	0.922
RC-BOTTOM-BALL	0.852	0.940	0.981
RC-BOTTOM-LINES	0.902	0.996	0.995
RC-BOTTOM-PENALTY	0.954	0.977	0.991
RC-TOP-BACKGROUND	0.933	0.780	0.968
RC-TOP-NAO	0.942	0.753	0.950
HH-BOTTOM-BALL	0.961	0.826	0.865
HH-BOTTOM-LINES	0.993	0.723	0.723
HH-BOTTOM-PENALTY	0.992	0.976	0.976
HH-TOP-BACKGROUND	0.929	0.903	0.985
HH-TOP-NAO	0.949	0.664	0.672

Table 4.1: *Balanced accuracy of the prediction results of all all classifiers (columns) for all images (rows) are shown in this table. A value of 1 would represent a perfect classification. The subscript $_{wb}$ was omitted for a more concise table.*

The balanced accuracy for all combinations of single images and classifiers is shown in table 4.1. The balanced accuracy of thresholding in g_{wb} results in a high balanced accuracy between 0.852 and 0.993 with a low variation. In contrast, thresholding in the I channel leads to a higher variation and to low accuracy values in images with strong illumination gradients. This classifier performs significantly better at frames from the bottom camera. The combination of both simple classifiers does not always lead to better results than any of the single classifiers alone. In the following three figures images with varying classification performances for the different classifiers are shown. Figure 4.4 shows the image with the lowest balanced accuracy score for the $g_{min} < g_{wb}$ classifier. Due to the combination with the thresholding in the intensity the balanced accuracy can be improved to 0.925. The image with the lowest balanced accuracy score for the $I_{min} < I < I_{max}$ is shown in figure 4.5. Due to the strong illumination gradient present on the field this classifier fails here. However, the threshold in g_{wb} results in a very good classification. Finally the classification prediction of the penalty spot image taken at the IranOpen are shown in figure 4.6. This is an example of the images where both classifiers result in a good result.

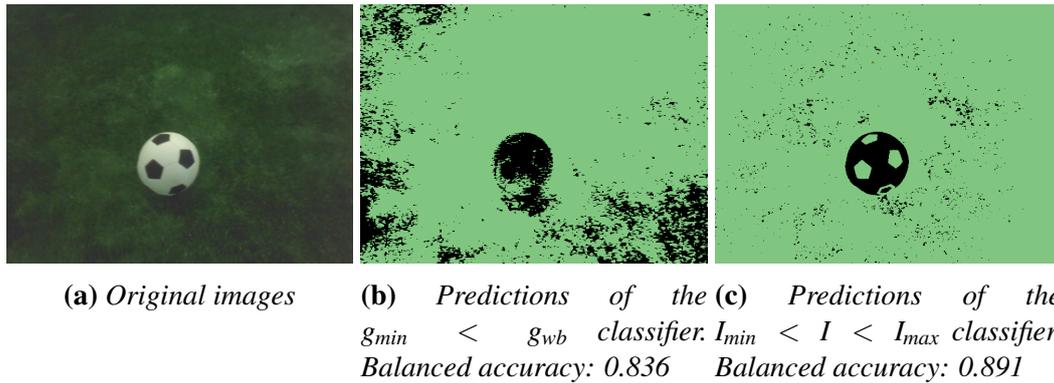


Figure 4.4: Classification predictions for the ball test image taken at the GermanOpen 2018. The g_{wb} thresholding does not result in satisfying results with a balanced accuracy of 0.836. Thresholding of the intensity results in a balanced accuracy of 0.891. Combining both classifiers however results in a balanced accuracy of 0.925.

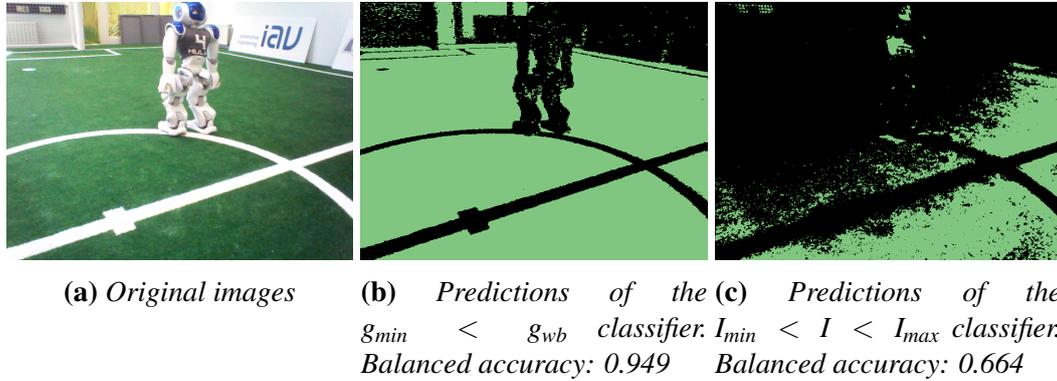


Figure 4.5: Classification predictions for the test image showing another NAO robot taken at HULKs laboratory with bright sunshine outside. The g_{wb} thresholding results in a near perfect classification except for the yellow object in the background with a balanced accuracy of 0.949. However, thresholding of the intensity fails completely and results in a poor balanced accuracy of 0.664. Combining both classifiers results in a balanced low balanced accuracy of 0.672.

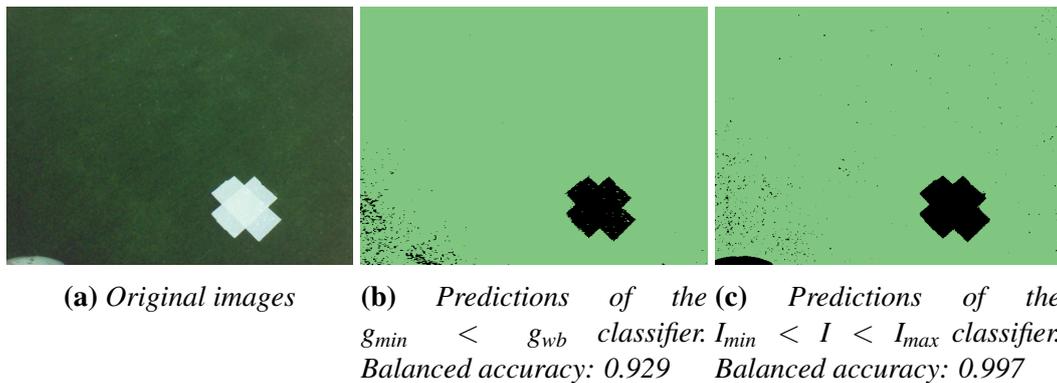


Figure 4.6: Classification predictions for the penalty spot test image taken at the IranOpen 2018. Both classifiers result in good balanced accuracy scores of 0.929 and 0.997. The combination of both classifiers also leads to a good balanced accuracy of 0.993.

4.3 Discussion

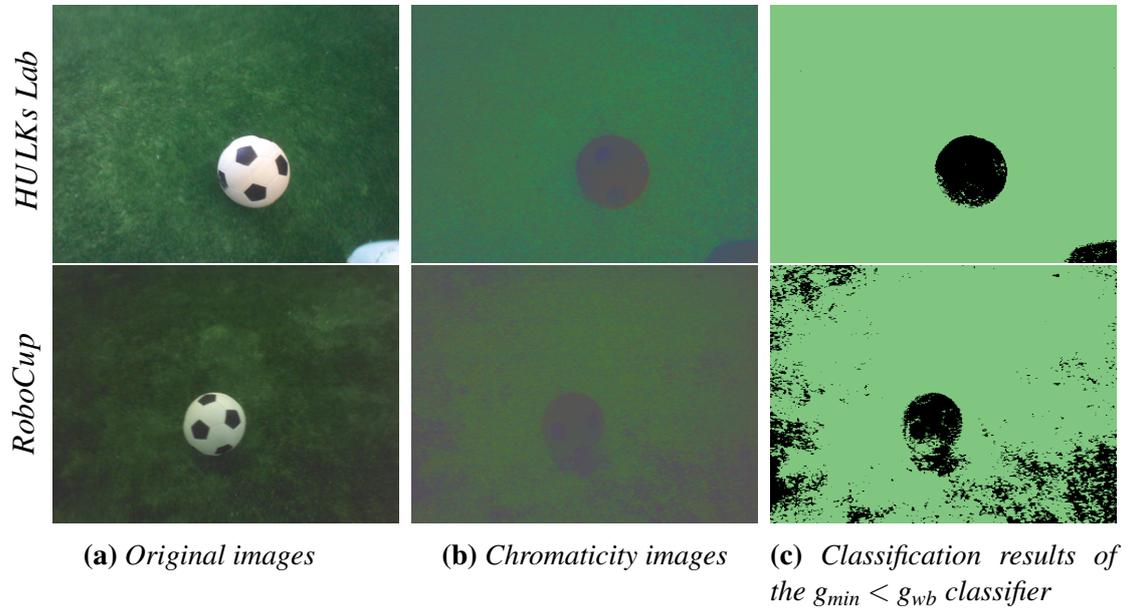


Figure 4.7: Illustration of lower classification performance at very dark and thus colorless pixels. The first row shows the ball test image taken at the HULKs test lab where the brightness varies but all field pixels still contain color information. The second row shows the ball test image taken at the RoboCup which contains field pixels that do not contain color information and are therefore miss-classified. The difference is apparent in the center column showing the chromaticity images for both test images.

A single threshold in the green chromaticity channel after white balancing g_{wb} results in a surprisingly good classification with a balanced accuracy between 0.852 and 0.993 and little variation on four different fields with four different lighting conditions. This is due to the fact that most pixels which do not show the field are neutrally colored i.e. black, gray or white. This classification's limit are pixels which are not bright enough to hold meaningful color information. This however, is a hard limit for single pixel based classification. If a pixel's R , G , and B values are equal or close to identical and the intensity of this pixel is low i.e. the pixel appears to be dark gray or black, there is no possibility to distinguish between a dark non-field part such as the black ball patches and the field. This is illustrated in figure 4.7 where two test images are compared regarding their chromaticity. The top row shows the ball test image taken at the HULKs Lab in Hamburg and the bottom row the ball test image from the RoboCup. The left column displays the images in RGB , the second column shows chromaticity images and the third one the classification result of the $g_{min} < g_{wb}$ classifier. The chromaticity images treat the rgb values as RGB so that all neutral areas are gray. The classification

of the image in row two fails at the dark gray pixels in the image as there is no color information as seen in the gray areas in the chromaticity image.

Chapter 5

Discussion

Both the comparison of different color channels and the classification results demonstrate the significance of the green chromaticity. This is not particularly surprising, as the green chromaticity is basically a measure of how green the pixel is. It is important to note that this is highly domain specific for soccer fields, due to their color. However, it is very important to point out that the single classification in g_{wb} outperforms all combinations of two channels used by other teams as seen in table 3.3 on this image set. The main reason for poor classification are pixel that contain no color information. However, this is a limitation of all classifiers that work on a single pixel basis which includes all approaches mentioned in chapter 2 except NAO Team HTWKs machine learning approach which operates on a local neighborhood of the pixel. As the computational power of the NAO is highly limited, a simplified white balance correction plus a single computationally inexpensive thresholding is superior regarding computing power.

In addition, algorithms using only color values of a single pixel have another advantage. The algorithm can be applied to the whole image or just a subset of the pixels. This is advantageous when a subsampled image, i.e. only every n^{th} pixel in each row and each column, or scan lines i.e. vertical or horizontal lines of pixels, are used. As the pixel's classification is not dependent on its vicinity it is possible to classify any pixel at any time. However, to calculate the white balance correction factors a histogram is needed and thus a representative number of pixels needs to be evaluated for every image first.

Due to the good classification results in combination with the low computational cost this approach can improve the results of other vision modules in the HULKs framework such as the detection of the field border. A robust field color detection also produces clear boundaries of objects on the field such as other NAOs which can facilitate the detection of these object. The current vision modules in the HULKs framework are all based on the $Y'CbCr$ colorspace so that the pixels have to be transformed to RGB first.

This can increase the runtime on the NAO. However, it is also possible to receive the image data directly in *RGB* from the camera. As the sixth version of the NAO will already be used in the upcoming events in 2019, an exact evaluation of the runtime of the proposed algorithms on the then outdated hardware was not valuable for this study.

Some areas in an image seen by a NAO show bright colors. If these colors are a mixture of green and red or green and blue they might be classified as field pixels due to their high g_{wb} value. These areas are mostly colorful jerseys of other NAOs, colorful clothing of spectators at events and lights from the LEDs of other NAOs. These areas can be excluded with simple thresholds in r and b as these colors are very distinct from green in these channels. Additionally, pixels that lie over the horizon and oversaturated pixels can be classified easily as non-field pixels. These additional "add-on" classifiers were not considered in the evaluation, as they do not give a general classification into field and non-field but only classify specific areas as non-field.

Chapter 6

Conclusion and Future Work

The aim of this thesis was to develop an algorithm that classifies pixels into either showing a part of the field or not showing a part of the field. This algorithm had to be robust against challenging lighting conditions while being computationally inexpensive. This was achieved by the comparison of different color representations used in related works regarding their ability to separate pixels into these two classes. The different channels of the color representations were compared systematically using the histogram intersection as a metric. It was shown that the green chromaticity channel of an image is the best single color channel for this classification as it is highly illumination invariant. A simplified approach to correct the images' white balance was introduced and it was shown that the white balance correction leads to a significantly better separability. Also, combinations of color channels were explored and it was found that the green chromaticity channel in combination with a representation of the pixel's brightness further improves separability. This pixel representation outperforms all other representations used by other RoboCup teams regarding separability. Based on these results, simple thresholding classifiers were introduced and their classification performance was measured against ground truth images displaying different game situations from multiple locations and under varying lighting conditions. It was shown that a simple thresholding in the white balance corrected green chromaticity channel alone results in good classification with a balanced accuracy between 0.852 and 0.993 with low variation in between different test images. Weaker classification results were analyzed and adjustments in the camera settings were proposed to eliminate the main cause for the miss-classifications.

As soon as the new NAO's API is available and a variety of pictures from different situations, fields and illumination variants are recorded, the algorithm developed in this thesis has to be tested on this new data. Furthermore, the performance of the developed algorithm has to be tested on the new hardware.

The limitation of the developed approach are underexposed areas of the field which

result into very low and noisy color information. This renders the distinction between these field pixels and for example the dark ball patches impossible for algorithms that work only on a single pixel. To overcome this problem the exposure and gain settings of the NAO's cameras should be adapted. It is possible to apply different exposure times to different image areas as the B-Human team already does. In addition, algorithms which work on the local neighborhood of a pixel might be explored to improve the classification results.

The simplified white balance correction introduced in this thesis calculates new correction factors for every image. The accuracy of this algorithm has to be improved by only updating the correction factors for images which contain a sufficient amount of white pixels. The correction factors should be adapted gradually when a new illumination situation is detected to avoid jumps. In addition, more emphasis should be put on the correct white balance settings of the camera system at the start of a match, to reduce the importance of the white balance correction described in this thesis.

In conclusion, the developed approach can improve the field detection significantly and outperforms other teams' field detection approaches. Finally, this will benefit other vision algorithms in the HULKS framework to improve the teams performance in the RoboCup.

Bibliography

- [1] Aldebaran Robotics. Nao - video camera. [Online]. Available: http://doc.aldebaran.com/2-1/family/robots/video_robot.html
- [2] Hsu and Feng-hsiung, "Ibm's deep blue chess grandmaster chips," *IEEE Micro*, vol. 19, no. 2, pp. 70–81, 1999.
- [3] Jet Propulsion Laboratory. (1997) A description of the rover sojourner. [Online]. Available: https://www.robocup.org//a_brief_history_of_robocup
- [4] HULKs. Hamburg ultra legendary kickers. [Online]. Available: <https://www.hulks.de>
- [5] Softbank Robotics Corp. Softbank robotics corp. [Online]. Available: <https://www.softbankrobotics.com/emea/en/index>
- [6] RoboCup Federation. (2016) A brief history of robocup. [Online]. Available: https://www.robocup.org//a_brief_history_of_robocup
- [7] RoboCup Technical Committee. (2018) Robocup standard platform league (nao) rule book. [Online]. Available: [http://spl.robocup.org/wp-content/uploads/downloads/2019/01/RoboCup-Standard-Platform-League-\(Nao\)-Rule-Book-2018.pdf](http://spl.robocup.org/wp-content/uploads/downloads/2019/01/RoboCup-Standard-Platform-League-(Nao)-Rule-Book-2018.pdf)
- [8] International Telecommunication Union. (2011) Recommendation itu-r bt.601-7. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.601-7-201103-I!!PDF-E.pdf
- [9] Semiconductor Components Industries, LLC. (2016) Mt9m114. [Online]. Available: <https://www.onsemi.com/pub/Collateral/MT9M114-D.PDF>
- [10] A. Fürtig, H. Friedrich, and R. Mester. (2010) Robust pixel classification for robocup. [Online]. Available: https://www.zbs-ilmeneau.de/fws2010/presentation/Fuertig_farbworkshop.pdf

- [11] Berlin United - Nao Team Humboldt. (2018) Team report 2017. [Online]. Available: <http://www.naoteamhumboldt.de/wp-content/papercite-data/pdf/naoth-report17.pdf>
- [12] Nao-Team HTWK. (2018) Team research report. [Online]. Available: http://naoteam.imn.htwk-leipzig.de/documents/TRR_2017.pdf?lang=en
- [13] D. Adikari, G. Felbinger, P. Gleske, C. Kahlefeldt, Y. Konda, R. Kost, P. Loth, K. Nölle, L. Peters, N. Riebesel, T. Schattschneider, M. Schmidt, E. Schröder, F. Warmuth, and F. Wege. (2018) Hulks team research report 2018. [Online]. Available: https://hulks.de/_files/TRR_2018.pdf
- [14] T. Röfer, T. Laue, Y. Bülter, D. Krause, J. Kuball, A. Mühlenbrock, B. Poppinga, M. Prinzler, L. Post, E. Roehrig, R. Schröder, and F. Thielke. (2017) Team report and code release 2017. [Online]. Available: <https://www.b-human.de/downloads/publications/2017/coderelease2017.pdf>
- [15] B-Human. (2018) The official b-human code releases 2018. [Online]. Available: <https://github.com/bhuman/BHumanCodeRelease>
- [16] Y. Qian and D. D. Lee, “Adaptive field detection and localization in robot soccer,” in *RoboCup 2016: Robot World Cup XX Lecture Notes in Computer Science, vol 9776*, S. Behnke, R. Sheh, S. Sariel, and D. D. Lee, Eds. Springer, Cham, 2017, pp. 218–229.
- [17] A. Koschan and M. Abidi, *Digital Color*. Wiley, 2008.
- [18] Aldebaran Robotics. Nao documentation. [Online]. Available: http://doc.aldebaran.com/2-1/home_nao.html
- [19] Softbank Robotics. Nao documentation. [Online]. Available: http://doc.aldebaran.com/2-8/home_nao.html
- [20] OmniVision Technologies. (2011) Ov5640 datasheet. [Online]. Available: https://cdn.sparkfun.com/datasheets/Sensors/LightImaging/OV5640_datasheet.pdf
- [21] Ecma International. (2017) The json data interchange syntax. [Online]. Available: <http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-404.pdf>
- [22] N. Riebesel. (2018) Center circle detection on the nao robotic platform for the robocup standard platform league. [Online]. Available: https://hulks.de/_files/PA_Nicolas-Riebesel.pdf

-
- [23] T. Jiang, D. Nguyen, and K.-D. Kuhnert, "Auto white balance using the coincidence of chromaticity histograms," in *2012 Eighth International Conference on Signal Image Technology and Internet Based Systems*. IEEE, 2012, pp. 201–208.
- [24] A. Gijsenij, T. Gevers, J. Van De Weijer *et al.*, "Computational color constancy: Survey and experiments," *IEEE Transactions on Image Processing*, vol. 20, no. 9, pp. 2475–2489, 2011.
- [25] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, Nov 1991. [Online]. Available: <https://doi.org/10.1007/BF00130487>